2003 International Symposium on Advanced Intelligent Systems
September 25-28, 2003, Jeju, Korea

Feature extraction with distance measures and fuzzy entropy

Sang-Hyuk Lee, Sung-Shin Kim, Hyeon Bae and Youn-Tae Kim

School of Electrical Engineering, Pusan National University
30 Jangjeon-dong, Geumjeong-gu, Busan 609-735, Korea
E-mail : leehyuk@pusan.ac.kr, Fax: +82-51-513-0212, Tel: +82-51-510-2497

Abstract - Representation and quantification of fuzziness are required for the uncertain system modelling and controller design. Conventional results show that entropy of fuzzy sets represent the fuzziness of fuzzy sets. In this literature, the relations of fuzzy enropy, distance measure and similarity measure are discussed, and distance measure is proposed. With the help of relations of fuzzy enropy, distance measure and similarity measure, fuzzy entropy is proposed by the distance measure. Finally, proposed entropy is applied to measure the fault signal of induction machine.

I. Introduction

Characterization and quantification of fuzziness are important issues that affect the management of uncertainty in many system models and designs. The results that entropy of a fuzzy set is a measure of fuzziness of the fuzzy set are known by the previous researchers[1-7]. Liu had proposed the axiomatic definitions of entropy, distance measure and similarity measure, and discussed the relations between these three concepts. Kosko viewed the relation between distance measure and fuzzy entropy. Bhandari and Pal gave a fuzzy information measure for discrimination of a fuzzy set A relative to some other fuzzy set B. Pal and Pal analyzed the classical Shannon information entropy. Also Ghosh used this entropy to neural network.

In this paper, we derived entropy with distance measure. Fuzzy entropy is illustrated in Theorem 3.1 by the arbitrary distance measure. By the proof of that theorem in [7], proposed entropy represents the different structure of Fan and Xie[6]. Furthermore, we apply this fuzzy entropy to the faulted induction machine. Feature extracting from the faulted induction machine is interesting area to the electrical and mechanical engineers[8-12]. Generally, major faults of induction machines can be classified as follows[8]:

- Bearing fault
- Rotor Bar fault
- Stator or amature fault
- Eccentricity(static, dynamic).

These faults produce one or more of the symptoms as given below:

- Unbalanced air-gap voltages and line currents
- Increased torque Pulsations
- Decreased average torque
- Increased losess and reduction in efficiency
- Excessive heating.

The above faults can be identified by the following diagnostic methods:

- Motor current
- Vibration monitoring
- Temperature measurements
- Infrared recognition
- Chemical analysis
- Radio frequency emission monitoring *etc.*

Various approach to extract fault characteristic are studied with vibration, sound or other approaches[8]. Among these method, analyzing current signal is interesting to the electrical engineer. Hence, we have derived feature extraction with wavelet transform and fuzzy entropy. Usefulness of proposed method is verified through entropy computation.

In the next section, the axiomatic definitions of entropy, distance measure and similarity measure of fuzzy sets are introduced and some basic relations between these measures are discussed. In section 3, Entropy is induced by the distance measure. Also in section 4, characterization of fault signal is measured by the proposed entropy measure. Conclusions are followed in section 5.

Notations: Through out this paper, $R^+=[0,\infty)$, $F(X)$ and $P(X)$ represent the set of all fuzzy sets and crisp sets on the universal set X respectively. $\mu_A(x)$ is the membership function of $A\in F(X)$, and the fuzzy set A, we use A^c to express the complement of A, i.e., $\mu_{A^c}(x)=1-\mu_A(x),\ \forall x\in X$. For fuzzy sets A and B, $A\cup B$, the union of A and B is defined as $\mu_{A\cup B}(x)=\max(\mu_A(x),\mu_B(x))$, $A\cap B$, the intersection of A and B is defined as $\mu_{A\cap B}(x)=\min(\mu_A(x),\mu_B(x))$. A fuzzy set A^* is called a sharpening of A, if $\mu_{A^*}(x) \geq \mu_A(x)$ when $\mu_A(x) \geq \frac{1}{2}$ and $\mu_{A^*}(x) \leq \mu_A(x)$ when $\mu_A(x) < \frac{1}{2}$. For any crisp sets D, A_{near} and A_{far} of fuzzy set A are defined as

$$\mu_{(1/2)D}(x) = \begin{cases} \frac{1}{2}, & x \in D \\ 0, & x \in D^c \end{cases}$$

$$\mu_{A_{near}}(x) = \begin{cases} 1, & \mu_A(x) \geq \frac{1}{2} \\ 0, & \mu_A(x) < \frac{1}{2}, \end{cases}$$

$$\mu_{A_{far}}(x) = \begin{cases} 0, & \mu_A(x) \geq \frac{1}{2} \\ 1, & \mu_A(x) < \frac{1}{2}. \end{cases}$$

II. Fuzzy entropy, distance measure and similarity measure

In this section, we introduce some preliminary results and discuss induced results. Liu suggested three axiomatic definitions of fuzzy entropy, distance measure and similarity measure as follows. By these definitions, we can induce entropy, and compare it with the result of Liu.

Definition 2.1 (Liu, 1992). A real function e : $F(X) \rightarrow R^+$ is called an entropy on $F(X)$, if e has the following properties:

(EP1) $e(D) = 0, \forall D \in P(X)$;

(EP2) $e([\frac{1}{2}]) = \max_{A \in F(X)} e(A)$;

(EP3) $e(A^*) \leq e(A)$, for any sharpening A^* of A;

(EP4) $e(A) = e(A^c), \forall A \in F(X)$.

For $S(x)$ is $S(x) = -x \ln x - (1-x)\ln(1-x)$, $0 \leq x \leq 1$. One of entropies can be represented by

$$e(A) = -\sum_{i=1}^{n} S(\mu_A(x_i)), \quad \forall A \in F(X),$$

where $X = \{x_1, x_2, \cdots, x_n\}$.

Definition 2.2 (Liu, 1992). A real function d : $F^2 \rightarrow R^+$ is called a distance measure on F if d satisfies the following properties:

(DP1) $d(A,B) = d(B,A), \forall A,B \in F$;

(DP2) $d(A,A) = 0, \forall A \in F$;

(DP3) $d(D,D^c) = \max_{A,B \in F} d(A,B)$, $\forall D \in P(X)$;

(DP4) $\forall A,B,C \in F$, if $A \subset B \subset C$, then $d(A,B) \leq d(A,C)$ and $d(B,C) \leq d(A,C)$.

One of fuzzy distances takes the following form

$$d(A,B) = \left(\frac{1}{n}\sum_{i=1}^{n} |\mu_A(x_i) - \mu_B(x_i)|^p\right)^{\frac{1}{p}}, \quad p \geq 1.$$

Fuzzy normal distance measure on F is obtained by the multiplication of $1/\max_{C,D \in F} d(C,D)$.

Definition 2.3 (Liu, 1992). A real function s : $F^2 \rightarrow R^+$ is called a similarity measure, if s has the following properties:

(SP1) $s(A,B) = s(B,A), \forall A,B \in F$;

(SP2) $s(D,D^c) = 0, \forall D \in P(X)$;

(SP3) $s(C,C) = \max_{A,B \in F} s(A,B), \forall C \in F$;

(SP4) $\forall A,B,C \in F$, if $A \subset B \subset C$, then $s(A,B) \geq s(A,C)$ and $s(B,C) \geq d(A,C)$.

Liu also pointed that there is a one-to-one correlation between all distance measures and all similarity measures, $d + s = 1$. Fuzzy normal similarity measure on F is also obtained by the division of $\max_{C,D \in F} s(C,D)$.

If we divide X into two parts D and D^c in $P(X)$, then the fuzziness of fuzzy set A be the sum of the fuzziness of $A \cap D$ and $A \cap D^c$. By this idea, following definition is followed.

Definition 2.4. (Fan and Xie) Let e be an entropy on F. Then for any $A \in F$,

$$e(A) = e(A \cap D) + e(A \cap D^c)$$

is σ-entropy on F.

Definition 2.5. (Fan and Xie) Let d be a distance measure on F. Then for any $A,B \in F, D \in P(X)$,

$$d(A,B) = d(A \cap D, B \cap D) + d(A \cap D^c, B \cap D^c)$$

be the σ-distance measure on F.

Definition 2.6. (Fan and Xie) Let s be a similarity measure on F. Then for any $A,B \in F, D \in P(X)$,

$$s(A,B) = s(A \cap D, B \cup D^c) + s(A \cap D^c, B \cup D)$$

be the σ-similarity measure on F.

Fan also defined an entropy which is defined as $e' = e/(2-e)$, where e is an entropy on $F(X)$. Entropy profile of $e' = e/(2-e)$ is illustrated in Fig. 1. To discriminate between entropies, we give another entropy using Fan's idea.

Theorem 2.1 If e is an entropy on $F(X)$, then $\hat{e} = e^k$ is also an entropy on $F(X)$, where $k \geq 1$.
Proof. It is founded in [7].

If entropy has the structure of Thm. 2.1, it can discriminate entropies at near one from the other area.

III. Fuzzy entropy induced by distance measure

In this section, we propose entropy that is induced by the distance measure. Among distance measures, Hamming distance is commonly used σ-distance measure between fuzzy sets A and B,

$$d_H(A, B) = \frac{1}{n} \sum_{i=1}^{n} |\mu_A(x_i) - \mu_B(x_i)|$$

where $X = \{x_1, x_2, \cdots, x_n\}$.
Next Proposition shows that the distance relation of between fuzzy set and crisp sets.

Proposition 3.1 (Fan and Xie). Let d be a σ-distance measure on $F(X)$: then

(1) $d(A, A_{near}) \geq d(A^*, A_{near})$

(1) $d(A, A_{far}) \leq d(A^*, A_{far})$

In the next Theorem, we propose fuzzy entropy induced by distance measure.

Theorem 3.1 Let d be a σ-distance measure on $F(X)$; if d satisfies

$$d(A^c, B^c) = d(A, B), \quad A, B \in F,$$

then $e(A) = 2d(A, A_{near}) - d((A \cap A_{far}), [0])$ is a fuzzy entropy.

Proof. Proof is also shown in [7].

IV. Illustrative Examples

In this section, we have considered the extraction of characteristics from the faulted motor with stator current. Data from the induction machine, 220V, 3450 rpm, 4 pole, 24 bar, 0.5 HP motor have been used to verify the results experimentally. Six cases of bearing fault, bowed rotor, broken rotor bar static eccentricity, dynamic eccentricity and healthy conditions are given. 3-phase motor has the sensitivity about 10mV/A, 100mV/A and 100mV/A, respectively.

Healthy and faulted current signal at full load is illustrated in Fig. 1. Input signal have 16,384 data points, respectively. Maximum frequency represents 3 kHz, data duration is 2.1333 s. As can be seen in figure, it is not easy to extract characteristic from the time series signals. Furthermore, it is also needed to be synchronized to process with the frequency approach. After preprocessing of 6 cases time signals is carried out, FFT(Fast Fourier Transform) is processed. Those results are also illustrated in Fig. 2. We briefly show four cases of Fourier transformation. Except broken rotor bar signal, other signals can not be distinguishable. Hence we conclude temporary that both of time and frequency analysis are not adequate to extract feature from the faulted signals. Therefore another approach is required to detect feature of faulted signals. Wavelet transformed had been used in [10], however only rotor bar broken case was considered.

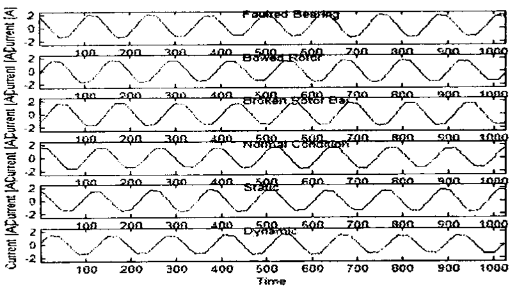
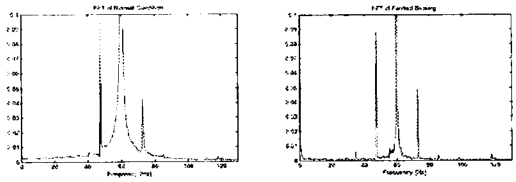
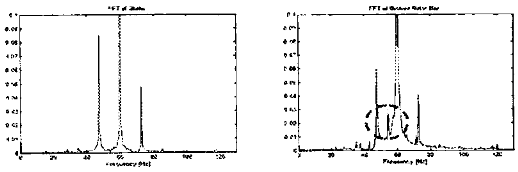


Fig. 1. Faulted and healthy current signals



(a)Healthy (b) Bearing fault



(c) Static eccentricity (d) Broken rotor bar

Fig. 2. FFT of healthy and faulted signals

Hence another method is required to extract the characteristic of faults, respectively. To extract feature of various cases, we consider the Wavelet transformation. This method give us time and frequency informations simultaneously.

The continuous wavelet transform $F(a, b)$ of a function f is defined by

$$F(a, b) = (f, \psi_{a,b}) = \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} f(t) \psi((t-b)/a) dt$$

where ψ is called the Mother wavelet.
Also in this paper, Coiflet Mother wavelet function is used. In Fig. 3, we represent Coiflet wavelet function.

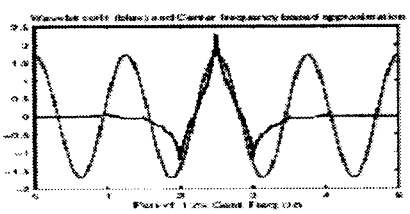


Fig. 3. Coiflet wavelet function

The Coiflet scale of wavelet decomposition is considered for the fault detection, 6th decomposition scale of 12 scales is analyzed for the fault detection. We can find the result in the Fig. 4.
Among the 6th decomposition results, around the 4th value represents good character to distinguish each faults. We can also check the zoom in values in Fig. 5.

We have performed these processing 20 times, these results are illustrate in Table 1.

In theorem 3.1, we have proposed the entropy with the distance measure, next membership function is formulated using the peak average values of Table 1. Membership function is shown in Fig. 6. Hence, from the result of Thm 3.1, we present the entropy of the normalized values.

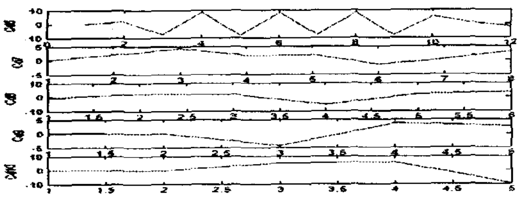


Fig. 4. Details and approximates signals

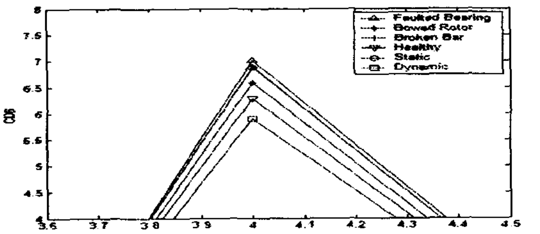


Fig. 5. Zoom in graph around the 4th value

Table 1. Gradient and peak values of the 4th value

No. \ Con.	Faulted Bearing		Bowed Rotor		Broken Rotor Bar		Healthy		Static Eccentricity		Dynamic Eccentricity	
	Gradient	Peak	Gradient	Peak	Gradient	Peak	Gradient	Peak	Gradient	Peak	Gradient	Peak
1	14.88	7.0172	13.71	6.5854	14.49	6.8991	13.04	6.2908	14.51	6.8780	12.29	5.9028
2	14.80	6.8379	13.55	6.5338	14.09	6.7671	13.40	6.4371	14.01	6.6925	11.79	5.7382
3	14.84	7.0021	13.27	6.4420	13.93	6.7198	13.42	6.4637	14.38	6.8110	12.21	5.8725
4	14.58	6.8780	13.32	6.4270	14.09	6.7804	13.27	6.3828	14.38	6.7807	11.69	5.6869
5	14.68	6.9143	13.38	6.4479	13.86	6.6965	13.38	6.3998	14.10	6.6862	12.68	6.1149
6	15.29	7.2088	12.66	6.2307	13.90	6.7061	13.08	6.3142	14.05	6.6699	12.48	5.9872
7	14.80	6.9874	14.23	6.7510	13.99	6.7557	13.20	6.3737	14.10	6.6882	12.84	6.0213
8	14.87	7.0154	13.90	6.6272	13.96	6.7282	13.14	6.3628	14.82	7.0270	12.37	5.9288
9	14.48	6.8168	13.34	6.4560	14.29	6.8634	13.44	6.4484	14.59	6.9077	11.96	5.7998
10	14.70	6.9285	13.60	6.5718	14.04	6.7558	13.51	6.4572	13.82	6.6394	11.98	5.8011
11	14.80	6.9789	13.72	6.5788	13.64	6.5992	13.13	6.3512	14.08	6.6932	11.63	5.6883
12	14.71	6.9284	13.57	6.5133	13.91	6.7101	13.32	6.4012	14.08	6.6880	11.90	5.7804
13	14.67	6.9238	13.38	6.4862	14.08	6.7501	13.42	6.4213	14.28	6.7739	12.39	5.9980
14	15.01	7.0807	12.25	6.0885	14.08	6.7694	13.40	6.4247	14.34	6.7925	12.18	5.8778
15	14.52	6.8524	13.85	6.6335	13.98	6.7465	13.28	6.3843	14.34	6.7926	12.32	5.9288
16	14.72	6.9335	13.42	6.4734	14.25	6.8368	13.25	6.3811	14.70	6.9562	12.55	6.0009
17	14.50	6.8388	14.08	6.6788	14.22	6.7934	13.42	6.4254	13.71	6.6383	12.68	6.0024
18	14.45	6.8124	13.23	6.4259	13.91	6.7165	12.82	6.2288	14.39	6.8116	11.83	5.7474
19	14.78	6.9592	13.78	6.6280	14.12	6.7986	13.17	6.3510	13.90	6.6699	12.13	5.8449
20	14.80	6.9924	13.80	6.6083	13.99	6.7382	13.39	6.4127	14.28	6.7662	12.14	5.8468
AVRG	14.72416	6.94626	13.50022	6.50711	14.06978	6.75560	13.27085	6.38259	14.24060	6.78749	12.16707	5.87602
Order	6		3		4		2		5		1	
STD	0.20825	0.09887	0.45419	0.15470	0.18330	0.06394	0.17154	0.06868	0.28803	0.10882	0.31817	0.11973
CORR	-0.20884	-0.15738	-0.00232	-0.00877	0.13360	0.04461	1.00000	1.00000	-0.25017	-0.19993	0.07075	0.02187
COVAR	-0.00889	-0.00085	-0.00017	-0.00008	0.00399	0.00016	0.02795	0.00326	-0.01174	-0.00121	0.00367	0.00015

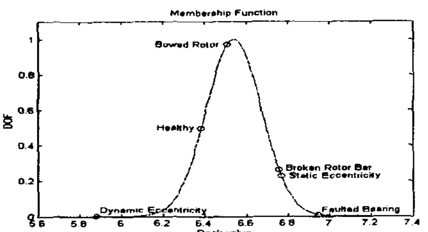


Fig. 6. Membership function of healthness

Table 2. Degree of membership function

Fault Case	Entropy
Faulted Bearing	0.0094
Bowed Rotor	0.9709
Broken Rotor Bar	0.2651
Healthy	0.4982
Static Eccentricity	0.2288
Dynamic Eccentricity	0.0000

V. Conclusions

We investigate the relations of entropy, distance measure and similarity measure. By the definition and results of Liu, we propose new entropy formula with the distance measure. For the faulted induction motor current signals, frequency and wavelet transform have been carried out. Through the wavelet transform, we can find the 4th value of 6th detail result from the 12 scales of wavelet decomposition is useful to analyze features of fault signals. Furthermore, proposed entropy computation is carried out to the faulted induction machine.

References

[1] D. Bhandari, N. R. Pal, Some new information measure of fuzzy sets, Inform. Sci. 67, 209-228, 1993.

[2] A. Ghosh, Use of fuzziness measure in layered networks for object extraction: a generalization, Fuzzy Sets and Systems 72, 331-348, 1995.

[3] B. Kosko, *Neural Networks and Fuzzy Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1992.

[4] Liu Xuecheng, Entropy, distance measure and similarity measure of fuzzy sets and their relations, Fuzzy Sets and Systems, 52, 305-318, 1992.

[5] N.R. Pal, S.K. Pal, Object-background segmentation using new definitions of entropy, IEEE Proc. 36, 284-295, 1989.

[6] J. Fan, W. Xie, Distance measure and induced fuzzy entropy, Fuzzy Set and Systems, 104, 305-314, 1999.

[7] S.H. Lee and S.S. Kim, On some properties of distance measures and fuzzy entropy, Proceedings of KFIS Fall Conference 2002, 9-12.

[8] P. Vas, Parameter Estimation, Condition Monitoring,and Diagnosis of Electrical Machines, Clarendron Press, Oxford, 1993.

[9] G. B. Kliman and J. Stein, "Induction motor fault detection via passive current monitoring,"International Conference in Electrical Machines, Cambridge, MA, pp. 13-17, August 1990.

[10] K. Abbaszadeh, J. Milimonfared, M. Haji, and H. A. Toliyat, "Broken Bar Detection In Iduction Motor via Wavelet Transformation," IECON'01: The 27th Annual Conference of the IEEE Industrial Electronics Society, pp. 95-99, 2001.

[11] Masoud Haji and Hamid A. Toliyat, "Patern Recognition-A Technique for Induction Machines Rotor Fault Detection Eccentricity and Broken Bar Fault," Conference Record of the 2001 IEEE Industry Applications Conference, vol. 3, pp. 1572-1578, 30 Sept.-4 Oct. 2001.

[12] S. Nandi, H. A. Toliyat, "Condition Monitoring and Fault Diagnosis of Electrical Machines A Review," IEEE Industry Applications Conference, vol. 1, pp. 197-204, 199