

# Online Defect Detection of Laser-Arc Hybrid Welding Based on Spectral Information and MSPC



Chengyuan Ma, Bo Chen, Caiwang Tan, Xiaoguo Song, and Jicai Feng

**Abstract** Online monitoring of welding quality is always a research hotspot. The online detection of welding defects can effectively improve the processing efficiency and ensure the welding quality. In this paper, a quality detection system for laser-arc hybrid welding process has been established by using a spectrometer. Principal component analysis (PCA) and multivariate statistical process control (MSPC) were used to analyze the collected spectral information. The results showed that the welding defects such as oxide inclusion and the lack of penetration could be effectively detected by selecting the specific elements, such as the spectral lines of Fe/Cr/Ni/Ar/O.

**Keywords** Spectral information · Laser-arc hybrid welding · Welding defects detection · PCA · MSPC

## 1 Introduction

With the development of laser, laser manufacturing including laser welding, laser cladding and laser additive manufacturing has been paid more and more attention by researchers [1]. A series of physical and chemical changes will occur in welding, including arc spectrum, acoustic spectrum, changes in welding current and voltage [1, 2]. It is well known that laser-induced plasma is produced during laser-arc hybrid welding process, and the quality of weld could be reflected by the characteristics of the plasma [3]. Therefore, real-time monitoring of signals can provide reliable basis for welding quality prediction.

---

C. Ma · B. Chen (✉) · C. Tan · X. Song · J. Feng  
State Key Laboratory of Advanced Welding and Joining,  
Harbin Institute of Technology, Harbin 150001, China  
e-mail: [chenb@hit.edu.cn](mailto:chenb@hit.edu.cn)

Shandong Provincial Key Laboratory of Special Welding Technology,  
Harbin Institute of Technology at Weihai, Weihai 264209, China

© Springer Nature Singapore Pte Ltd. 2020  
S. Chen et al. (eds.), *Transactions on Intelligent Welding Manufacturing*,  
Transactions on Intelligent Welding Manufacturing,  
[https://doi.org/10.1007/978-981-15-6922-7\\_4](https://doi.org/10.1007/978-981-15-6922-7_4)

To detect defects in welding from spectral information, many researchers have done a lot of work. Shea et al. used spectrum to monitor hydrogen pollution in welding arc in real time to control hydrogen and hydrogen pores in weld metal [4]. Morgan et al. studied the relationship between the non-fusion defects caused by laser spot offset from the joint and plasma spectral characteristics in narrow gap laser welding. The results of offline analysis of spectral intensity and electronic temperature showed that the combination of spectral characteristics and visual sensing is beneficial to the hybrid sensing tracking of weld seam [5]. Kong et al. used fiber laser with the maximum output power of 4 kW to weld 1.2 and 1.5 mm high-strength galvanized steel plate DP980. Their research found that the strength of spectral line in the welding process was related to welding spatter caused by zinc vapor, and the electronic temperature of plasma calculated by Boltzmann plot method could reflect whether there were pores in the weld [6]. Sibillano et al. conducted an online spectral study of butt welding of AA5083 aluminum alloy, which compared the effects of different shielding gas nozzles and the flow rate on the spectral strength during welding. Their results showed that when the state of shielding gas changes, it can be well reflected in the spectral characteristics [7]. Ribic et al. studied YAG laser-TIG hybrid welding plasma by means of spectroscopy, and studied the influence of welding current and heat source spacing on electron temperature and conductivity of plasma. The composition of plasma under different welding currents was calculated, and the fusion depth under different welding currents was deduced according to the heat transfer model and fluid model [8]. Harooni et al. studied the influence of oxide film on arc electron temperature and weld quality of AZ31B Al-Mg alloy during zero-gap laser lap welding. It was found that, compared with the complete cleaning of the oxide film, when the oxide film was not cleared, the electron temperature and fusion depth obtained based on the spectral lines of MgI at wavelengths of 383.83 and 517.26 nm were relatively high, but pores appeared on the surface [9].

In addition to time domain analysis of electron temperature and strength of characteristic spectral line, some researchers have applied statistical and machine learning algorithms to spectral data processing. By using spectral information and principal component analysis (PCA), Colombo et al. proposed a monitoring scheme to measure the gap in laser lap welding of galvanized steel. This scheme can not only monitor the assembly gap value but also evaluate the position where spectral emitted. This study also found a strong correlation between laser back reflection and assembly gap [10]. Zhang et al. studied the key technology of online detection of internal pores of Al-Mg alloy in pulsed GTAW by means of spectral detection. The internal pores in aluminum alloy welding could be measured online. The samples with different porosity were tested and characterized by SEM and EDS. The spectral lines of metal and hydrogen were selected before extracting the characteristics. The relationship between the principal component coefficients of HI and MgI lines and porosity was studied quantitatively. On this basis, an improved characteristic parameter, the absolute value coefficient of the H spectrum in first principal component, was proposed for the detection of internal porosity. Finally, feature reduction and visualization were realized through PCA and *t*-distribution random neighborhood embedding, and pores

density was successfully classified [11]. Lacroix and Jeandel [12] studied the spectral characteristics of plasma in pulsed Nd: YAG laser welding. The change of electron temperature directly related to material heat transfer under different processing modes was studied. The maximum temperature observed was 7150 K, which was lower than the CO<sub>2</sub> laser welding. Welding shielding gas has a cooling effect on plasma. The higher the velocity of shielding gas, the stronger the cooling effect. The temperature at the top of the keyhole is higher than that of the rest of the plasma.

In this paper, the detection of defects in 316LN welding was studied. 316LN is austenitic stainless steel, which is a kind of structural material commonly used in nuclear power field. N element plays the role of solution strengthening and aging strengthening to improve the strength, hardness, corrosion resistance and wear resistance of steel. 316LN stainless steel has excellent mechanical properties at high temperature, good weldability, corrosion resistance and fatigue resistance. It is a widely used material for nuclear reactor structures. Therefore, the quality of welding is of great significance. In this paper, a quality detection system for laser-arc hybrid welding process was established by using a spectrometer. Due to the large amount of spectral information, it is necessary to reduce the dimensionality. Thus, the principal component analysis (PCA) and multivariate statistical process control (MSPC) were used to analyze the collected spectral information. Since MSPC is a data-driven approach, compared with other machine learning algorithms, MSPC does not need any wrong data during the training phase. In addition, MSPC chart can continuously display variable status and be updated over time to perform dynamic monitoring.

## 2 Experiment

The experiments were conducted based on laser-arc hybrid welding system, the system was consisted of an Ytterbium-doped fiber laser (IPG YLS-6000), an arc welding power source (LINCOLN R350) and a KUKA six-axis robot with a self-designed laser-arc hybrid welding torch. The maximum output power of the laser used is 6000 W with wavelength of 1060–1070 nm; the beam parameter product (BPP) is greater than or equal to 4.0. An AvaSpec multichannel fiber optic spectrometer matched with a collecting fiber optic probe, of which the spectral resolution is  $0.052 \pm 0.001$  nm, was used to collect the emission spectra ranging from 315 to 420 nm. The schematic diagram of the experimental system is shown in Fig. 1.

The butt-welding experiment was carried out with 316LN stainless steel plate cut by flame, and the dimensions of plate were  $150 \times 75 \times 6$  mm. To explore the influences of the oil and oxide pollution on the spectrum, in the first experiment, the groove was cleaned before welding, but the oxide film generated by flame cutting was retained, as shown in Fig. 2a. In the second experiment, the welding gap was enlarged to 2 mm in the middle of the plate, as shown in Fig. 2b. The filler metal is 316L.

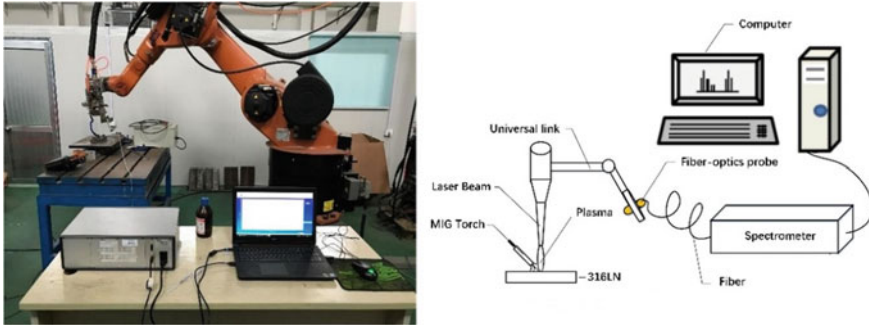


Fig. 1 Schematic diagram and photograph of welding system and spectral acquisition system

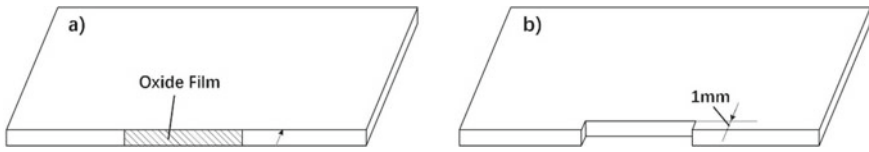


Fig. 2 Schematic diagram of base metal. a Retain the oxide film. b Welding gap variation

In this paper, the laser power, welding current, welding speed and the laser-arc distance were, respectively, set to 3300 W, 150 A, 0.8 m/min and 3 mm. The spectrometer was used to collect spectral signals during welding. The integral time was 10 ms, and the number of averages was 2.

### 3 Results and Discussion

#### 3.1 Multivariate Statistical Process Control

In practice, multivariate statistical analysis often encounters two situations: One is to monitor the stability of a given sample of multivariate observations, and the other is to set a control field for future observations. The commonly used multivariate statistical control charts include square prediction error (SPE) chart,  $T^2$  chart. The SPE chart depicts the deviation of the measured value from the predicted value at a certain moment, but it is difficult to give the predicted value to the spectral signal in the welding process. Therefore,  $T^2$  chart based on Hotelling statistics was selected to conduct real-time monitoring of the spectral signals of the welding process. A  $T^2$  control chart can handle no less than two variables, so that multiple characteristic spectral lines can be monitored simultaneously.

In monitoring the stability of a given sample, historical normal data are used to train a dimensionality reduction model such as PCA, in which high-dimensional feature data are represented by low-dimensional features (such as dominant principal components and independent components) [13]. The algorithm of PCA is briefly elaborated below.

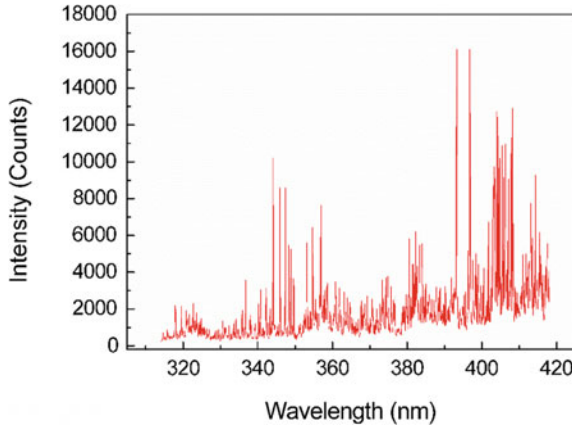
Principal component analysis (PCA) is the most widely used in data reduction and data compression. The main idea of PCA is to find a  $k$ -dimensional space to map the original  $n$ -dimensional data into the  $k$ -dimensional space and retain the main information of the data. Each axis in the  $k$ -dimensional orthogonal eigenspace is also called the principal component (PC). The choice of the new axis is closely related to the data itself. Among them, the selection of the first new coordinate axis is the direction with the largest difference in the original data, the selection of the second new coordinate axis is the plane with the largest variance in the orthogonal plane with the first coordinate axis, and the third axis is the plane with the largest difference in the orthogonal plane with the first and second axis. And so on, you get  $k$ -dimensional space of these axes. With the new axis obtained in this way, we find that most of the variance is contained in the first  $k$  axis, and the variance in the latter axis is almost 0. Therefore, we can ignore the remaining axis and only keep the first  $k$  axis that contains most of the variance. In fact, this is equivalent to only retaining the dimension features that contain most of the variance, while ignoring the feature dimensions that contain almost zero variance, so as to realize the dimensionality reduction processing of data features.

When implementing multivariate statistical process control, it is necessary to establish a principal component model to reflect the normal process. The historical data reflecting the normal process were collected, the principal component analysis was carried out on these data, and the principal component model was established. Since the results of principal component analysis are affected by the data scale, the data should be standardized first, the mean value of each variable should be subtracted, and then divided by its standard deviation. It could be expressed as following equation:

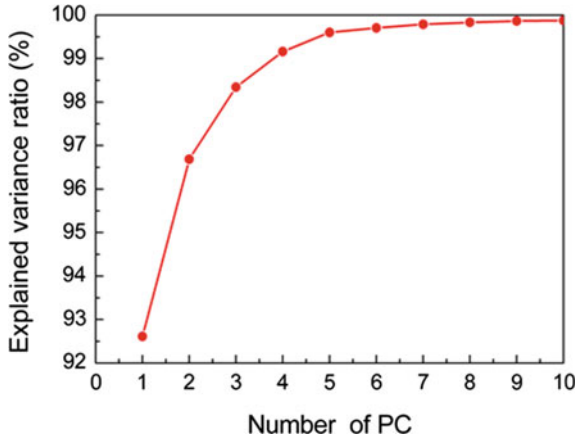
$$\bar{X}_s = [X - (1 \ 1 \ \dots \ 1)^T M] \text{diag} \left( \frac{1}{s_1} \ \frac{1}{s_2} \ \dots \ \frac{1}{s_m} \right) \quad (1)$$

where  $M = [m_1 \ m_2 \ \dots \ m_m]$  is the mean value of  $X$ ,  $s = [s_1 \ s_2 \ \dots \ s_m]$  is the standard deviation of  $X$ .

In this paper, 52 spectral lines from Fe/Cr/Ni/Ar/O elements were selected from the original data as the input variable of PCA. The original spectral data in an integration time are shown in Fig. 3. The data analysis methods library provided by Scikit-learn [14] was used to standardize the data and reduce the dimension by principal component analysis, so as to obtain the interpretation of data changes by principal component models with different numbers of PC. As shown in Fig. 4, PCA results showed that the first six principal components contributed more than 99.5% of the variance. Furthermore, the original 52-dimensional data can be reduced to



**Fig. 3** Original spectral data in an integration time



**Fig. 4** Interpretation of data changes by principal component models with different principal component numbers

six-dimensional data through PCA without losing most of the information in the data.

After PCA dimension reduction, six linearly independent principal components of spectral data are obtained, and MSPC control chart can be drawn by using these principal components.

Stability monitoring is considered first. Assume that monitored vector  $X_1, X_2, \dots, X_n$  independent identical distribution, and distribution obey  $N_p(\mu, \Sigma)$ , then calculated the  $T^2$  for the  $j$  point:

$$T_j^2 = (x - \bar{x})^T S^{-1} (x - \bar{x}) \tag{2}$$

where  $x$  is the observations of sample,  $\bar{x}$  is the mean value of the observations,  $S$  is

the covariance matrix of the variables in the sample, and  $S = \begin{bmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & \ddots & \vdots \\ s_{m1} & \cdots & s_{mn} \end{bmatrix}$ , where

$s_{ij}(i, j = 1, 2, \dots, p)$  is the covariance of the variable  $x_i$  with  $x_j$ .

Equation (1) has standardized the  $X$ , so  $\bar{x} = 0$ ; consequently, the above equation can be simplified as:

$$T_j^2 = x^T S^{-1} x \tag{3}$$

Then, the  $T^2$  value was drawn on the time axis, and the lower control limits was set as zero, while the upper control limits could be set as  $\chi_p^2(0.01)$  and its value can be obtained by looking up the table. When the  $T^2$  value is below the upper control limits (UCL), the process can be considered stable within the 99% confidence interval.

Then consider setting up control fields for future observations. If the current observed process is stable, the control fields of a single observed value in the future can be predicted based on the current observed value. Assume that monitored vector  $X_1, X_2, \dots, X_n$  independent identical distribution, and the distribution obey  $N_p(\mu, \Sigma)$ ; let  $X$  be a future observation from the same distribution, then [15]

$T^2 = \frac{n}{n+1}(X - \bar{X})^T S^{-1}(X - \bar{X})$  obeys  $\frac{(n-1)p}{n-p} F_{p,n-p}$  distribution.

So, for each new observation  $x$ , apply Eq. (3) in chronological order to get  $T^2$

$$T = \frac{n}{n+1} x^T S^{-1} x \tag{4}$$

where  $x$  is the sample data,  $S$  is the covariance matrix of the stable data, and  $n$  is the number of samples of the stable data.

It is important to note that process changes are not necessarily caused by changes that adversely affect the process. If the  $T^2$  statistic or other statistic exceeds the control limits, and it can only indicate that there are sample points in the system that are far away from the data aggregation range, that is, specific points [15]. The existence of such points is sometimes caused by the characteristics of the process itself. When analyzing the statistical process control chart, it is necessary to combine the specific process knowledge, that is the characteristics of the welding process itself, which is of great significance for obtaining correct results.

The basic drawing steps of the statistical process control chart used to monitor the stability of welding process are as follows:

- (a) Determine the variables to be monitored and collect spectral data from the normal welding process.
- (b) The monitored variables were dimensionalized by PCA, and the  $T^2$  value was calculated according to Eq. (3). After the UCL was determined, the control chart was drawn to observe the process stability.
- (c) On the basis of stable process, UCL of future observations could be set as:

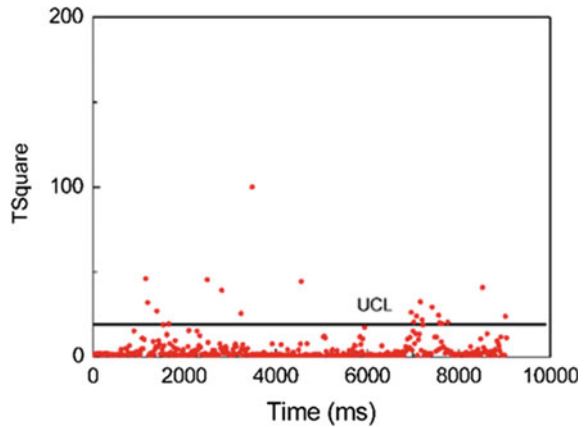
$$UCL = \frac{(n - 1)p}{n - p} F_{p, n-p}(0.01)$$

The newly collected data were dimensionally reduced by PCA, and the  $T^2$  value was calculated in time sequence according to Eq. (4).

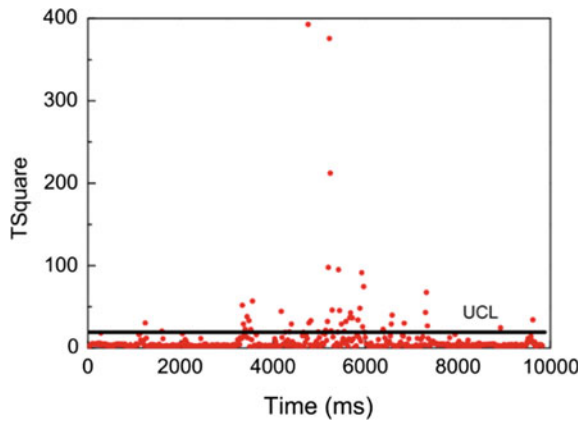
### 3.2 Explanation of Statistical Process Control Chart

The  $T^2$  value was calculated according to the above steps, the control chart of normal welding process was shown in Fig. 5, and the control chart of welding process with oil and oxide film was shown in Fig. 6.

**Fig. 5** Control chart of normal welding process



**Fig. 6** Control chart of welding process with oil and oxide film







**Fig. 7** Photograph of weld with oxide film

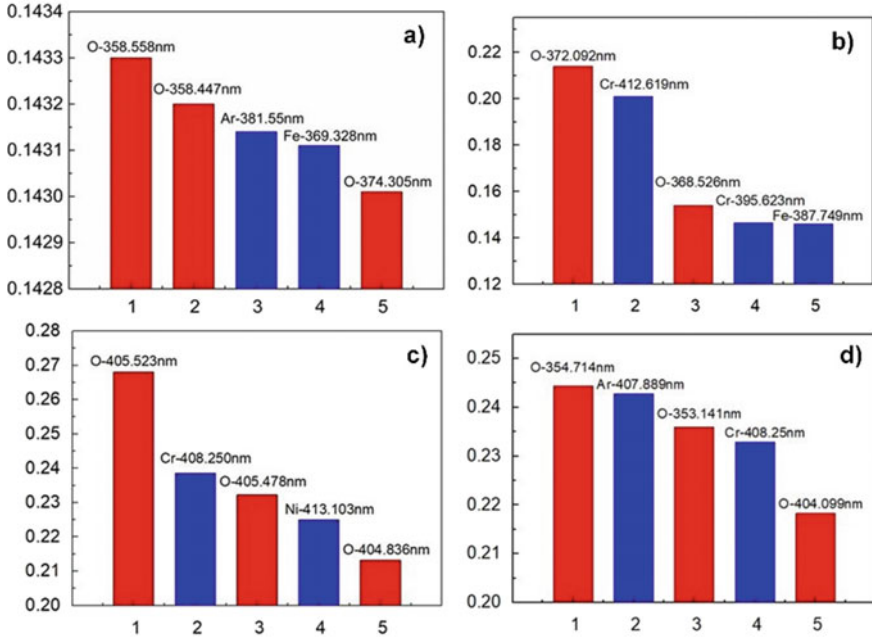
As shown in Fig. 5, in the normal welding process, only a small part of sporadic data exceeded the control limits, which was evenly distributed throughout the whole welding process with a small extent of deviation, which was a normal phenomenon due to the random fluctuation caused by the unstable transition of melt drops in MIG welding. However, in the middle of Fig. 6, there are a large volume of data beyond the control limit. These data are concentrated in the middle of the weld, and the range of data beyond the control limit is very large. It can be considered that there is something abnormal in the welding process here.

Abnormal causes can be analyzed by the PCA load matrix. The data analysis methods library provided by Scikit-learn [14] were used to calculate the load matrix of principal components, and the five spectral lines with the largest load among the four principal components were selected and drawn as Fig. 8, in which the spectral lines of oxygen elements were marked red and the spectral lines of other elements were marked blue. As shown in Fig. 8, three of the five spectral lines with the largest load on the first PC came from O element. Two of the five spectral lines with the largest load on the second principal component came from the O element. Three of the five spectral lines with the largest loading on the third and fourth principal components also came from O element. It can be seen that the characteristic spectral line of oxygen has the largest load on the first, second, third, and fourth principal components. Combined with the actual welding situation (as shown in Fig. 7), it can be judged that the oxygen content of the weld is too high.

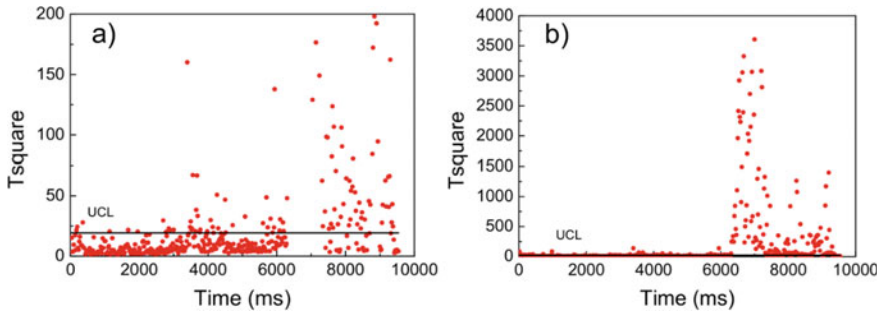
The control chart of welding gap variation was calculated according to the above steps, as shown in Fig. 9, and the photograph of the weld was shown in Fig. 10.

In Fig. 9a, it can be found that the  $T^2$  value of few sampling points exceeded the control limits before 3000 ms. However, there was a large number of points exceeded the control limits between 3000 and 6200 ms. After 6200 ms, the  $T^2$  value of the point increased significantly and exceeded the control limits rapidly. In Fig. 9b, due to the axis scale, the upper control limits ( $\chi_6^2(0.01) = 16.812$ ) almost overlapped with the coordinate axis. This control limits for the indices was calculated on a confidence limit of 99%. Some of these samples exceeded the control limits, these data are classified as the abnormal data.

By comparing the appearance of the weld (Fig. 10), it could be confirmed that after 2000 ms, due to the sudden change of the gap, there was an underfill defect. After 6200 ms, due to the sudden change of gap, the lack of penetration occurred.



**Fig. 8** Five spectral lines with the largest load in the first four principal components. **a** First PC **b** second PC **c** third PC **d** fourth PC

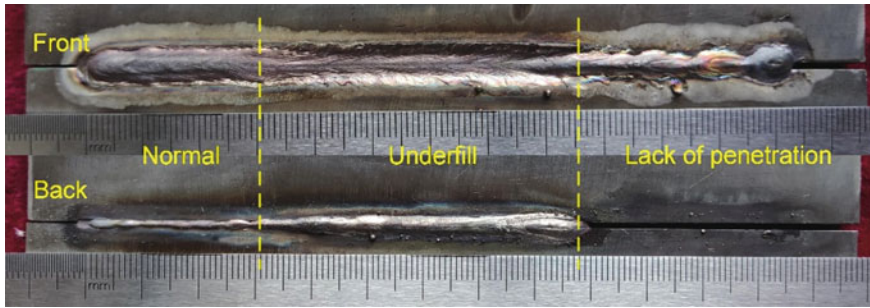


**Fig. 9** Control charts of underfill and lack of penetration. **a**  $T^2 \leq 200$  **b**  $T^2 \leq 4000$

## 4 Conclusions

In this paper, the detection of oxide inclusion and the lack of penetration was carried out. The conclusions are as follows:

- (a) For highly correlated spectral data, PCA before MSPC can reduce the dimensionality of the original data, reduce the computation and the requirements on computing power, which is conducive to the realization of online monitoring.



**Fig. 10** Appearance of underfill and lack of penetration

- (b) Multivariable statistical process control can accurately identify the adverse effects of oxide inclusions and gap changes on the welding process. As a data-driven approach, MSPC can simultaneously detect multiple defects in the welding process, but it cannot classify them.

**Acknowledgements** This work was supported by the National Key Research and Development Program of China under the Grant (No. 2018YFB1107900), Shandong Provincial Natural Science Foundation, China, under the Grant (No. ZR2017MEE042), Shandong Provincial Key Research and Development Program under the Grant (No. 2018GGX103026).

## References

1. You D, Gao X, Katayama S (2014) Multisensor fusion system for monitoring high-power disk laser welding using support vector machine. *IEEE Trans Industr Inf* 10(2):1285–1295
2. Steen WM, Mazumder J (2010) *Laser material processing*. Springer, London
3. Blackmon DR, Kearney FW (1983) A real time approach to quality control in welding. *Weld J* 3739(8):37
4. Shea JE, Gardner CS (1983) Spectroscopic measurement of hydrogen contamination in weld arc plasmas. *J Appl Phys* 54(9):4928–4938
5. Nilsen M, Sikström F, Christiansson A-K et al (2017) Vision and spectroscopic sensing for joint tracing in narrow gap laser butt welding. *Opt Laser Technol* 96:107–116
6. Ma J, Kong F, Carlson B et al (2013) Two-pass laser welding of galvanized high-strength dual-phase steel for a zero-gap lap joint configuration. *J Mater Process Technol* 213(3):495–507
7. Sibillano T, Ancona A, Berardi V et al (2006) A study of the shielding gas influence on the laser beam welding of AA5083 aluminum alloys by in-process spectroscopic investigation. *Opt Lasers Eng* 44(10):1039–1051
8. Ribic B, Burgardt P, DebRoy T (2011) Optical emission spectroscopy of metal vapor dominated laser-arc hybrid welding plasma. *J Appl Phys* 109(8)
9. Harooni M, Carlson B, Kovacevic R (2014) Detection of defects in laser welding of AZ31B magnesium alloy in zero-gap lap joint configuration by a real-time spectroscopic analysis. *Opt Lasers Eng* 56:54–66
10. Colombo D, Colosimo BM, Previtali B et al (2012) Through the optical combiner monitoring in remote fiber laser welding of zinc coated steels. In: *High power laser materials processing: lasers, beam delivery, diagnostics, and applications*

11. Zhang Z, Zhang L, Wen G (2019) Study of inner porosity detection for Al-Mg alloy in arc welding through on-line optical spectroscopy: correlation and feature reduction. *J Manuf Process* 39:79–92
12. Lacroix D, Jeandel G (1997) Spectroscopic characterization of laser-induced plasma created during welding with a pulsed Nd: YAG laser. *J Appl Phys* 81(10):6599–6606
13. Jin X, Fan J, Chow TWS (2019) Fault detection for rolling-element bearings using multivariate statistical process control methods. *IEEE Trans Instrum Meas* 68(9):3128–3136
14. Pedregosa F et al (2011) Scikit-learn: machine learning in python. *JMLR* 12:2825–2830
15. Johnson RA, Wichern DW (2008) *Applied multivariate statistical analysis*