The Application of LCS and Information Entropy as a Novel Fusion Algorithm for Degradation Feature Extraction

Haotian Wang[1] – Jian Sun[2],* – Xiusheng Duan[1] – Ganlin Shan[1] – Wen Yang[2]

[1] Mechanical Engineering College, China
[2] Luoyang Electronic Equipment Test Center of China, China

Feature extraction is significant for the prognostics and health management (PHM) of hydraulic pumps. In this paper, a novel fusion algorithm is proposed based on local characteristic-scale decomposition (LCD), composite spectrum (LCS). and information entropy. To make full use of feature information, the LCS is proposed based on the modification of traditional composite spectral algorithm. LCS high-order power entropy and high-order singular entropy, which are relatively defined in Shannon entropy (SE) and Tsallis entropy (TE), are extracted as initial features. Furthermore, the method of feature fusion is presented to modify the features' conciseness and to improve the performance. Results of the analysis in the experiment indicate that the proposed method is available, and the fused feature is effective in evaluating the pump degradation process.
Keywords: degradation feature extraction, information fusion, LCS, information entropy, hydraulic pump.

Highlights
- *A novel fusion algorithm based on LCS and information entropy for the degradation feature extraction of hydraulic pumps is proposed.*
- *The fusion method based on information entropy is proposed to make fusion of initial features.*
- *The LCS algorithm gains the ability to extract more sensitive features than usual algorithms and it is effective in evaluating the pump degradation process.*

0 INTRODUCTION

Conditional-based maintenance (CBM) is a modern maintenance concept [1]. As an effective method to realize CBM, prognostics and health management (PHM) has become the focus of the research [2]. One typical example is the application of PHM in joint strike fight (JSR) F-35 [3]. To establish the appropriate PHM of equipment, it is first necessary to conduct the research of degradation feature extraction for prognostics [4]. The degradation experiment requires a long time and the vibration signals during the degradation are very complex [5]. However, features extracted by traditional methods [6] are normally based on the single monitoring signal. Tran et al. extracted features by the analysis of the monitoring signal in the time domain [7]. Zhao used the empirical mode decomposition (EMD) in vibration signal analysis and extracted the approximate entropy as the degradation feature [8]. Dong et al. chose the non-extensive wavelet feature scale entropy to be the feature for degradation evaluation [9]. In this situation, some important fault information may be lost. For example, it is known that the loose slipper fault is the typical failure mode of hydraulic pumps [10]. It is caused by the wear between the head of the piston and the hat of the slipper [11]. Finally, the interaction could be monitored on the pump shell in the form of strikes.

However, the strikes of the interaction are not always from only one direction [12]. Analysis based on a single vibration signal can hardly achieve the whole information needed for diagnosis and prognosis. As a result, multi-direction vibration signals are required for processing. Additionally, vibrations of hydraulic pumps caused by the striking from pistons are directly influenced by the coupling effect [13]. It seems to be extremely difficult to effectively extract obvious fault information. To solve this issue, a modified method is required to fuse multi-direction signals of the hydraulic pump to extract the degradation features more effectively.

The traditional fusion algorithms for multiple signals include the weighted fusion algorithm [14], the Kalman filtering algorithm [15] and the wavelet analysis fusion algorithm [16]. Based on the analysis of these methods [17], it is known that some important fault information may be lost due to the inherent disadvantages during signal processing [18]. As a novel fusion algorithm, composite spectrum (CS) can calculate the correlative relations between neighbouring signals for information fusion [19]. Considering the structural characters of hydraulic pumps, some modifications must be made to extract the sensitive features effectively [20]. First, the complex computation caused by the multiplication the CS algorithm must be reduced. On this basis,

---

an appropriate method is required for CS fusion to obtain detailed components from the original signals, which contain sensitive feature information. An LCD algorithm is an efficient nonlinear signal processing method [21]. The feature information in the original signal can be detailed into various intrinsic scale components (ISCs) [22]. The sensitive information is relatively easy to be achieved for CS fusion. Therefore, based upon the modification of the earlier CS, the LCD composite spectrum (LCS) algorithm is proposed for the fusion of various vibration signals to extract sensitive degradation features.

Features extracted by various methods can reveal fault information in different aspects to improve completeness. However, it may also increase information redundancy degree and even result in dimensional disaster [23]. Common methods to solve this problem include decreasing dimensions by simplifying signal features [24]. The typical algorithms are principal component analysis (PCA) and the locality-preserving projection (LPP). PCA performs better only with linear data. If the high-dimensional data needed processing is nonlinear, the PCA's ability for feature dimensionality reduction would be limited [25]. LPP is a widely used manifold learning algorithm. Based on the Laplace mapping, LPP can deal with the nonlinear data [26]. However, LPP always requires many samples [27]. Furthermore, these dimensionally reduction methods may delete parts of features directly, leading to the loss of sensitive information. As analysed, the basis for feature dimensionality reduction is to evaluate the correlations between various vectors [28]. Correlation coefficient is one traditional factor [29]. Because of the structural character of hydraulic pump, the data of the extracted feature are nonlinear to a large extent. In this situation, the correlation coefficient can hardly evaluate the correlations. Based on information entropy theory, the relative entropy can reveal the correlation between nonlinear data by the analysis of the complexity [30]. Therefore, the feature fusion algorithm is proposed to make full use of the extracted features' information and to improve features conciseness.

Consequently, the method for feature fusion using LCS and relative entropy is proposed. This article is organized as follows: In Section 0, the LCS algorithm is presented for multi-channel signal fusion; In Section 1, LCS high-order power entropy and the high-order singular entropy, which are relatively defined in SE and TE, are extracted as initial features; in Section 2, a feature fusion strategy based on relative entropy is put forward in detail and the originally extracted features

are fused into the required new feature; in Section 3, the experiment is presented and the results are verified by the application of the proposed method; in Section 4, some conclusions are drawn.

1 THE PROPOSED LCS FUSION ALGORITHM

1.1 Signals Decomposed by LCD

Assuming that $x(t)$ denotes the original signal, the LCD algorithm is detailed as follows.

Step 1. Confirm the extremum point X_k of $x(t)$ and its correspondent time τ_k, $k = 1, 2, ..., M$. M is the total number of all extremum points.

Step 2. Calculate A_{k+1} ($k = 2, 3, ..., K-1$) and L_k:

$$A_{k+1} = X_k + \frac{\tau_{k+1} - \tau_k}{\tau_{k+2} - \tau_k}(X_{k+2} - X_k), \qquad (1)$$

$$L_k = aA_k + (1-a)X_k, \quad k = 2, 3, ..., K-1. \qquad (2)$$

Step 3. Fit $L_1 \sim L_k$ by the cubic spline function and acquire the baseline $BL_1(t)$. Separate the baseline.

$$h_1(t) = x(t) - BL_1(t), \qquad (3)$$

if $h_1(t)$ meets the ISC condition [21], output $ISC_1 = h_1(t)$.

Step 4. Separate ISC_1 from the original data and get the new signal. Then repeat the steps until $u_n(t)$ is monotonous or a constant function. Get ISC_1, ISC_2, ..., ISC_n and the trending component $u_n(t)$ $u_n(t)$.

1.2 Selection of ISCs

Normally, the correlation coefficient is applied for the reflection of the nonlinearity between variables, and the mutual information is used for the reflection of the linearity relations. In this paper, based on the SE, the mutual information I is defined:

$$\begin{cases} I(y_1, ..., y_n) = \sum_{i=1}^{n} H(y_i) - H(y_1, ..., y_n) \\ H(y_i) = -\int p(y_i) lgp(y_i) dy_i \end{cases}, \qquad (4)$$

where, $H(y_i)$ is the entropy of y_i. $p(y_i)$ is the probability density function for p_i. For the solution of joint probability density function, the probability of the samples falling into multi-dimensional space can be used for approximate estimation.

In this paper, on the basis of the mutual information, the sensitive factor ε is defined as the ISCs selection index. Assume that the sampling parameters are the same, the sampled normal signal is y_{nor}, and the sampled fault signal is y_f. The fault

Wang, H. – Sun, J. – Duan, X. – Shan, G. – Yang, W.

signal is further decomposed by LCD and the ith ISC component is marked as ISC_i. The sensitive factor of ISC_i is ε_i, which is calculated by Eq. (5).

$$\varepsilon_i = I_{1i}\text{-}I_{2i}, \qquad (5)$$

where, I_{1i} denotes the mutual information between ISC_i and y_f. I_{2i} denotes the mutual information between ISC_i and y_{nor}. Based upon the analysis above, ε_i can evaluate the sensitive information contained in ISC_i. The higher the ε_i is, the more obvious the difference between the mutual information I_{1i} and I_{1i} will be. This means that the sensitive fault information contained in the ISC_i is supposed to be more. Therefore, the ISC component with the highest ε_i will be selected from each vibration signal.

1.3 LCS Algorithm

Based on the n_s division of every ISC, the CS is defined as [19]:

$$S_{CS}(f_k) = \sum_{r=1}^{n_s} X_{CS}^r(f_k) X_{CS}^{r*}(f_k) / n_s, \qquad (6)$$

where $X_{CS}^r(f_k)$ denotes the Fourier coefficient and $X_{CS}^{r*}(f_k)$ denotes its complex conjugate in the rth segment in all B signals. $X_{CS}^r(f_k)$ is defined by Eq. (7).

$$X_{CS}^r(f_k) = \sqrt{(S_{x_1\gamma_{12}x_2}^r(f_k)S_{x_2\gamma_{23}x_3}^r(f_k)...S_{x_{B-1}\gamma_{(B-1)B}x_B}^r(f_k))^{1/(B-1)}}, \quad (7)$$

where $\gamma_{12}, \gamma_{23}, ..., \gamma_{(B-1)B}$, denote the coherences. $S_{x_1\gamma_{12}x_2}^r(f_k), S_{x_2\gamma_{23}x_3}^r(f_k), ..., S_{x_{B-1}\gamma_{(B-1)B}x_B}^r(f_k)$ are defined by Eq. (8).

$$S_{x_1\gamma_{12}x_2}^r(f_k) = [X_1^r(f_k)\gamma_{12}X_2^{r*}(f_k)]. \qquad (8)$$

The multiplication of the Fourier coefficient and the complex conjugate exists in CS. If Eq. (8) is combined with Eq. (7), terms like $X_2^{r*}(f_k)X_2^r(f_k)$, $X_3^{r*}(f_k)X_3^r(f_k), ...$ are achieved, which can be merged together. This phenomenon leads to the increasing of the computation complexity. To solve this problem, Eq. (6) in LCS algorithm is redefined by Eq. (9).

$$S_{CS}(f_k) = (\sum_{r=1}^{n_s} X_1^r(f_k)\gamma_{12}^2 X_2^r(f_k)\gamma_{23}^2 X_3^r(f_k)...$$
$$...X_{B-1}^r(f_k)\gamma_{(B-1)B}^2 X_B^r(f_k))^{1/B} / n_s. \qquad (9)$$

Compared with Eq. (1), the computation of LCS algorithm is only around half of the earlier CS algorithm. Furthermore, the high sensitivity

information contained in the selected ISCs is introduced and further processed by LCS.

2 THE PROPOSED METHOD FOR DEGRADATION FEATURE EXTRACTION

2.1 Extraction of LCS high-Order Power Spectral Entropies

Information entropy is an effective factor to quantitatively reveal the uncertain information in the signal by putting up the contained characteristics. As indicated in our previous research in reference [31], the high order entropy is sensitive to non-symmetrical and nonlinear components and it could reveal the information which cannot be obtained by second-order statistical analysis. Therefore, the high-order entropy is used in this paper as the LCS spectrum analysis factor. It can describe the degradation trend. SE is a kind of information entropy that is defined in logarithmic form. For a system which contains Q microstates, the SE is defined by Eq. (10) [32].

$$SE = -Bolt\sum_{i=1}^{Q} p_i \ln(p_i), \qquad (10)$$

where $Bolt$ is the Boltzmann constant, p_i means the occurrence probability of the ith status. SE is usually used to reveal the property of the scope system and it can reflect the general trend. However, the factual system may be more or less correlated in time and space, and SE can hardly reveal all the inner characteristics of the system. Being the generalization of SE, TE is able to reveal the non-scope information which cannot be obtained by SE. TE is defined by Eq. (11) [32].

$$TE = -Bolt(1 - \sum_{i=1}^{Q} p_i^q) / (1-q), \qquad (11)$$

where q is the non-scope parameter. The common settled value of q is 3. Compared with SE, the values of TE range only from 0 to 1. However, there may be the correlations in time and space or certain degradation stages. TE is required to be the supplement of SE. As a result, to reveal the performance degradation of hydraulic pumps effectively and sensitively, the Shannon LCS power spectral entropy (SLPE) and Tsallis LCS power spectral entropy (TLPE) are both selected to be the features.

Based on reference [31], taking the three-channel vibration signals as example, the procedures are detailed as follows.

① LCD is applied in each of the three vibration signals; after selection, three ISCs remained.

② Divide each ISC into N parts and then carry out Fourier transform on each part and obtain $X_i^j, i = 1,2,3, j = 1,...,N$.

③ Calculate the composite spectrum $S_{CF}(f_k)$:

$$S_{CS}(f_k) = (\sum_{j=1}^{N} X_1^j(f_k)\gamma_{12}^2 X_2^j(f_k)\gamma_{23}^2 X_3^j(f_k))^{1/3} / N. \quad (12)$$

④ Make standard procession on the obtained $S_{CF}(f_k)$ and the result is $\{u(t), t=1,2,...,N\}$. Calculate the triple correlation of DCT coefficients.

$$b_i(\lambda_1, \lambda_2) = \frac{1}{\mu^2} \sum_{k_1=-s_1}^{s_1} \sum_{k_2=-s_2}^{s_2} Y^{(i)}(\lambda_1 + k_1)Y^{(i)}...$$
$$...(\lambda_2 + k_2)Y^{(i)}(\lambda_1 + \lambda_2 + k_1 + k_2), \quad (13)$$

where $Y^{(i)}$ is the DCT coefficient, $\mu = f_k/N$, s_1 and s_2 are the 3-order accumulation range, λ_1 and λ_2 are positive and the sum should not be more than half of the sampling frequency f_s.

⑤ Calculate the proportion of each coefficient:

$$p_k = |b_k(\lambda_1, \lambda_2)|^2 / \sum_{l=1}^{N} |b_l(\lambda_1, \lambda_2)|^2, \quad (14)$$

where K means number of the frequency bands.

⑥ Calculate SLPE and TLPE:

$$SLPE = -Bolt(\sum_{k=1}^{K} p_k \ln p_k) / \ln K, \quad (15)$$

$$TLPE = -Bolt(1 - \sum_{k=1}^{K} p_k^q) / (1-q). \quad (16)$$

Based on the analysis, it can be seen that LCS high-order power spectral entropies (SLPE and TLPE) are sensitive to the total energy changes theoretically.

2.2 Extraction of LCS High-Order Singular Spectral Entropies

As another effective spectrum parameter, singular spectral entropy is able to reveal complexity of signal components. Compared with power spectral entropy, it tends to be more sensitive to response energy changing [33]. The first four calculation steps of LCS high order singular entropy are the same as LCS high-order power spectral entropy, and the following steps are:

⑤ Calculate the DCT high order estimation matrix B.

$$B(\lambda_1, \lambda_2) = \sum_{i=1}^{N} b_i(\lambda_1, \lambda_2) / N. \quad (17)$$

⑥ Make singular values decomposition on U and calculate the proportion of each:

$$U = EDW, \quad (18)$$

where $\mathbf{E}$ is $V \times O$ matrix, $\mathbf{D}$ is $O \times O$ matrix and $\mathbf{D} = \text{diag}(x_1, x_2, ..., x_O)$. $\mathbf{W}$ is $O \times T$ matrix, and O is the order of U.

$$p_u = x_u / \sum_{o=1}^{O} x_o. \quad (19)$$

⑦ Calculate LCS high-order singular entropies respectively defined in SE and TE, which is marked as $SLSE$ and $TLSE$:

$$SLSE = -Bolt \sum_{u=1}^{O} p_u \ln p_u, \quad (20)$$

$$TLSE = -Bolt(1 - \sum_{u=1}^{O} p_u^q) / (1-q). \quad (21)$$

LCS high-order singular entropies (SLSE and TLSE) are able to reveal distributions of various mode components in composite spectrum sensitively. If the degradation degree is lighter, distributions of different mode components are in balance and the signal complexity will be higher and, consequently, the entropies will be higher.

2.3 The Proposed Features Fusion Algorithm

The main purpose of features' fusion is to improve the completeness and conciseness. As a typical evaluating factor, the traditional definition for relative entropy is shown in Eq. (22) [34]:

$$I(y_1,...,y_n) = \sum_{i=1}^{n} H(y_i) - H(y_1,...,y_n), \quad (22)$$

where I represents for the relative entropy of vector $y_1,...,y_n$, $H(y_i)$ is the SE of each vector.

Assume that the features data collection extracted from various degradation stages after normalization is $\mathbf{X} = [\mathbf{X}_1,...,\mathbf{X}_n,...,X_N]$, where N is the number of features. X_n is the vector of the nth feature, and $\mathbf{X}_n = [s_1^n,...,s_m^n]^T$, where s_m^n is the nth element of the mth sample.

Step 1. For the extracted N features, calculate $I(\mathbf{X}_{n1}, \mathbf{X}_{n2})$ between two arbitrary feature vectors:

$$I(\mathbf{X}_{n1}, \mathbf{X}_{n2}) = H(\mathbf{X}_{n1}) + H(\mathbf{X}_{n2}) - H(\mathbf{X}_{n1n2}), \quad (23)$$

where, $\mathbf{X}_{n1n2}$ is the new reconstructed feature by the fusion of $\mathbf{X}_{n1}$ and $\mathbf{X}_{n2}$. s_m^{n1n2} is the element of $\mathbf{X}_{n1n2}$, which is defined as:

$$s_m^{n1n2} = p_m^{n1} s_m^{n1} / (p_m^{n1} + p_m^{n2}) + p_m^{n2} s_m^{n2} / (p_m^{n1} + p_m^{n2}). \quad (24)$$

Step 2. Choose $\mathbf{X}_{nx}$ and $\mathbf{X}_{ny}$ with the maximum I_{nxny} and delete these two feature vectors. Then add the new fused feature $\mathbf{X}_{nxny}$ into $\mathbf{X}$ instead. Then the dimension of $\mathbf{X}$ reduces to $N-1$.

Step 3. Repeat the steps above until dimension X becomes 2.

Step 4. Carry out the final fusion of the only two features in X. The final achieved feature is the fused one needed for use.

Based on the gradual fusion strategy, all extracted features are finally fused into only one new feature, effectively improving the concision of computation.

3 EXPERIMENTAL VALIDATION

3.1 Experimental Rig

To verify the proposed method, an experiment was conducted on the test-bed shown in Fig. 1. Three-channel vibration signals were sampled.

Fig. 1. *Installation of the vibration sensors*

According to the technical documentation from the pump manufacturer, the volumetric efficiency (VE) is applied to evaluate the degradation performance. When VE is less than 80 %, the pump is confirmed to be failed. In this experiment, VE was less than 80 % when the total time was 582 h 15 min 17 s. After disassembly, the piston was verified to occur the loose slipper failure, as seen in Fig. 2.

Fig. 2. *The failure of loose slipper*

3.2 Results and Analysis

The whole degradation processes are roughly divided, as shown in Table 1.

Table 1. *Various degradation stages*

| VE	Operating time	Degradation state	Mark
95 % to 100 %	0 to 532 h 55 min 11 s	normal	
94 % to 95 %	532 h 55 min 12 s to 549 h 54 min 46 s	initial failure	F1
92.5 % to 94 %	549 h 54 min 47 s to 562 h 0 min 25 s	slow development	F2
90 % to 92.5 %	562 h 0 min 26 s to 572 h 49 min 33 s	weak failure	F3
85 % to 90 %	572 h 49 min 34 s to 579 h 26 min 52 s	accelerated degradation	F4
80 % to 85 %	579 h 26 min 53 s to 582 h 15 min 17 s	loose slipper formation	F5
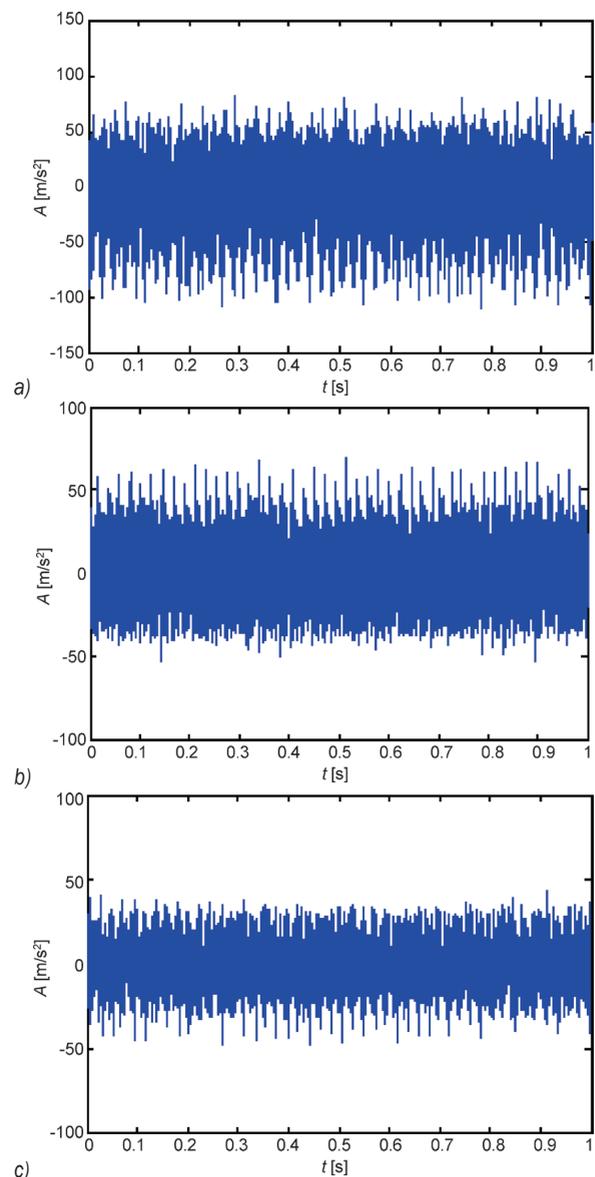
Fig. 3. *Time domains of vibration signals x_{F4}; a) x_{F4-X}, b) x_{F4-Y}, and c) x_{F4-Z}*

One sample of the F1 to F5 period is taken. $\{x_{F1}, x_{F2}, …, x_{F5}\}$ is used to denote the five samples. To simplify the data processing, the initial 1 s of the 10 s sampled data are used for analysis. Every signal contains 5200 points. x_{F4} is applied as a detailed example for analysis. The time domain is shown in Fig. 3.

In Fig. 3, we can see that these original signals can hardly provide enough complete feature information and the influences by noises and other disturbances are obvious. Based on the LCD algorithm, each signal has been decomposed and 10ISCs remained. After selection, the 6th ISC of x_{F4-X}, the 7th ISC of x_{F4-Y} and the 4th ISC of x_{F4-Z} remained. These are equally divided into segments. After Fourier transform, the coefficients X_i^j are achieved. Furthermore, $S_{CS}(f_k)$ can be calculated by Eq. (12).

3.2.1 DCS Power Spectral Entropy

Based on the procedures presented, the following result can be obtained: $SLPE = 3.863$ and $TLPE = 0.725$. Similarly, SLPE and TLPE of the remaining four stages can also be obtained. Data are considered to be one group of results. Then, SLPE and TLPE of the other nine samples in F1 to F5 are calculated. Results are shown in Figs. 4 and 5.
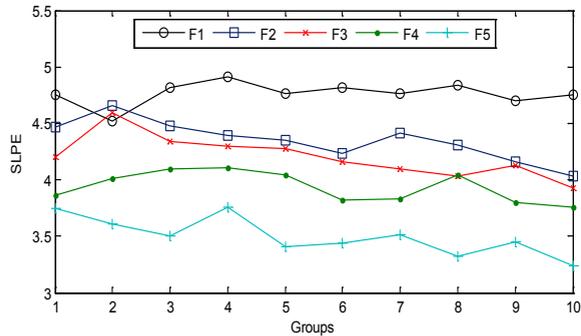
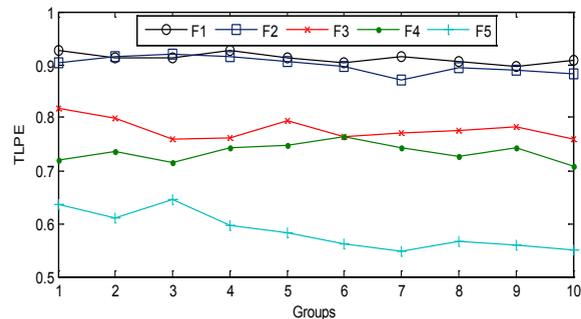**Fig. 4.** *SLPE in F1 to F5 degradation stages*

Fig. 5. *TLPE in F1 to F5 degradation stages*

Fig. 4 shows SLPE in various degradation stages. With the aggravation, SLPE tends to be reduced. This reveals the tracking ability of degradation changing. However, crossovers exist in neighbouring curves obviously, meaning that the distinguish ability of SLPE still needs to be improved. Fig. 5 shows the TLPE in various degradation stages. The crossover and overlapping of neighbouring curves become less, which reveals the complementary of TE to SE.

3.2.2 LCS High-Order Singular Spectral Entropy

Based on the procedures presented, the result can be obtained that: $SLSE = 3.227$ and $TLSE = 0.521$. Finally, totally 10 groups of results are achieved, which are shown in Figs. 6 and 7.
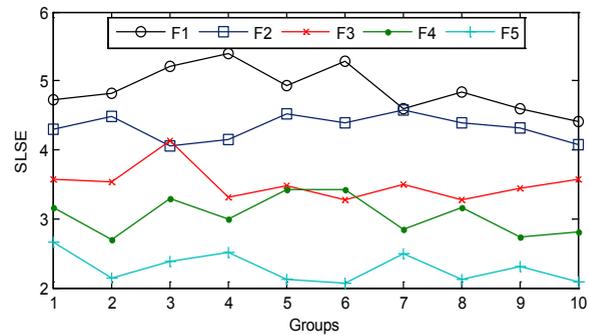
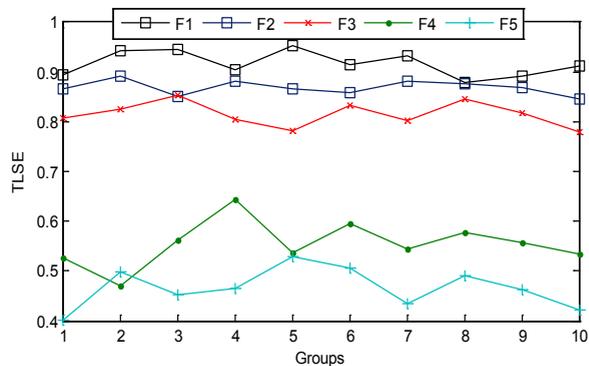**Fig. 6.** *SLSE in F1 to F5 degradation stages*

Fig. 7. *TLSE in F1 to F5 degradation stages*

Compared with LCS high-order power spectral entropies, crossover and overlapping in neighbouring curves of SLSE and SLTE are obviously reduced. This phenomenon means these two entropies possess better distinguishing ability than others. However, in detail, fluctuations in curves of SSE and STE are slightly more obvious, which shows the disadvantages in the stability of singular entropy.

3.2.3 Fusion of Features

The fusion of the ten groups of SLPE, TLPE, SLSE, TLSE in F1 to F5 are carried out, and the results are shown in Fig. 8.
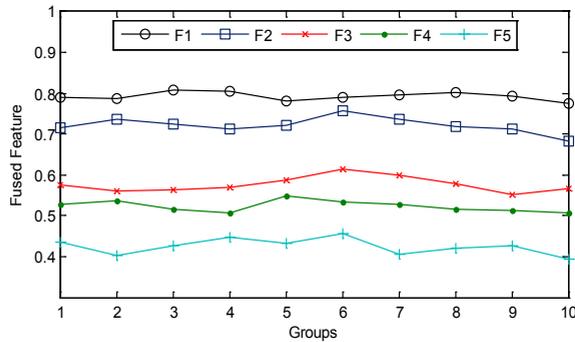

Fig. 8. *Fusion feature in F1 to F5 degradation stages*

3.2.4 Analysis and Comparison of Experimental Results

The double sample Z test **[35]** is applied for quantitative evaluation. Its computation is explained by:

$$Z = \left| \bar{X}_1 - \bar{X}_2 \right| \Big/ \left(\mu_{X_1}^2 \big/ n_1 + \mu_{X_2}^2 \big/ n_2 \right), \qquad (25)$$

where X_1 and X_2 denote sample sets of degradation feature X in various degradation stages. $\bar{X}$ and μ_X^2 respectively mean the average and variance of X. n is the number of samples. In comparison, the wavelet algorithm is used to process the three-channel signals. The amplitude and power of the characteristic frequencies of the constructed signal are extracted as

the traditional features, which are marked as TF-1 and TF-2. Sensitivity is calculated. and results are shown in Table 2.

In Table 2, it can be seen that the traditional features TF-1 and TF-2, based on the wavelet fusion algorithm, could roughly describe the degradation. However, limited by the disadvantages of wavelet transform, the sensitivity of TF-1 and TF-2 are the lowest. Compared with TF-1 and TF-2, the performances of four features extracted by the proposed LCS fusion algorithm are improved. According to the sensitivity of various double samples, the extracted features are complementary to each other. However, some obviously redundant information still remains. Compared with the initial four features, the fused feature keeps the better performances of original features. Sensitivity to each double sample and the mean sensitivity of the fused feature are all the highest. Furthermore, there is only one feature after fusion, so that the conciseness would be greatly improved. This feature's fusion can effectively avoid contradiction disturbances by overloading the feature's information during the prognostic policy-decision stage.

Furthermore, the support vector machine (SVM) is applied for verifying the identification performance of the features. The four extracted features, the fused feature based on LPP algorithm **[26]** and the fused feature based on the proposed algorithm are made comparison. 30 samples of each stages of F1 to F5 have been used. There are a total of 150 values for each feature. The SVM identification results are shown in Table 3.

Table 2. *Sensitivity of various features*

Double Samples	Sensitivity						
	SLPE	TLPE	SLSE	TLSE	TF-1	TF-2	Fused feature
(F1, F2)	10.17	6.38	14.28	12.35	6.38	8.26	17.32
(F2, F3)	12.94	20.58	11.87	15.21	10.94	12.37	24.26
(F3, F4)	8.28	13.15	10.94	24.19	11.27	12.02	25.33
(F4, F5)	19.56	26.31	22.42	19.07	18.95	17.59	28.59
Mean sensitivity	12.74	16.61	14.88	17.70	11.89	12.56	23.88

Table 3. *Degradation stages identification accuracy of various features*

	F1 [%]	F2 [%]	F3 [%]	F4 [%]	F5 [%]	Total accuracy
SLPE	76.7	86.7	73.3	76.7	70	76.7
TLPE	73.3	90	76.7	70	83.3	78.7
SLSE	93.3	86.7	70	66.7	80	79.3
TLSE	76.7	83.3	80	80	83.3	80.6
Fused feature (LPP)	86.7	86.7	83.3	80	86.7	84.6
Fused feature (relative entropy)	93.3	93.3	90	86.7	86.7	90

Table 3 shows the result of the degradation stage identification with various features. As analysed previously, the sensitivity of SLPE to degradation is not very high. Therefore, the identification accuracy of SLPE to various degradation stages is the lowest. The accuracy of TLSE is the best of the four extracted features. However, the total accuracy of SLPE, TLPE, SLSE, and TLSE are close, ranging from 76 % to 80 %. Based on the LPP algorithm, the accuracy of the fused feature is improved to 84.6 %. Limited by the disadvantages, the identification of some stages has not changed so much before or after fusion. The proposed relative entropy fusion algorithm can make full use of the extracted feature information. Consequently, the identification accuracy of the fused feature to each stage and the total accuracy are all the highest. This means that the fused feature can evaluate various degradation stages of hydraulic pumps.

Above all, the comparison results show that the presented LCS algorithm gains the ability to extract more sensitive features. Meanwhile, unlike the traditional feature dimension reduction method, the relative entropy algorithm can make full use of each feature's information, so that the performance and concision of the fused feature will be both improved.

4 CONCLUSIONS

A novel fusion algorithm for extracting degradation feature is proposed here. Following the application in the degradation experiment, major conclusions are drawn, as follows:
1. LCS method for vibration signals fusion is presented to take advantages of fault information. LCS high-order power spectral entropy and high-order singular entropy defined in SE and TE are extracted as initial features.
2. The fusion method based on information entropy is proposed to make fusion of the initial features. The concision and performance are both effectively improved.
3. Experimental results showsthat the fused feature performances much better describe the degradation of hydraulic pump effectively.

5 ACKNOWLEDGEMENTS

8 REFERENCES

[1] Ahmad, R., Kamaruddin, S. (2012). An overview of time-based and condition-based maintenance in industrial application. *Computers & Industrial Engineering*, vol. 63, no. 1, p. 135-149, **DOI:10.1016/j.cie.2012.02.002**.

[2] Coraddu, A., Oneto, L., Ghio, A., Savio, S., Figari, M., Anguita, D. (2015). Machine learning for wear forecasting of naval assets for condition-based maintenance applications. *International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles*, p. 1-5, **DOI:10.1109/ESARS.2015.7101499**.

[3] Lee, J., Wu, F., Zhao, W., Ghaffari, M., Liao, L., Siegel, D. (2014). Prognostics and health management design for rotary machinery systems-reviews, methodology and applications. *Mechanical Systems and Signal Processing*, vol. 45, no. 1-2, p. 314-334, **DOI: 10.1016/j.ymssp.2013.06.004**.

[4] Dong, S.J., Luo. T.H. (2013). Bearing degradation process prediction based on the PCA and optimized LS-SVM model, *Measurement*, vol. 46, no. 9, p. 3143-3152, **DOI:10.1016/j.measurement.2013.06.038**.

[5] Zhu, X. Zhang, Y., Zhu, Y. (2013). Bearing performance degradation assessment based on the rough support vector data description. *Mechanical Systems and Signal Processing*, vol. 34, no. 1, p. 203-207, **DOI:10.1016/j.ymssp.2012.08.008**.

[6] Wang, Y., Kang, S., Jiang, Y., Yang, G., Song, L., Mikulovich, V.I. (2012). Classification of fault location and the degree of performance degradation of a rolling bearing based on an improved hyper-sphere-structured multi-class support vector machine. *Mechanical Systems and Signal Processing*, vol. 29, p. 404-414, **DOI:10.1016/j.ymssp.2011.11.015**.

[7] Tran, V.T., Pham, H.T., Bo-Suk Yang, B.-S., Nguyen, T.T. (2012). Machine performance degradation assessment and remaining useful life prediction using proportional hazard model and support vector machine. *Mechanical Systems and Signal Processing*, vol. 32, p. 320-330, **DOI:10.1016/j.ymssp.2012.02.015**.

[8] Zhao, S.F., Liang, L., Xu, G.H., Wang, J., Zhang, W.M. (2013). Quantitative diagnosis of a spall-like fault of a rolling element bearing by empirical mode decomposition and the approximate entropy method. *Mechanical Systems and Signal Processing*, vol. 40, no. 1, p. 154-177, **DOI:10.1016/j.ymssp.2013.04.006**.

[9] Dong, S.J., Tang, B.P., Chen, R.X. (2013). Bearing running state recognition based on non-extensive wavelet feature scale entropy and support vector machine. *Measurement*, vol. 46, no. 10, p. 4189-4199, **DOI:10.1016/j.measurement.2013.07.011**.

[10] Sun, J., Li, H., Xu, B. (2016). Degradation feature extraction of the hydraulic pump based on high-frequency harmonic local characteristic-scale decomposition sub-signal separation and discrete cosine transform high-order singular entropy. *Advances in Mechanical Engineering*, vol. 8, no. 7, p. 1-12, **DOI:10.1177/1687814016659601**.

[11] Wang, J.P., Hu, H.T. (2006). Vibration-based fault diagnosis of pump using fuzzy technique. *Measurement*, vol. 39, no. 2, p. 176-185, **DOI:10.1016/j.measurement.2005.07.015**.

[12] Sun, J., Li, H.R., Xu, B.H. (2016). The morphological undecimated wavelet decomposition - discrete cosine

transform composite spectrum fusion algorithm and its application on hydraulic pumps. *Measurement*, vol. 94, p. 794-805, **DOI:10.1016/j.measurement.2016.09.024**.

[13] Du, J., Wang, S.P., Zhang, H.Y. (2013). Layered clustering multi-fault diagnosis for hydraulic piston pump. *Mechanical Systems and Signal Processing*, vol. 36, no. 2, p. 487-504, **DOI:10.1016/j.ymssp.2012.10.020**.

[14] Angelov, P., Yager, R. (2013). Density-based averaging - A new operator for data fusion. *Information Science*, vol. 222, p. 163-174, **DOI:10.1016/j.ins.2012.08.006**.

[15] Rodger, J.A. (2012). Toward reducing failure risk in an integrated vehicle health maintenance system: A fuzzy multi-sensor data fusion Kalman filter approach for IVHMS. *Expert Systems with Applications*, vol. 39, no. 10, p. 9821-9836, **DOI:10.1016/j.eswa.2012.02.171**.

[16] Geng, H., Liang, Y., Liu, Y.R., Alsaadi, F.E. (2018). Bias estimation for asynchronous multi-rate multi-sensor fusion with unknown inputs. *Information Fusion*, vol. 39, no. 1, p. 139-153, **DOI:10.1016/j.inffus.2017.03.002**.

[17] Wei, C.M., Blum, R.S. (2010). Theoretical analysis of correlation-based quality measures for weighted averaging image fusion. *Information Fusion*, vol. 11, no. 4, p. 301-309, **DOI:10.1016/j.inffus.2009.10.006**.

[18] Yang, W.B., Li, S.Y. (2014). A switch-mode information fusion filter based on ISRUKF for autonomous navigation of spacecraft. *Information Fusion*, vol. 18, no. 1, p. 33-42, **DOI:10.1016/j.inffus.2013.04.012**.

[19] Elbhbah, K., Sinha, J.K. (2013). Vibration-based condition monitoring of rotating machines using a machine composite spectrum. *Journal of Sound and Vibration*, vol. 332, no. 11, p. 2831-2845, **DOI:10.1016/j.jsv.2012.12.024**.

[20] Yunusa-Kaltungo, A., Sinha, J.K., Elbhbah, K. (2014). An improved data fusion technique for faults diagnosis in rotating machines. *Measurement*, vol. 58, no. 1, p. 27-32, **DOI:10.1016/j.measurement.2014.08.017**.

[21] Jinde, Z., Junsheng, C., Yu, Y. (2013). A rolling bearing fault diagnosis approach based on LCD and fuzzy entropy. *Mechanism and Machine Theory*, vol. 70, no. 1, p. 441-453, **DOI:10.1016/j.mechmachtheory.2013.08.014**.

[22] Junsheng, C., Yi, Y., Yu, Y. (2012). Local characteristic-scale decomposition method and its application to gear fault diagnosis. *Journal of Mechanical Engineering*, vol. 48, no. 9, p. 45-49, **DOI:10.3901/JME.2012.09.064**.

[23] Zhu, K.H., Song, X.G., Xue, D.X. (2014). A roller bearing fault diagnosis method based on hierarchical entropy and support vector machine with particle swarm optimization algorithm. *Measurement*, vol. 47, p. 669-675, **DOI:10.1016/j.measurement.2013.09.019**.

[24] Li, H.J., Xie, X.P. (2015). Gear fault pattern recognition based on kernel feature fuzzy clustering and fuzzy association entropy. *Chinese Journal of Scientific Instrument*, vol. 36, no. 4, p. 849-854.

[25] Gharavian, M.H. Almas Ganj, F., Ohadi, A.R., Heidari Bafroui, H. (2013). Comparison of FDA-based and PCA based features in fault diagnosis of automobile gearboxes. *Neurocomputing*, vol. 121, no. 1, p. 150-159, **DOI:10.1016/j.neucom.2013.04.033**.

[26] Deng, X.G., Tian, X.M. (2013). Sparse kernel locality preserving projection and its application in nonlinear process fault detection. *Chinese Journal of Chemical Engineering*, vol. 21, no. 2, p. 163-170, **DOI:10.1016/S1004-9541(13)60454-1**.

[27] Dornaika, F., Assoum, A. (2013). Enhanced and parameterless locality preserving projections for face recognition. *Neurocomputing*, vol. 99, no. 1, p. 448-457, **DOI:10.1016/j.neucom.2012.07.016**.

[28] Krawczak, M., Szkatula, G. (2014). An approach to dimensionality reduction in time series. *Information Sciences*, vol. 260, no. 1, p. 15-36, **DOI:10.1016/j.ins.2013.10.037**.

[29] Kristoufek, L. (2014). Measuring correlations between non-stationary series with DCCA coefficient. *Physica A: Statistical Mechanics and its Applications*, vol. 402, p. 291-298, **DOI:10.1016/j.physa.2014.01.058**.

[30] Zarinbal, M., Fazel Zarandi, M.H., Turksen, I.B. (2014). Relative entropy fuzzy c-means clustering. *Information Sciences*, vol. 260, p. 74-97, **DOI:10.1016/j.ins.2013.11.004**.

[31] Sun, J., Li, H.R., Xu, B.H. (2016). Degradation feature extraction of the hydraulic pump based on high-frequency harmonic local characteristic-scale decomposition sub-signal separation and discrete cosine transform high-order singular entropy. *Advances in Mechanical Engineering*, vol. 8, no. 7, p. 1-12, **DOI:10.1177/1687814016659601**.

[32] Rufeil Fiori, E., Plastino, A. (2013). A Shannon-Tsallis transformation. *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 8, p. 1742-1749, **DOI:10.1016/j.physa.2012.12.037**.

[33] Caraiani, P. (2014). The predictive power of singular value decomposition entropy for stock market dynamics. *Physica A: Statistical Mechanics and its Application*, vol. 393, p. 571-578, **DOI:10.1016/j.physa.2013.08.071**.

[34] Zarinbal, M., Fazel Zarandi, M.H., Turksen, I.B. (2014). Relative entropy fuzzy c-means clustering. *Information Science*, vol. 260, no. 1, p. 74-97, **DOI:10.1016/j.ins.2013.11.004**.

[35] Zhe, C. (2011). *Theory and Method on Damage Modeling and Prognostics for Planetary Gear Set of Helicopter Transmission System*. PhD Thesis, National University of Defense Technology, Changsha.