The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

A cognitive system for autonomous robotic welding

Georg Schroth, Ingo Stork genannt Wersborg, and Klaus Diepold

Abstract— **Currently, there is a high demand for autonomous industrial production systems. This paper outlines the development of a cognitive system for autonomous robotic welding. This system is based on dimensionality reduction techniques and Support Vector Machines, allowing the system to learn to separate between acceptable and unacceptable welding results within one batch, and to transfer this ability to a batch with different workpiece properties. It does not aim at a complete and general relationship between all process variables and result quantities, since it has been demonstrated that this is not necessary to reduce significantl the costs of calibrating the welding system. The main objective is to examine a cognitive system that stabilizes robotic welding processes by learning how to improve at least one process steering variable. In order to evaluate and improve the cognitive system, an extensive experimental setup is realized and described. The ability to learn and autonomously adapt to changes in workpiece properties allows the system to reduce the time an expert needs, and relaxes the requirements with respect to workpiece tolerances.**

I. INTRODUCTION

Robotic welding combines several similar welding techniques, such as arc welding, laser-hybrid welding, or laser beam welding. This paper focuses on laser beam welding for demonstration purposes. However, the cognitive system that is examined in this paper applies to many different kinds of robotic welding processes that may be steered by influencin at least one process parameter. In this application, robotic welding and particularly welding with laser light permit superior processing of workpieces compared to many other tools by allowing a high production rate, fl xibility and quality. The possibilities for precision in laser beam welding are very high if the various parameters that are related to system excitation and response can be accurately controlled. Variations in the response of the workpiece during welding are a dominant source of instability in the overall welding process. Laser welding produces acoustic and optical emissions. Detection of these emissions is one way in which the welding conditions can be monitored. Learning about and recognizing the components of these signals, which are diagnostic of specifi fault conditions, offers the possibility of control in order to optimize welding and eliminate weld defects.

Modeling and monitoring the process, which includes aspects such as multiple reflections absorption, heat conduction, melting, vaporization, melt pool movement and solidification has been studied in numerous papers [1], [2], [3], [4], [5]. However, model understanding is not yet sufficien to characterize process behavior reliably and explicitly, which is currently described as chaotic. This paper therefore studies the development of a cognitive approach for robotic welding or laser beam welding in particular.

The quality of the weld can be described using result quantities, such as welding depth, seam width, and the existence of pores. Due to strong process emissions, it is difficul to measure the stability of a weld seam reliably without destroying the workpiece. Therefore, it is important to learn from weld defects and to prevent them before they occur. In addition to many other welding parameters, we concentrate in this work on the most influentia variable in laser beam welding - laser power - to create traceable results. We consider the welding process as a black box, and try to discover relationships between the result quantities and the input values.

This work does not aim to achieve a complete and general relationship between all process variables and result quantities for two reasons. Firstly, we believe that it is not possible to create a general model that describes all of the different kinds of welding processes completely; and secondly, it has been demonstrated that this is not necessary to reduce the time and costs of calibrating the welding system significantl . This is because an expert is not only needed once for the initial setup of a completely new system task. The workpieces are produced as batches with similar properties. However, there may arise differences among the properties within individual batches. These small but noticeable changes can lead to an unacceptable increase in the rejection rate, meaning that the system has to be reconfigured This involves not only the labor costs of an expert, but also downtime for the production line. It is thus of great interest to reduce the recalibration time and allow the system to adapt autonomously to small changes in the properties of a workpiece.

The cognitive approach for an intelligent laser beam welding system can be sub-divided into three main steps. Chapter II describes the learning phase, in which the system identifie the characteristics of the process that allow differentiation between acceptable and unacceptable welds. In the monitoring phase, outlined in chapter III, the system uses the knowledge it has acquired for fault detection. This allows the initialization of the adaption phase, described in chapter IV, in which alterations in the process variables are determined in order to adapt the system to slight changes in the properties of the workpiece. Following this, there is a chapter presenting the detailed experimental simulation and evaluation of the approach.

CoTeSys—Cognitive Technical Systems—LoMAMoS project funded.
G. Schroth, I. Stork genannt Wersborg and K. Diepold are with the Insitute of Data Processing, Department of Electrical Engineering and Information Technology, Technische Universität München, Munich, Germany istork@tum.de, www.cotesys.org

II. LEARNING PHASE

In order for the system to learn how it should deduce the weld quality from sensor measurements, an expert performs an initial setup of the entire welding process, including all of the process variables. The system then starts to weld a workpiece with linearly increasing laser power, as shown in Fig. 1. The optimal laser power, which is previously determined by an expert, corresponds to the mean of the power range. The system thus starts at a power value which is too low for an acceptable welding result, passes through the optimal power value and ends with a laser power that is too high. The resulting workpiece therefore includes all possible welding results that can be obtained by a change in the laser power. At this point, the system has recorded all of the welding results within the applied sensor array. Furthermore, the expert has determined which results are acceptable and which are not so that the system can differentiate between these two classes.

The most simple approach for comparing welds is to compute the Euclidean distance between the individual sensor value vectors. This would compare not only the relevant information, but also the noise and other irrelevant disturbances. Due to the non-linear properties of the welding process, sensor value vectors close to each other in the measurement space do not necessarily indicate similar welding results. Additionally, the amount of data to be compared from the extensive set of sensors would result in an extremely time-consuming task. This can be avoided by extracting features that are known in the analytical model of the welding process.

There are a variety of algorithms that are capable of analyzing the data and measuring important physical properties that are identified such as the size of the melt pool [5]. Unfortunately, these algorithms do not adapt to the data by themselves and have to be reconfigure for each individual welding process. This makes them very specialized and can result in faulty measurements. In addition, most of these algorithms are limited to a few clearly recognizable features, and can only extract a fraction of the information in the data. Many of these algorithms assume that predefine measurements are sufficien for the classificatio of welding results.

A. Dimensionality reduction

Statistic dimensionality reduction techniques can cope with the requirements described. These algorithms can learn the relevant information in the data, and can therefore separate this from the redundancies, disturbances and noise. Techniques for dimensionality reduction can be subdivided into various groups. The main distinction is between linear and nonlinear techniques. Linear techniques assume that the data lies on or near a linear subspace of the high-dimensional space.

1) Linear Dimensionality Reduction: Principal Component Analysis (PCA) is an excellent linear dimension reduction technique [6]. In essence, PCA seeks to reduce the dimensions of the data by identifying a few orthogonal linear combinations, the Principal Components (PCs), of the original variables with the largest variance. How to obtain the PCs is discribed in [6]. The number of PCs that should be retained can be determined by firs fixin a threshold λ_0, and retaining only those eigenvectors whose corresponding eigenvalues are greater than λ_0.

2) Nonlinear Dimensionality Reduction: As already indicated, the laser beam welding process is known to be highly non-linear, and it is therefore reasonable to apply techniques that are capable of non-linear dimensionality reduction. A linear approach such as PCA is not capable of detecting the true non-linear geometry behind the samples, projecting points that are far away on the manifold to nearby points in the computed feature space. The true relationships between the points, and thus the observations, are not preserved, and it is therefore not possible to classify the observations in this feature space.

The need for algorithms that are capable of handling non-linear data has led to the creation of several approaches, such as Laplacian Eigenmaps [7], the Locally Linear Embedding Algorithm (LLE) [8] and the Isomap Algorithm [9], to list the most renowned. Nonlinear techniques for dimensionality reduction can be split into two main categories: techniques that attempt to preserve the global properties of the original data in low-dimensional representation (Isomap); and techniques that attempt to preserve the local properties (LLE, Laplacian Eigenmaps). Due to limited space and its fina application (chapter V), we will only describe the Isomap in detail.

Isometric feature mapping, or Isomap, seeks to map from $D-$dimensional observation space X to a low-dimensional feature space Y. The core idea is to fin an efficien way of computing the true geodesic distance between observations, given only their Euclidean distances in the high-dimensional observation space. The Isomap procedure consists of three main steps. As a preparation for computing manifold distances, Isomap firs constructs a topology-preserving network by determining which points are neighbors on the manifold M, based on the Euclidean distances between pairs of points. Two methods are used to connect each point to all points within a f xed radius ε, or to all of the k nearest neighbors.

Given this network representation, the geodesic distance between any two nodes in the network is computed using Dijkstra's shortest path algorithm. This provides a good approximation, and results show that correlation to the true manifold distances is very high with sufficien observations, but tends to overestimate by a constant factor due to the discretization introduced by the graph.

Finally, using these manifold distances, we are able to construct a global, geometry-preserving map of the observations in a low-dimensional Euclidean space using Multidimensional Scaling (MDS). MDS maps the high-dimensional data representation onto a low-dimensional representation while retaining the pairwise distances between the data points as far as possible. The quality of the mapping is expressed in the stress function, which is a measure of the error between the pairwise distances in the low-dimensional and high-

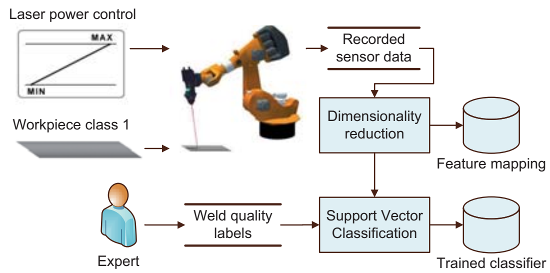

Fig. 1: Learning phase

dimensional representations of the data. A simple example of a stress function is the raw stress function

$$\Phi(Y) = \sum_{ij} (||x_i - x_j|| - ||y_i - y_j||)^2, \tag{1}$$

in which $||x_i - x_j||$ is the distance between the high-dimensional data points x_i and x_j, and $||y_i - y_j||$ is the distance between the low-dimensional data points y_i and y_j. Minimization of the stress function can be achieved using various methods, such as the eigenvalue decomposition of a pairwise distance matrix.

However, Isomap is known to fail on manifolds with "holes" and intrinsic curvature [10]. Furthermore, Isomap is sensitive to an effect called "short-circuiting". If neighborhoods are define too widely, data points can be selected that do not belong to it. Therefore, links are generated in the neighborhood graph which allow shortcuts in the computation of the geodesic paths. A single faulty shortcut can lead to a significan error in the estimation of numerous geodesic paths, and thus to incorrect mapping. Hence, the neighborhood size has to be carefully determined.

3) Comparative view: In [10], an evaluation of Isomap and LLE was performed using both artificia datasets and real datasets representing a variety of application domains. Experiments on the artificia data showed that non-linear algorithms considerably outperform linear techniques such as PCA. In all the artificia datasets, Isomap created more exact mappings than both LLE and Laplacian Eigenmaps. For the natural datasets, the results of various papers differ significantl . An experimental evaluation of the algorithms for this specifi application was therefore necessary. These approaches are combined with PCA to account for prewhitening and to improve the performance of the algorithms by linearly reducing the dimensions in advance. This hierarchical structure should not only reduce computational complexity, but also account for the different data relationships among the sensors.

As described in [11], the Isomap algorithm has two computational bottlenecks. The firs is calculating the $N \times N$ shortest-paths distance matrix D_N. Using Floyd's algorithm, this is $O(N^3)$; it can be improved to $O(kN^2 logN)$ by implementing Dijkstra's algorithm with Fibonacci heaps, where k is the neighborhood size. The second bottleneck is the MDS eigenvalue calculation, which involves a full $N \times N$ matrix and has a complexity of $O(N^3)$. Random sampling techniques provide a powerful alternative for approximate spectral decomposition and only operate on a subset of the matrix. Recently, the Nyström approximation [12] has been shown to be most applicable for Isomap; for details, see [13].

B. Support Vector Machines

Comparison of the welds can now be performed by computing the Euclidean distance in the feature space, which can be compared with the techniques used at face recognition. However, the use of a Support Vector Classifie (SVC) [14] can facilitate a more reliable identificatio of the weld quality. Because still, non-linear relations can be expected in the feature space, a classifie should not base the decision boundary on linear features alone. Rather, the original attributes x shall be mapped onto non-linear features, e.g. x, x^2, x^3. Since the algorithm can be written entirely in terms of the inner products of input vectors, the so-called 'kernel trick' can be used to allow SVCs to learn in the high dimensional feature space given by $\Phi(x) = \left[x, x^2, x^3\right]^T$ without ever having to fin or represent vectors $\Phi(x)$ explicitly. In our application, the well-known Gaussian kernel, which corresponds to an infinit dimensional feature mapping ϕ, will be used:

$$K(x_1, x_2) = \exp\left(-\frac{\|x_1 - x_2\|^2}{2\sigma^2}\right). \tag{2}$$

Here, σ^2 corresponds to the variance of the Gaussian distribution. This classifie is trained with the initially recorded and labeled gradient data.

III. MONITORING PHASE

In the subsequent welding processes, the system can now continuously check whether or not the results are within the acceptable range. This is achieved by embedding the current sensor reading into the previously learned feature space via an out-of-sample extension [13]. Since every welding process differs from every other, the optimal or sufficien number of dimensions for the feature space has to be discovered through trials or the intelligent guesses of a welding expert. However, section V will demonstrate that this number may remain unchanged for several specifi welding processes, either within one batch or between different batches. Subsequently, the embedded feature data is passed to the SVC for an indication of the weld quality based on the probability of it belonging to a certain class. This allows a supervisor to observe the process continuously, or even to insert a control loop that can adjust the laser power to fine-tun the weld results.

If there are multiple consecutive bad result detections, this indicates that a batch change has occurred and the properties in the batches have changed to an unacceptable degree. The system then has to adapt to these changes.

IV. ADAPTION PHASE

When a new batch is detected, a fresh laser power gradient is applied to the new workpiece, as shown in Fig. 2. Again, all of the possible weld variations for this new batch, which

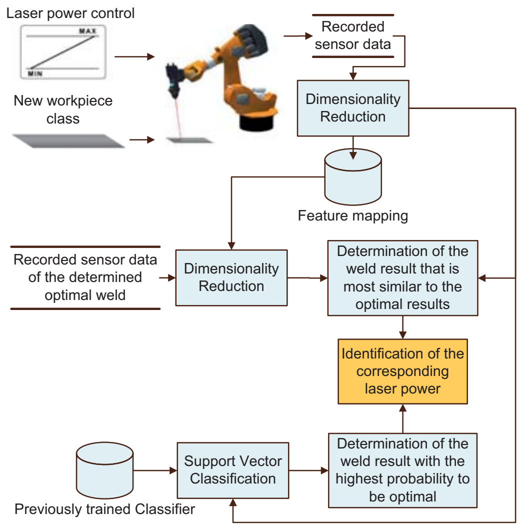

Fig. 2: Adaption phase

result from a change in the laser power, can be recorded by the sensor array, and a new feature space is constructed with the aid of dimensionality reduction techniques. The sensor values from the optimal welding result that was initially determined are now compared to all of the possible new results within this new feature space. This exploits the assumption that minor changes in the workpiece properties result in comparable welds. The system can therefore use the relationships that are learned from certain workpiece properties to make generalizations that can be adapted to the new properties. Hence, it is possible to identify directly the welding result that is most similar to the optimal one, and also to determine the used laser power. Alternatively, the fina comparison process can be carried out with the aid of SVC. A comparison of both approaches is given in section V-B.

The system can thus re-determine the power value that leads to a weld which is most similar to the optimal value within one trial. Furthermore, this complies with the desire that the initial learning does not require a huge dataset by concentrating on a local patch of the general laser welding model.

V. EXPERIMENTS AND RESULTS

In order to evaluate and improve the approaches that are described in chapters II to IV, an experimental setup that includes a largely redesigned and extended process monitoring system has been developed. In the following experiments, the thickness of the workpiece (WP) is define as the property that changes with the batches, since such variation is very common and can be reproduced reliably. The performance of the approaches has been evaluated using six different WP sizes, although only the adaption process from 0.8 mm to 1.0 mm and vice versa is described here. This validates the capability of handling a change of 20%.

A. Experimental setup

The experimental setup includes a six-axis articulated robot carrying a laser processing head as well as parts of the sensor system. During the welding, the robot, mounted at the table, moves the processing head linearly over the WP at a constant speed. The monitoring system includes three photo diode sensors integrated in the processing head, which measure the electromagnetic radiation of the process at three wavelengths. This allows the system to observe the metal vapor radiation, the temperature and the degree of laser back reflection representing state-of-the-art process monitoring. The system is extended by directed and undirected microphones, as well as two solid-borne acoustic sensors mounted on the WP, which record the acoustic emissions resulting from the vaporization process. These scalar recordings are Fourier transformed so that a high-dimensional vector in the frequency domain is generated for each individual acoustic sensor at 1kHz. In addition, the welding is observed by a video camera, which takes images at 1 kHz.

B. Experimental simulation

As described in the previous chapter, the central idea behind this cognitive approach to welding is to ascertain features in the measurement data that allow description of the quality of the weld, and thus compare different weld results with each other. This would allow us to calculate the laser power that results in comparable weld results for workpieces of different properties. These features are not necessarily physical properties, but are determined by dimensionality reduction techniques that best describe the variance with respect to the weld quality, and therefore provide most detailed information about the process.

The extraction of relevant features is performed on the video data first In order to allow analysis, a data matrix X of 25,600 rows and 7,000 columns is created. Despite this, the enormous size of the data matrix would result in an unfeasible eigenproblem, since the resulting covariance matrix has dimensions of 25,600 x 25,600. This can be avoided by computing $X^T X$, rather than using the covariance matrix XX^T, due to the fact that the eigenvalues of both products are identical. This reduces the size of the eigenproblem to 7,000 by 7,000. The resulting eigenvectors can be transformed to a matrix representation in order to illustrate them as images. At eigenvectors corresponding to lower eigenvalues, high frequency features can be observed, including noise and disturbances. Hence, a dimensionality of around 200 should be sufficien to represent the data. Investigating the dimensionality reduced (embedded) data, we can clearly identify strong but non-linear correlations along the dimensions. Using the techniques presented in section II-A.2, these redundancies can be further reduced. Most importantly, by unfolding the nonlinear manifold onto a linear plane, the true relationships among the data vectors can be represented through Euclidean distances.

Because the manifold does not imply holes or intrinsic curvature (see section II-A.2), the Isomap algorithm seems an appropriate choice from an analytical point of view.

However, a detailed experimental evaluation of Isomap, LLE, and Laplacian Eigenmaps revealed Isomap as superior to the alternatives with respect to computational complexity, especially for the out-of-sample extension, and with lower residual variance at the same number of embedding dimensions. Hence, only the results of the Isomap algorithm are presented here.

Over 100 validation experiments showed that a neighborhood size of 10 is optimal for the video data mapping, as judged by welding experts. This neighborhood size can prevent the problem of short-circuiting. The variance is determined from the covariance matrix of the embedded data. It was possible to reduce the dimensionality further, from 200 with PCA to approximately 20 using Isomap.

Regarding the air-borne acoustic data, it was unfortunately difficul to identify any trend that correlated to the quality of the weld [13]. However, this does not mean that the air-borne acoustic data is of no use for the general comparison. Rather, it is necessary for improving either the preprocessing or for altering the experimental setup in order to increase the signal-to-noise ratio.

The solid-borne acoustic data revealed a strong relationship between the frequency distribution and the quality of the weld. However, the PCA embedding was relatively noisy, and the number of dimensions were reduced to a very limited amount. On the other hand, Isomap was capable of understanding the nonlinear relations and reduced the dimensions from 1,000 to 10. Unfortunately, the information that was provided by the solid-borne acoustic data was ambiguous because the magnitudes did not continuously increase with the laser power, but decreased after a certain point.

The ambiguity can be dissolved by combining the data from all information sources, including video and the three-diode sensors. No further dimension reduction is applied to the resulting 33-dimensions feature vector, since the sources imply little redundancy and the required computational complexity would not result in significan advantages. It is thus possible to reduce the dimensionality of the input vector from 26,603 dimensions to only 33 dimensions by splitting the data based on sensor type, by performing prewhitening to eliminate a huge portion of the noise, and by findin nonlinear relations with the aid of Isomap. These 33 dimensional feature vectors represent the essential information on the current state of the welding process, and should allow differentiation between the welding results and thus the welding quality.

Finally, the feature vectors of the workpiece are used to learn the embedding, and the feature vectors of the out-of-sample mapped workpiece are compared using a simple Euclidean distance in the feature space. For each vector of the WP 1 with 1.0 mm thickness, the corresponding welding result on the WP 2 with 0.8 mm thickness is determined by determining the vector with the smallest Euclidean distance. This process corresponds to the nearest neighbor search for the dimensionality reduction techniques.

As described above, the high dynamics in the welding process necessitate comparison not only of the feature vector of one time step with all others, but also the use of multiple consecutive time steps. In order to allow fast recognition of errors and to ensure reliable results, a frame of 40 feature vectors corresponding to 1.67 mm on the workpiece is suggested. The resulting relationships are shown in Fig. 3. The convexity of the curve impressively illustrates the nonlinear relationships between two workpieces of different thickness. As described in chapter II, the position on the workpieces corresponds linearly to the applied laser power. An identical laser power gradient has been applied on both workpieces.

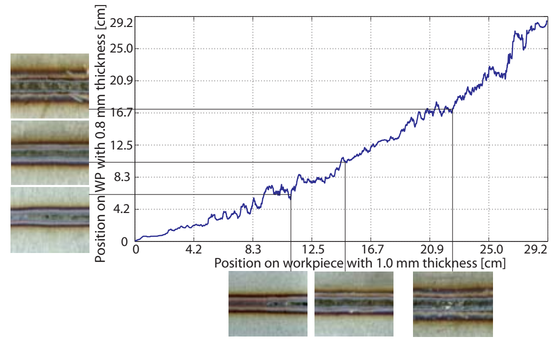

Fig. 3: Relation of similar welding results on workpieces with different thickness vs. the applied laser power

Three characteristic points of WP 1 plotted, and the identifie regions on WP 2 are shown. At the firs point, the laser power is barely sufficien to achieve a full penetration weld. This point is extremely difficul to determine because even very small local properties in the workpiece can alter the outcome significantl . For this reason, it is impressive that the relationship is precise enough to fin this region on WP 2. The second characteristic point marks a region of optimal weld quality on WP 1. This is characterized by concave seam geometry, a smooth seam width, and an absence of pores or splatters. The region identifie on WP 2 complies with all these constraints, and is located in approximately the middle of the comparable results. Hence, if an expert has determined a region of optimal weld quality on WP 1, this figur proves that the system is capable of findin an area of comparable weld results, and can thus determine the necessary alteration of laser power to achieve this optimal result on new batches of workpieces. The third point characterizes the upper limit of the acceptable weld results, as the firs small splatters can be observed. Again, the corresponding region on WP 2 is localized at a very similar welding result.

The physical interpretation of the relationship at the beginning and end of the workpiece confirm the results. At the very low laser powers at the beginning of both workpieces, the results must be roughly the same since only conduction welding occurs and the surface of the material is processed, which should be almost identical. At the very high laser powers at the end of both workpieces, the results saturate at a certain limit. Hence, the results become more and more comparable, as the relationship in Fig. 3 shows. However, the saturation limit of WP 2 is slightly lower than that for

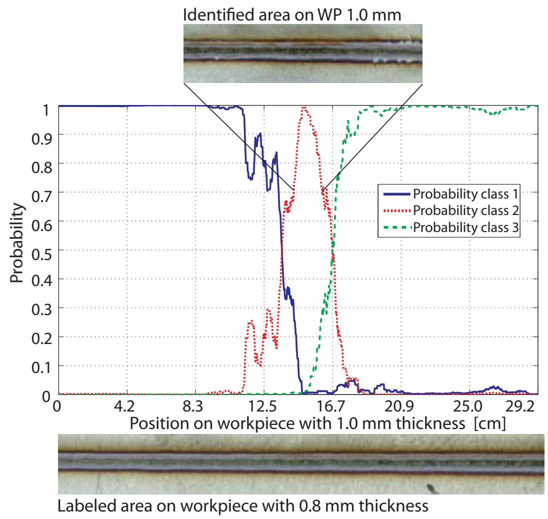

Fig. 4: Support Vector Classificatio of good welding results

WP 1, which explains why the relationship does not end in the upper right corner of the graph.

Although the results seem very satisfying, it is worth investigating a more sophisticated approach for comparing the weld results than using the Euclidean distance in the feature space. In the approach based on Support Vector Classificatio (SVC), described in chapter II-B, the welding results that an expert determines to be optimal are used to categorize all welding results into three classes. The firs class includes all welds that result from a laser power that is too low; class two contains all optimal welds; and class three has all welds with a laser power that is too high. The 33-dimensional feature vectors of the initial workpiece are used to allow the SVC to fin the hyperplane that maximizes the margin between classes. Because the application of a continuous laser power gradient results in continuously changing weld results, there is little margin between the three classes, which makes the classificatio task difficult In order to cope with this problem, *C*-Support Vector Classificatio is applied. This regularization approach is capable of dealing with outliers so that the optimal hyperplanes can be found in continuous transitions. For the kernel function, the Gaussian kernel (also known as the radial basis function (RBF)) is applied. In order to identify the optimal parameters for classification a grid-search on the parameters is performed where the parameters are set in an exponentially growing order.

Fig. 4 shows the probability of belonging to a certain class against the position on the workpiece (WP). Here, a range on WP 2 is labeled as belonging to class two, whereas the other areas belong to classes one and three. The area on WP 1 that is identifie as having a probability higher than 70% of belonging to class two is illustrated at the top of Fig. 4.

This area corresponds well with the region labeled as optimal by an expert. However, for the described procedure, only the position of maximum probability has to be determined in order to calculate the necessary alteration of laser power to achieve optimal welds on a new batch of workpieces.

In summary, this approach is capable of robustly identifying and learning both good and faulty weld results on workpiece batches with unknown but similar properties to the initial workpiece batch. The major advantage of the SVC approach is that it does not require the use of multiple consecutive feature vectors to facilitate a reliable classification

VI. CONCLUSIONS AND FUTURE WORKS

In this paper, a cognitive system for autonomous robotic welding is developed and verifie for laser beam welding. This is capable of efficientl learning to separate between acceptable and unacceptable welding results within one batch of workpieces, and transferring this ability to a batch of different workpiece properties. For this purpose, two techniques - dimensionality reduction and Support Vector Machines, and also their hybridization - have been applied. In order to improve and verify the approach, an extensive experimental setup has been realized and described, with the details given in respect to the sensors used, the data that is record, and the analysis and preprocessing of the data. In contrast to previous efforts, this system is capable of autonomously and reliably adapting the welding process to changes in the workpiece properties. Hence, the intended objective, which is to reduce significantl the time of the expert in the recalibration process, has been achieved. Future work will concentrate on extending the system further to input additional variables to the laser power.

REFERENCES

[1] M. Jaeger and F. Hamprecht. Principal Component Imagery for the Quality Monitoring of Dynamic Laser Welding Processes. *IEEE Transactions on Industrial Electronics*, 56:1307–1313, 2009.

[2] M. Zhou, W. Liu, and L. Wan. Multi-information Fusion and Identificatio System for Laser Welding. *ISNN*, LNCS 5551:986–992, 2009.

[3] M. Kogel-Hollacher. *Sensorik und Datenauswertung zur industriellen Prozessüberwachung beim Lasersstrahlschweien*. PhD thesis, Technische Universität Berlin, April 2008.

[4] J. Müller-Borhanian, C. Deininger, F.H. Dausinger, and H. Hugel. Spatially resolved on-line monitoring during laser beam welding of steel and aluminum'. *Proc. of the 23rd ICALEO 2004*.

[5] S. Kaierle, P. Abels, and C. Kratzsch. Process monitoring and control for laser materials processing-an overview. *Proceedings of Lasers in Manufacturing*, 3:101–105, 2005.

[6] I.T. Jolliffe. *Principal component analysis*. Springer-Verlag New York, 1986.

[7] M. Belkin and P. Niyogi. Laplacian eigenmaps for embedding and clustering. *Advances in Neural Processing Systems*, 14:585–591, 2002.

[8] S.T. Roweis and L.K. Saul. Nonlinear Dimensionality Reduction by Locally Linear Embedding, 2000.

[9] J.B. Tenenbaum. Mapping a manifold of perceptual observations. *Advances in Neural Info. Proc. Sys.*, 10:682–687, 1998.

[10] L. van der Maaten, E. Postma, and H. van den Herik. Dimensionality Reduction: A Comparative Review. 2007.

[11] V. de Silva and J.B. Tenenbaum. Global versus local methods in nonlinear dimensionality reduction. *Advances in Neural Information Processing Systems*, 15:705–712, 2003.

[12] C. Williams and M. Seeger. Using the Nystrom method to speed up kernel machines. *Advances in Neural Information Processing Systems*, 13(14):682–688, 2001.

[13] G. Schroth. A cognitive approach and experimental simulation for laser beam welding. Master's thesis, Diplomarbeit, Technische Universität München, 2008.

[14] J. Shao and et al. An intelligent system for the on-line quality inspection of micro laser spot welding processes. In *J. Phys.: Conf. Ser. 76*, 2007.