

Joint Identification of Optical and Electro-Mechanical Transfer Functions in Laser Scanners with Application to Remote Ablation Cutting

2nd International Conference on Automation, Control, Engineering and Computer Science (ACECS-2015)
Proceedings of Engineering & Technology (PET)

Steffen Sonntag and Ulrich Vogl
OTH Amberg-Weiden

Kaiser-Wilhelm-Ring 23, 92224 Amberg, Germany
Email: U.Vogl@ieee.org

Bernhard Kiesbauer
ARGES GmbH

Werk 4, 92442 Wackersdorf, Germany
Email: kiesbauer@arges.de

Abstract—Laser scanners basically deflect a laser beam in 2 dimensions by electro-mechanically driven mirrors. Due to the mechanical nature of galvanometers (so-called galvos) and mirrors the dynamic is limited. For most applications however, dynamic enhancement without loss of precision is desirable. To boost the dynamics, a precise model is indispensable. Furthermore the optical (geometric) mappings need to be considered. To estimate both, we use a composite mapping algorithm. As stimulus signals plane filling (fractal) curves prove to be very useful. Once a reliable dynamic and geometric model is obtained, we can optically rectify the set-point trajectory and apply optimality criteria to find the control sequence for a given contour. This will be demonstrated by a remote ablation cutting (RAC) application, where constant high speed trajectories are crucial for good cutting results.

I. INTRODUCTION

Laser scanner systems have a broad range of applications, including laser eye surgery (LASIK), inscription on various surfaces, laser cutting and welding. Common to all applications is the need of high precision and high speed. Higher speed obviously reduces the production time and thus increases productivity. For the relatively new method of laser cutting, so-called Remote Ablation Cutting (RAC), a (constant) high speed along the cutting contour is even crucial. If speed is too low, the percentage of liquid melt produced is too high (as compared to the vapor portion), which results in a partial re-closure of the kerf. On the other hand, the electromechanical actuators (in principle two galvanometers) needed for highly brilliant laser beams with 1-10 kW power, comprise mirrors with up to 50 mm in diameter, resulting in a high moment of inertia. With the common setup of an analog PID controller for the galvanometer, the dynamics do not meet the requirements. Due to the high speed needed, small details of contours can not be resolved to exact size. Enhancing the dynamics, i.e. increasing the bandwidth of the system, thus equals to rising the resolution and make smaller scales accessible.

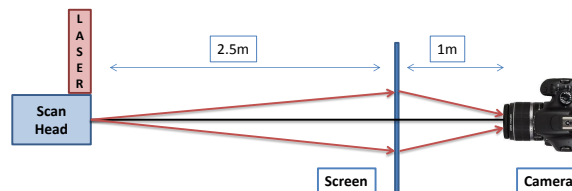


Fig. 1. Geometry of measurement setup.

Standard galvanometer laser scanner systems (galvos) comprise a magnetic axis, which drives the mirror on one side, and an angle sensor at the other. The angle sensor is used to feed the PID controller. At high dynamic operation, however, vibrational modes of the mirror lead to a discrepancy between the angle position as measured by the sensor and the true angle of the mirror [2]. Thus it is necessary to find a model of the total system, which includes the **dynamic** transfer functions from the input voltage of the electric driver to the true mirror angle.

Furthermore, there is a **static** geometric mapping function, which maps the mirror angles from their respective positions to the 2D-plane. In many cases special optical component (e.g. F-Theta objectives) are used. They also contribute to the global static geometric mapping. In addition, the optical transfer function from the 2D-plane to the camera must be taken into account. The measuring setup is shown in Fig. 1. The scan head is driven by a Digital Signal Processor (DSP) using sample rate $f_0 = 48$ kHz. The test patterns are projected onto a screen, which is photographed by a digital camera aligned to the optical axis.

II. SYSTEM MODELING

System modeling can be done by starting from first principles, e.g. solving Maxwell's equations to find an equivalent electrical circuit diagram, and perform mechanical FEM calculations [1]. For better reproducibility of results and also due

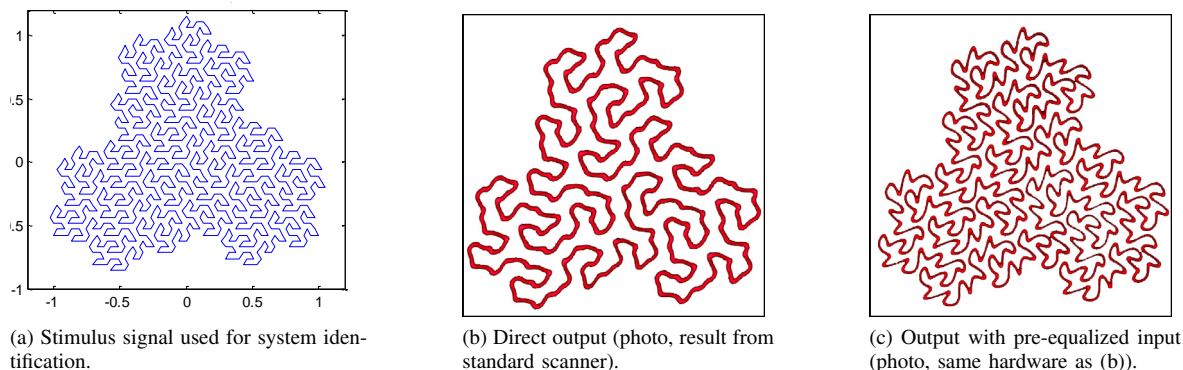


Fig. 2. Triple flow snake curves at different scale resolutions.

to variation in production we prefer here a so-called gray-box approach. This means that the principal structure (e.g. order and special poles of the dynamic model) are fixed, the detailed parameter values however are adapted to measuring results.

A. Stimulus Signal

The choice of the stimulus signal is crucial in system recognition applications. First, broad frequency spectrum is required. Since in our case the system output is a static 2D-figure, we need in addition periodicity and a large number of discriminable measuring dots. Furthermore, approximate homogeneous filling of the 2D-plane is of advantage for estimation of the geometric distortions.

1) *Plane filling periodic curves*: Plane filling curves are known since their discovery by Giuseppe Peano in 1890, [3]. Since we have the additional constraint of periodicity, only few classical curves are suitable here. For example, three Gosper curves [4], can be arranged to form a periodic, 3-symmetric curve. This so-called triple flow snake is shown in Fig. 2a at construction level 3, comprising $3 * 7^3 = 1029$ sample points. At a sample rate of 48 kHz this results in a repetition period of about 21.4 ms. We propose to use the triple flow snake as stimulus (3FS) for several reasons:

- The 3FS has very good spectral behavior, with relatively slow decay for higher frequencies. Due to symmetry, only every 3rd spectral line vanishes.
- As for any plane filling curve, we achieve highest yield of measuring points (and thus best SNR) in the 2D-plane.
- 3FS has good autocorrelation properties with a unique maximum.

The self-similar structure also allows for direct measurement of scale resolution, and thus offers a quantitative means for assessing dynamic-enhancements due to e.g. improved control algorithms. Applying the 3FS stimulus of Fig. 2a to our laser scanner system, the curve displayed in Fig. 2b is observed. Clearly, the scanner can not resolve the 3FS curve with the input resolution (level 3), rather a 3FS curve of construction level 2 is produced.

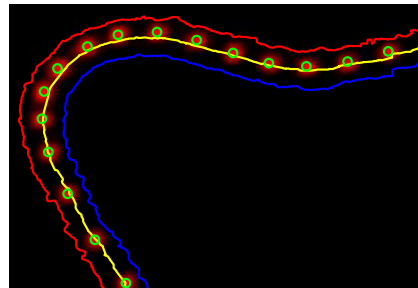


Fig. 3. Retrieving digital signal samples from photographed laser pulses (red dots). Green circles: Detected sample positions. Yellow: Detected contour.

2) *Acquiring the output signal from photo*: By pulsing the laser with the sample frequency of $f_0 = 48$ kHz, we can retrieve the output signal using standard image processing techniques, like the circular Hough transform [5]. A detail of the pulsed curve is displayed in Fig. 3, together with detected sample dots and trajectory. Due to good auto-correlative properties of the stimulus signal, the phase (and thus the starting point) of the signal can then be found easily.

B. The Composite Mapping Algorithm

To identify the optical mapping \mathcal{G} consistently with the dynamic model \mathcal{M} , we apply an iterative algorithm. Let the input samples for the 2D-space be denoted by a matrix $X = (\mathbf{x}_1, \mathbf{x}_2)$, where $\mathbf{x}_i \in \mathbb{R}^n$ are n -dimensional stimulus vectors for each dimension. Similarly, the matrix $Y = (\mathbf{y}_1, \mathbf{y}_2)$ represents the sample points extracted from the photo. The dynamic model \mathcal{M}_a , depending on the parameter vector \mathbf{a} , is then considered a mapping

$$\mathcal{M}_a : \mathbb{R}^{n \times 2} \rightarrow \mathbb{R}^{n \times 2}, \mathcal{M}_a(X) = Y. \quad (1)$$

The parameters \mathbf{a} are constrained to ensure a gain factor of 1 for each dimension of this model. In the same way, the optical distortion function is a map

$$\mathcal{G}_b^{-1} : \mathbb{R}^{n \times 2} \rightarrow \mathbb{R}^{n \times 2}, \mathcal{G}_b^{-1}(Y) = Y'. \quad (2)$$

Typically, second order polynomial functions suffice to model the optical behavior. The algorithm, a variant of a composite mapping algorithm [7], starts then with

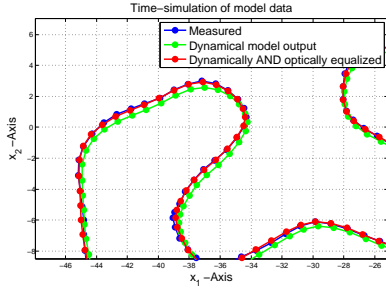


Fig. 4. Example for re-simulated signal (red) in comparison to measured signal (blue). Green: Optical distortion not considered.

- (1) Find $\mathbf{a}^{(0)}$ which minimizes

$$\mathbf{a}^{(0)} = \arg \min_{\mathbf{a}} \|\mathcal{M}_{\mathbf{a}}(X) - Y^{(0)}\|_F, \quad (3)$$

where $Y^{(0)} = Y$ and $\|\cdot\|_F$ represents the Frobenius norm. Next we re-simulate the output signal from this first guess of parameter set $\mathbf{a}^{(0)}$: $Y^{(1)} = \mathcal{M}_{\mathbf{a}^{(0)}}(X)$.

- (2) The following step consists of finding optimal geometric rectification parameters $\mathbf{b}^{(0)}$, such that

$$\mathbf{b}^{(0)} = \arg \min_{\mathbf{b}} \|\mathcal{G}_{\mathbf{b}^{(0)}}^{-1}(X) - Y^{(1)}\|_F \quad (4)$$

is minimized. Again, using $\mathbf{b}^{(0)}$, we rectify the geometry by $Y^{(2)} = \mathcal{G}_{\mathbf{b}^{(0)}}^{-1}(Y)$.

- (3) Now we can iterate step (1) with $Y^{(0)}$ replaced by $Y^{(2)}$.

Typically, convergence is achieved after 5 to 10 times repeating steps (1) and (2), and we end with a model $\mathcal{M}_{\mathbf{a}^{(\infty)}}(X)$, representing the dynamics, and $\mathcal{G}_{\mathbf{b}^{(\infty)}}^{-1}(X)$, corresponding to the total geometric distortions, including the optical transfer function of the camera. To get the rectification map \mathcal{R}^{-1} needed for latter pre-distortion, we seek the static optical transformation \mathcal{T} from the projection screen to the camera. This can be done using standard registration techniques [6]. Thus we have $\mathcal{G} = \mathcal{T} \circ \mathcal{R}$, from which \mathcal{R}^{-1} can be calculated.

Since all signals are time-periodic, there is no need to consider initial states in the dynamic model \mathcal{M} , and the minimization in step (1) is efficiently done in Fourier space, using FFT. Since we choose unity gain for the dynamic model, the mappings \mathcal{G} and \mathcal{M} are independent. This implies, that none of the parameters in \mathbf{a} can compensate for geometric distortions and \mathbf{b} can not mimic any dynamic behavior. It is thus guaranteed, that the algorithm converges to a unique solution. Since already the first estimation step leads to a dynamic model with more than 90% accuracy, only local convergence of the algorithm is required. As the stimulus signal is a plane filling curve, geometric distortions are weighted uniformly in 2D-image space, whereas dynamic weighting is $\sim 1/f$ due to the self-similarity property of fractals. It is straightforward to use different spectral weighting too, since the model estimation technique is performed in frequency space. In Fig. 4 we show a detail of the re-simulated signal compared to the measured ones after 4 iterations.

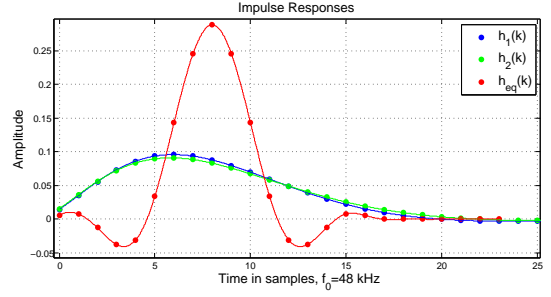


Fig. 5. Estimated discrete time impulse responses of system (blue: x_1 – mirror axis, green: x_2 – mirror axis), and responses after dynamic equalization (red).

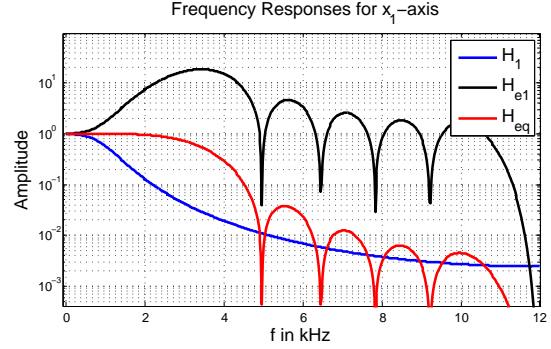


Fig. 6. Frequency responses for x_1 -mirror axis (blue), equalizer transfer function (black) and response after dynamic equalization (red).

C. Results

1) *Geometric Distortion*: Staying in the range of small amplitudes, which is best practice to get locally reliable results for a linearized dynamic model, higher order optical distortions are small. Nonetheless, optical correction improves the estimation of the *dynamic* model substantially, as can be seen in Fig. 4.

2) *Dynamic Model*: In Fig. 5 we show the impulse responses for the models of two scanner axes.

Obviously, the systems have nonlinear phase (no symmetry in the impulse response), thus originally symmetric objects will appear asymmetrical. Responses are different between x_1 – mirror and x_2 – mirror axes, which is another source of contour distortion. Typical response-time is 10 samples.

D. Model Validation

1) *Equalizer Construction*: A good way to validate a dynamic model, is to construct an equalizer, which pre-distorts the signal in such a way, that the serial connection of equalizer and galvo behaves like a symmetric finite impulse response system. Obviously, the equalizers must have zeros at the pole positions of the galvo systems in the Z -plane. In addition, we apply the constraint, that the total responses of both axes are identical, which leads to a coupled channel optimization problem. Note, that we do not require the equalizer to act like the complete inverse of the galvo systems. This would lead to extremely high amplitudes in the control signal, which are not realizable in practice. System order of the equalizers as well

as their frequency responses are design parameters, so that the gain in dynamics can be traded off with control amplitudes. Mathematical details of optimal equalizer design will be described elsewhere. Fig. 5 shows the resulting combined impulse response of equalizers and galvos (red), which are the same for both axes. Obviously we have a constant group delay of 8 samples and a typical response time of only 5 samples in this example. The Frequency responses as displayed in Fig. 6 clearly shows that the effective bandwidth of the system has been extended to more than 2 kHz as compared to about 800 kHz. On the other hand, it makes no sense to "over-compensate" the galvo system for higher frequencies, which would lead to non tractable behavior and noise amplification. As seen, our method leads to regularized equalizers with low gain in the high frequency region.

2) *Assessment of dynamic gain:* Here we give a new method for quantifying the dynamic gain achieved, e.g. by inclusion of an equalizer into the galvo path. The idea is based on multi-scale analysis and gives an objective measure by a scale resolution curve. As an example consider Fig. 2c, which has been photographed using the same scanner hardware, but with an equalizer in series to the control signal. With the same stimulus signal, (Fig. 2a), the triple snow flake curve could have been resolved to a much finer scale (about one construction level). For a quantitative assessment, a multi-scale-analysis of the measured output signals using wavelet transforms are a natural choice (see e.g. [8]). The Wavelet coefficients are calculated by

$$X_d(k) = \sum_n \psi_d(k-n)x(k), \quad (5)$$

where $\psi_d(k)$ are suitable wavelet functions with scale parameter d . Due to their self-similarity, the fractal stimulus signals $u(k)$ have enough signal energy on all scales: $E(d) = \sum_k |U_d(k)|^2$, where $U_d(k)$ denote the wavelet coefficients of the stimulus. A meaningful quantity to analyze the resolution of a scanner system is then the ratio

$$\rho(d) := \frac{\sum_k |X_d(k)|^2}{\sum_k |U_d(k)|^2}. \quad (6)$$

Values $\rho(d) \simeq 1$ imply a high accuracy for scale d . Deviations ($\rho(d) < 1$) indicate a loss of precision. In Fig. 7 the blue curve represents the result for a conventional scan head, the green and red curves result from pre-distorted control signals with differently parametrized equalizers. Clearly, at a level of $\rho = 0.95$, an improvement of about one order of magnitude is possible. The scale (in meters) shown here is referred to a fixed geometric setup, and can of course be converted to angles (mrad), if different scanner systems are to be compared.

III. APPLICATION: REMOTE ABLATION CUTTING

Remote laser technologies using long focal lengths become more and more attractive in industry. Remote welding and cutting are two examples of these processes [9]. One advantage of remote technology is clearly the possible enhancement in productivity as compared to conventional laser processes. For

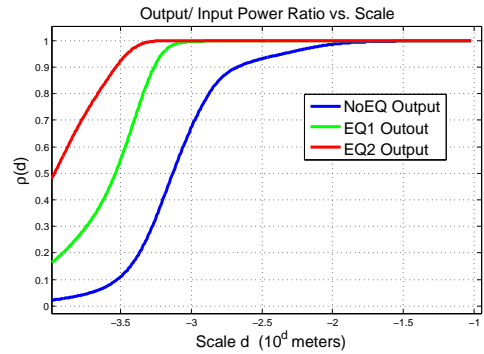


Fig. 7. Scale behaviour of uncompensated scanner system (blue). Scale refinement due to equalization (red and green).

remote ablation cutting (RAC) process, a high power laser (1-10 kW) is focused on the workpiece with a very small diameter of about 50-100 μm at a focal distance of around half a meter. The resulting beam intensity is sufficient to directly evaporate steel material. However, to maintain a high percentage of vapor, the laser spot velocity must be in the range of 5 m/s and above. Thus, highly dynamic scanner systems are enabling technology for RAC.

A. Model based Optimal Trajectories

With a valid dynamic model at hand, it is straightforward to construct optimal control signals to get the desired contour. We start out with a model in discrete state space form,

$$\mathbf{x}(k+1) = A\mathbf{x}(k) + Bu(k) \quad (7a)$$

$$\mathbf{y}(k) = C\mathbf{x}(k) + Du(k), \quad (7b)$$

where $\mathbf{x}(k)$ denotes the n dimensional state vector of the system at sample instance k . In the following, we discuss solutions for the period- N case ($N > n$), since this is what we need for the RAC application. Let

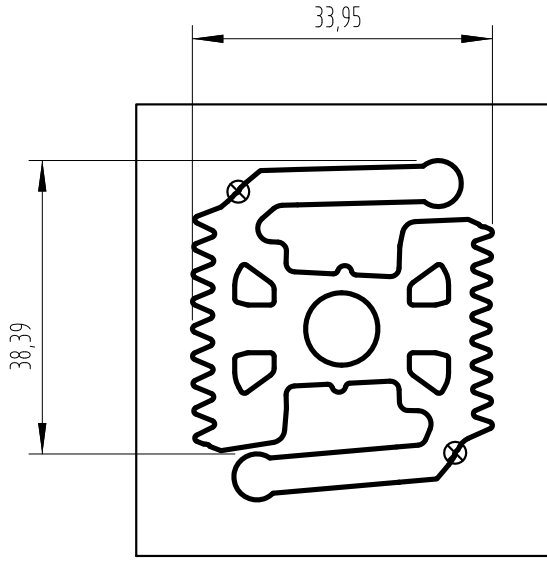
$$\mathbf{u} = \begin{pmatrix} u(0) \\ \vdots \\ u(N-1) \end{pmatrix} \text{ and } Q_N = [A^{N-1}B \ \dots \ A^2B \ AB \ B] \quad (8)$$

be the input control sequence and the (extended) control matrix respectively. State evolution then reads

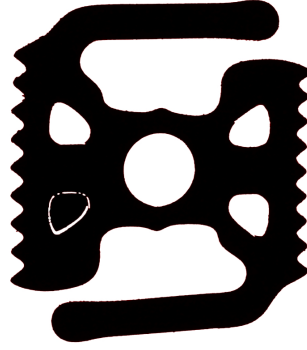
$$\mathbf{x}(N) = A^N \mathbf{x}(0) + Q_N \mathbf{u} = \mathbf{x}(0), \quad (9)$$

where the second equation is due to periodicity. This leads immediately to the initial state $\mathbf{x}(0) = -(A^N - I)^{-1} Q_N \mathbf{u}$, provided the system has no eigenvalue at $z = 1$. The system output vector \mathbf{y} for N samples can be written as

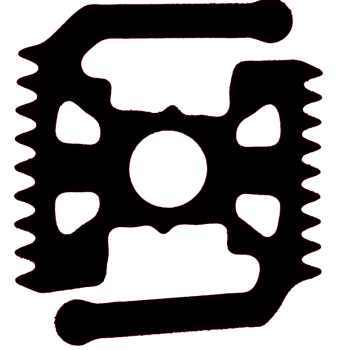
$$\mathbf{y} = Y_0 \mathbf{x}(0) + Y_1 \mathbf{u}, \quad (10)$$



(a) Set-point contour (sizes in mm).



(b) Result with standard method, $T_c = 230$ ms.



(c) Cutting result with model based optimal pre-distortion, $T_c = 99$ ms.

Fig. 8. Results of RAC using an industry-contour on a 0.2mm steel sheet.

where the matrices Y_0 and Y_1 are given by

$$Y_0 = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{N-1} \end{pmatrix}, Y_1 = \begin{pmatrix} D & & & & \mathbf{0} \\ CB & D & & & \\ CAB & CB & D & & \\ \vdots & \vdots & \vdots & \ddots & \\ CA^{N-2}B & \dots & \dots & \dots & D \end{pmatrix}, \quad (11)$$

and \mathbf{y} is defined in analogy to \mathbf{u} in (8). With (9) and (10), the control signal \mathbf{u} can be readily obtained for any given output signal $\mathbf{y} = \mathbf{y}_s$. However, in practice, the set-point signal \mathbf{y}_s , as constructed e.g. by a CAD program, contains (mostly unintentionally) high frequency components. This leads to extremely high amplitudes in the control signal \mathbf{u} , making it obsolete for practical usage. Instead, we allow for a deviation error \mathbf{e} between the set-point signal \mathbf{y}_s and the resulting output \mathbf{y} :

$$\mathbf{y}_s = \mathbf{y} + \mathbf{e} = Y_0\mathbf{x}(0) + Y_1\mathbf{u} + \mathbf{e}. \quad (12)$$

Furthermore, the system model is augmented by an additional output $v(k)$, whose energy is to be minimized. As example, we could take $v(k)$ to be the (discrete) 2nd derivative of $u(k)$. In analogy to (10) we find

$$\mathbf{v} = V_0\mathbf{x}(0) + V_1\mathbf{u}, \quad (13)$$

and we are ready to formulate a functional

$$\mathcal{L}_1 = \mathbf{v}^T\mathbf{v} + \alpha \mathbf{e}^T\mathbf{e} + [\mathbf{x}(0)^T(A^N - I)^T + \mathbf{u}^T Q_N^T] \boldsymbol{\lambda}. \quad (14)$$

Here α is a weighting parameter, controlling the allowed deviation of \mathbf{y} from its set-point, and $\boldsymbol{\lambda}$ is a Lagrange multiplier. In fact, the term in square brackets ensures the boundary conditions in (9). Minimizing the functional (14) with respect

to \mathbf{u}^T , $\mathbf{x}(0)^T$ and $\boldsymbol{\lambda}^T$, leads to the following linear system of equations:

$$\begin{pmatrix} V_1^T V_1 + \alpha Y_1^T Y_1 & V_1^T V_0 + \alpha Y_1^T Y_0 & Q^T \\ V_0^T V_1 + \alpha Y_0^T Y_1 & V_0^T V_0 + \alpha Y_0^T Y_0 & (A^N - I)^T \\ Q & A^N - I & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \mathbf{x}(0) \\ \boldsymbol{\lambda} \end{pmatrix} = \alpha \begin{pmatrix} Y_1^T \mathbf{y}_s \\ Y_0^T \mathbf{y}_s \\ 0 \end{pmatrix}. \quad (15)$$

Solving this (symmetric) set of equations using standard techniques, leads to the desired control signal \mathbf{u} and the initial state $\mathbf{x}(0)$.

B. Model based Optimal Transients

If the object we want to cut is not simply connected, transients are required, where the laser beam is switched off. Our goal is here to find a N -point control sequence \mathbf{u} , connecting the given states $\mathbf{x}(0)$ and $\mathbf{x}(N)$. Again, we augment the model with an output $v(k)$, and with (9) and (13) we can set up an energy functional

$$\mathcal{L}_2 = \mathbf{v}^T\mathbf{v} - [\mathbf{x}(0)^T A^{N^T} + \mathbf{u}^T Q_N^T - \mathbf{x}(N)^T] \boldsymbol{\lambda}. \quad (16)$$

Inserting (13) and minimization with respect to \mathbf{u}^T again yields a linear system of equations,

$$\begin{pmatrix} -V_1^T V_1 & Q^T \\ Q & 0 \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} V_1^T V_0 \mathbf{x}(0) \\ \mathbf{x}(N) - A^N \mathbf{x}(0) \end{pmatrix}, \quad (17)$$

which can be readily solved for the desired control signal vector \mathbf{u} .

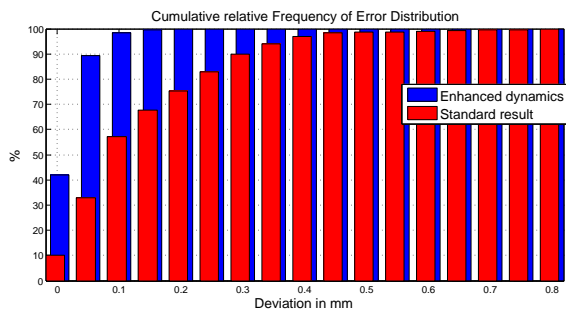


Fig. 9. Comparison of cumulative contour error: Standard method (red), enhanced dynamics (blue).

C. Results

We want to cut the contour shown in Fig. 8a from a 0.2 mm steel sheet. This contour is from an industrial application, drawn with a CAD tool. Essentially all curves are polygons, leading to the high frequency problem discussed above. A 1 kW laser together with a 35 mm aperture scanner system has been used. Clearly, this contour is non-simply connected, and contains lots of small-scale details. Applying the current industry-standard algorithm results in Fig. 8b. Cutting period for the whole contour was $T_c = 230$ ms, lower speed (meaning higher T_c) leads to the re-welding. This effect has already partly occurred in the lower left "eye" of the contour. Small scale details are obviously slurred, and amplitudes of the "sine-waves" are too low. In contrast, cf. Fig. 8c, applying the methods discussed in this paper, we achieve a contour which is true to size also at smaller scales, while applying more than twice the cutting speed: $T_c = 99$ ms. This leads to a true cutting speed above 4 m/s and there is no danger of re-welding. For a more quantitative investigation, we show in Fig. 9 the cumulative relative frequency of the contour error. The blue curve indicates, that with our method the contour error is less than $100 \mu\text{m}$ for 97% of all measurements (blue), as compared to $400 \mu\text{m}$ at the same level for the standard method (red). Alternatively, the root mean square contour error can be evaluated. We find $e_{rms} = 200 \mu\text{m}$ for the standard contour, to be compared to $e_{rms} = 50 \mu\text{m}$ for the dynamically enhanced curve.

IV. CONCLUSION

We presented a novel method for estimating dynamic and geometric models for laser scanners within an unified framework, based on a composite mapping algorithm. As stimulus signals for gray-box modeling fractal, plane filling curves proved very useful. While geometric distortions can be easily rectified by a static mapping, increasing the dynamic (and thus making small scales at high speed accessible) is more involved. We outlined methods for optimally creating control sequences for prescribed periodic contours, and for connecting them by transients. In a remote ablation cutting application, using industrial set-point curves, we showed the successful operation of the methods presented.

ACKNOWLEDGEMENT

The authors would like to thank G. Mandel for technical assistance, and the Bayerische Forschungsstiftung, reference AZ-1045-12, for financial support of this project.

REFERENCES

- [1] S. Pieczona, "Modellbildung eines Galvanometer-Laserscanners", to appear in wt Werkstattstechnik online, 2015.
- [2] H. W. Yoo, S. Ito, M. Verhaegen, and G. Schitter, European Control Conference (ECC), Zrich, Switzerland, 2013.
- [3] G. Peano, "Sur une courbe, qui remplit toute une aire plane", *Mathematische Annalen* 36 (1): 157160, 1890.
- [4] Eric W. Weisstein, "Peano-Gosper Curve", MathWorld. Retrieved 31 October 2013.
- [5] H.K. Yuen, J. Princen, L. Illingworth and J. Kittler "Comparative study of Hough transform methods for circle finding", *Image and Vision Computing*. Vol. 8, Num 1, pp 71-77, 1990.
- [6] A. Ardeshir Goshtasby: "2-D and 3-D Image Registration for Medical, Remote Sensing, and Industrial Applications", Wiley Press, 2005.
- [7] J.A. Cadzow, "Signal Enhancement- A composite Property mapping Algorithm". *IEEE trans. Acoust., Speech, Signal Processing*, 36(1): 49-62, January, 1988.
- [8] I. Daubechies: "Ten Lectures on Wavelets". CBMS-NSF regional conference series in applied mathematics, 61, 2006.
- [9] M. Zaeh, J. Moesl, J. Musiol and F. Oefele, "Material processing with remote technology revolution or evolution?" In: F. Schmidt, et al. (Hrsg.): *Laser Assisted Net Shape Engineering* 6 pp 19-33, 2010.