Computers in Industry 94 (2018) 75–81

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.elsevier.com/locate/compind

Real-time monitoring of high-power disk laser welding based on support vector machine

Juequan Chen[a], Teng Wang[a,*], Xiangdong Gao[b], Wei Li[c]

[a] School of Computer, South China Normal University, Guangzhou, 510631, China
[b] School of Electromechanical Engineering, Guangdong University of Technology, Guangzhou, 510090, China
[c] School of Physics and Telecommunication Engineering, South China Normal University, Guangzhou, 510631, China

ARTICLE INFO

Article history:
Received 9 May 2017
Received in revised form 22 September 2017
Accepted 11 October 2017
Available online xxx

Keywords:
Laser welding
Support vector machine
Real-time monitoring
Feature selection
Welding quality

ABSTRACT

In this paper, an efficient quality monitoring system for monitoring high-power disk laser welding in real time was developed. Fifteen features of laser-induced metal vapor plume and spatters were extracted and support vector machine was adopted to establish a classifier to evaluate the welding quality. Feature selection method was employed to choose suitable features. The experiment results demonstrated that this method had satisfactory performance and could be applied to real-time monitoring application.

1. Introduction

Laser welding is one of the most important welding techniques due to its advantage in production and quality [1,2]. However, high-power laser welding is a complex process with highly unstable heat transfer, which makes it formidable to monitor. Consequently, effective and accurate method for monitoring the laser welding process plays a very important role in guaranteeing the process stability and the product quality.

Among various monitoring methods, computer aided analysis technique has attracted considerable attention recently [3,4]. Especially since the rapid development in machine learning and pattern recognition, many approaches have been applied to monitor manufacturing process. Some examples are as follows. Wu et al. [5] investigated the relationships between the weld penetration and keyhole characteristics with different welding conditions during variable polarity plasma arc welding. Particle swarm optimization and adaptive network were employed to establish a system to predict the joint penetration from keyhole images. Wan et al. [6] introduced an efficient quality monitoring system for small scale resistance spot welding based on dynamic resistance. Back propagation neural network were used to estimate the weld quality and showed a better performance than regression analysis.

Support vector machine (SVM) is a supervised learning model based on the statistical learning theory developed by Vapnik et al. [7,8]. As an increasingly popular tool for classification and regression [9], SVM has been widely applied in modern welding industry. Zhang and Chen [10] established a SVM classifier to automatically evaluate seam quality in gas tungsten arc welding. You et al. [11] introduced an innovative approach to perform laser welding process monitoring and weld defect diagnosis, SVM classification model was built to effectively identify weld defects. Mekhalfa and Nacereddine [12] used SVM to automatically classify four types of weld defects in radiographic images. Nevertheless, the performance of SVM is highly effected by the features used. Lack of precise priori knowledge and choosing features blindly may result in redundant or irrelevant features.

In this work, images of laser-induced metal vapor during high-power disk laser welding were captured by high-speed camera. After image processing, fifteen features of the metal vapor plume and spatters were extracted. SVM classifiers were generated to automatically evaluate the welding quality. Feature selection method was employed to identify the optimal feature subset. The experiment results were recorded and analyzed afterwards.

The rest of this paper is organized as follows. Section 2 presents the basic principles of SVM and feature selection. Section 3 describes in detail the four steps of our experiment, including welding, image processing, feature extraction, feature selection

* Corresponding author.
E-mail address: towangteng@263.net (T. Wang).

https://doi.org/10.1016/j.compind.2017.10.003
0166-3615/

and classification. Section 4 analyses the experiment results and conclusions are presented in Section 5.

2. Methods

2.1. SVM

The basic idea of SVM can be briefly described as follows [13–15].

Given a set of training data:

$$T = \{(x_i, y_i) | x_i \in R^d, y_i \in \{-1, 1\}, i = 1, \dots, N\}$$

where N represents the number of training data and d denotes the number of dimensions of input data, SVM attempts to identify the hyperplanes that separate data points of different classes. The separating hyperplane in multidimensional space is defined as:

$$w \cdot x + b = 0, w \in R^d, b \in R$$

If any hyperplane that satisfies this equation exists, data set T is linearly separable.

The margin of a separating hyperplane is calculated as $2/||w||$, where $||w||$ is the Euclidean norm of w. As the generalization ability of a separating hyperplane is determined by its margin, for linearly separable case, the support vector algorithm simply looks for the separating hyperplane with largest margin. This can be formulated as follows:

$$\min_{w,b} \frac{1}{2} ||w||^2$$
$$\text{s.t.} y_i(wx_i + b) - 1 \geq 0, i = 1, 2, \dots, N$$

Typically, this constrained optimization problem, which is referred as the primal optimization problem, can be written by introducing its Lagrange function:

$$L(w, b, \alpha) = \frac{1}{2}||w||^2 - \sum_{i=1}^{N} \alpha_i y_i (w \cdot x_i + b) + \sum_{i=1}^{N} \alpha_i$$

where $\alpha = \{\alpha i | \alpha i \geq 0,\ i = 1, \dots, N\}$ Lagrange multiplier vector.

And thus the dual problem is:

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j (x_i \cdot y_j) - \sum_{i=1}^{n} \alpha_i$$
$$\text{s.t.} \sum_{i=1}^{n} \alpha_i y_i = 0$$

This problem can be solve by using quadratic programming method, and the solution α^*can be used to calculate $w^*$and b^*. Finally, the optimal separating hyperplane is:

$$w^* \cdot x + b^* = 0$$

and the SVM classifier function can be written as:

$$f(x) = \text{sign}(w^* \cdot x + b^*)$$

The method above is linear SVM, which uses the training data set to generate an optimal separating hyperplane. But in many cases, the data points are not linearly separable, and thus the function of the optimal separating surface may be nonlinear. SVM can be extended to handle this kind of problems by using a method called kernel trick. The basic idea is mapping the input data into a higher dimensional space called feature space and then performing linear classification in that higher dimensional space. These can be done implicitly by replacing the inner product with kernel function.

Several kernel functions have been explored, such as polynomial kernel, sigmoid kernel and radial basis function(RBF) kernel. RBF kernel is one of the most widely used kernel functions, usually in the Gaussian form:

$$k(x, x') = \exp(-\frac{||x - x'||^2}{2\sigma^2})$$

The parameter σ controls the radial range of the function.

Because of the generally satisfactory performance and simple parameter setting, RBF kernel function is applied in this study.

2.2. Feature selection

Feature selection [16–18] is an important and widely employed technique in the field of machine learning and pattern recognition. The purpose of feature selection is to choose a small subset of features from the original data set according to a certain evaluation criterion, which usually results in better performance, such as higher classification accuracy, lower computational cost and better model interpretability. As a practical approach, feature selection has been widely applied to many fields [19–21].

Generally, feature selection algorithms fall into three categories: the filter model, the wrapper model, and the hybrid model. The filter model relies on measures of the general characteristics of the training data to evaluate and select feature subsets so that the learning algorithm does not involved in this phase. The wrapper model requires a predetermined learning algorithm and uses the prediction accuracy to determine the quality of selected features, which means it searches for more suitable features for the learning algorithm and thus improve the prediction accuracy, but it also tends to be more computationally expensive than the filter model. The hybrid model attempts to combine the advantages of both previous models by exploiting different evaluation criteria in different search stages.

As shown in Fig. 1, a typical feature selection algorithm consists of three basic steps:

1) Subset Generation: In the first step, subset generation, a candidate feature subset will be chosen based on a given search strategy. Many strategies have been explored, some elementary ones are: (a) complete search, which guarantees to find the most suitable feature subset according to the evaluation criterion used; (b) random search, which means generating

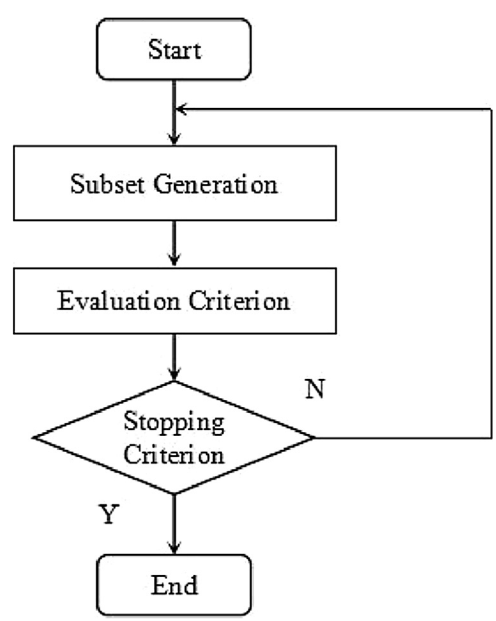

Fig. 1. Flow chart of a typical feature selection algorithm.