Digital Signal Processing on Optoelectronic for SHM

W. Flores-Fuentes, M. Rivas-Lopez, O. Sergiyenko, F. Gonzalez-Navarro J. Rivera-Castillo, and D. Hernandez-Balbuena

Abstract— This paper proposes a particular advantage of Support Vector Machine (SVM) to predict measurements errors for Accuracy Enhancement in Optical Scanning Systems, for position detection in real life application for Structural Health Monitoring (SHM) by a novel method, based on the Power Spectrum Centroid Calculation in determining the energy center of an optoelectronic signal. The Energy Signal Centre is found in the Power Spectrum Centroid and the SVM Regression Method is used as a digital rectified to increase measurement accuracy for Optical Scanning System.

Index Terms— Support Vector Machines, Power Spectrum Centroid, Energy Signal Centre, SHM.

I. INTRODUCTION

HM is a technology focused on data acquisition and analysis regarding man-made structures. The data is obtained from non-destructive sensors and its analysis allows the detection of anomalies to evaluate its health [1].

In this paper, an Optical Scanner System (OSS) is subject to be enhanced by increasing its accuracy by modeling possible errors in position measurements by means of digital processing techniques.

The Power Spectrum Centroid and Support Vector Machine techniques are introduced in this research to process the OSS optoelectronic output signal in order to calculate the energy centre of the signal by Power Spectrum Centroid and to use error approximation functions to perform the digital rectification of systematic and random measurement error by using a well know machine learning regression model, the Support Vector Machine (SVM).

In an OSS system, right after data acquisition stage takes place; a data management process is carried. It consists of five steps as shown in Figure 1.

Manuscript received July 16, 2013; revised July 16, 2013. This work was supported by the Engineering Institute of Autonomous University of Baja California, Mexicali, Baja California, Mexico.

W. Flores-Fuentes, M. Rivas-Lopez, O. Sergiyenko, F. Felix-Navarro, J. Rivera-Castillo, D. Hernandez-Balbuena are with the Engineering Institute of Autonomous University of Baja California, Mexicali, Baja California, Mexico. (phone: 011-52-686-566-4150; fax: 011-52-686-565-4431; e-mail: wendy.flores@uabc.edu.mx, mrivas@uabc.edu.mx, srgnk@uabc.edu.mx, fernando.gonzalez@uabc.edu.mx, javier.rivera.castillo@uabc.edu.mx, dhernan@uabc.edu.mx Step 1. Data processing through Fourier Transformation to translate data from time domain to frequency domain. Step 2. Power Spectrum computation in order to represent

the data in power per frequency unit.

Step 3. Centroid calculation, which corresponds to the energy centre of the signal (in frequency domain).

Step 4. Convert this frequency domain coordinate to time domain coordinate to obtain the position where the light emitting source under scanning is positioned.

Step 5. A digital rectification in order to obtain accuracy enhancement.



Fig. 1. Digital signal processing flow.

II. OVERALL OSS FUNCTIONING

In this chapter, an OSS [2-4] will be described in detail. Theoretical Principles, measurements concepts and the Power Spectrum Centroid Computing Method will be described, as long as the accuracy enhanced proposal.

A. Scanning

The OSS of this study (a 45° Sloping Cylindrical Mirror Scanner Aperture) shown in Fig 2 consists of the following elements:

- 1. An incoherent light source emitter (non-rotating) mounted on the structure under monitoring. That could be placed at different angles to detect its position as shown in Figure 2 at 90°, 120°, and 150°.
- 2. A passive rotating optical aperture sensor designed with a 45° sloping mirror, and embedded into a cylindrical micro rod. The beam of light is deviated by the 45°sloping mirror to a double convex lens and filtered, in order to remove any interference and enhance the focus.

Proceedings of the World Congress on Engineering and Computer Science 2013 Vol I WCECS 2013, 23-25 October, 2013, San Francisco, USA

3. A photodiode to capture the beam of light while the cylindrical micro rod mounted on a dc electrical motor shaft is rotating.

This last element generates the targeted signals to be analyzed by the proposed method. In the case of that the incoherent light emitter source, mounted on the structure, changes its position, the possibility of deterioration or damage is present. The light emitter source is set at a distance from the receiver; the receiver is compound by the mirror, which spins with an angular velocity ω . The beam emitted arrives with an incident angle β with respect to the perpendicular mirror, and is reflected with the same angle β , according to the reflecting principle to pass through the lens that concentrates the beam to be captured by the photodiode, which generates a signal with a shape similar to the Gaussian function. When the mirror starts to spin, the sensor "s" is synchronized with the origin generating a pulse that indicates the starting of measurement that finishes when the sensor "s" releases the next start signal. This signal is captured by National Instrument USB-6008 data acquisition module to start the signal processing method. Figure 3 shows the timing diagram and how the light energy intensity increments in the centre of the signal generated by the scanner. The sensor "s" generates a starting signal when t_{α} = 0, then the signal is processed and its energy centre is detected.



Fig. 2. OSS with 45° slopping cylindrical mirror

	Signals Data Acquisiti	on: Blue-S	itarting Puls	e Green – Oj	otoelectronicSig	nal Red-	Energy Signal Cen	tre
Voltage	0.01	0.02	0.03	ta	T ₂₇₇	0.06	, i	0.08
			- m	m2	Time (secs)	m	1	

Fig. 3. Timing Diagram

B. Energy Signal Centre

Experimental readings showed that, in order to find the position of a light source, the targeted signal resemblances a Gaussian shaped signal. This is mainly observed when the light source searched by the optoelectronic scanning is a punctual light source. This last assertion corresponds to the fact that when the punctual light source expands its radius, a cone-like or an even more complex shape is formed depending on the properties of the medium through which the light is travelling. To reduce errors in position measurements, the best solution is taking the measurement in the energy centre of the signal generated by the scanner [2].

The Energy Centre of the signal concept employs different mathematical methods, as a way to assess which one yields the most precise measurement.

The geometric centroid, the power spectrum centroid, the energy peaks, and saturation method are some of them [3].

The Energy Centre of the signal can be found by the signal processing of the optoelectronic scanner sensor output. This mathematical treatment can be made through specialized computer software e.g. MATLAB, and some others.

III. AN ACCURACY ENHANCEMENT METHOD

The prosed method is based on the assumption that the signal generated by optical scanners for position measurement is a Gaussian shaped signal. However, during experimentation it was seen that the optoelectronic scanning sensor output is a Gaussian shaped signal with some noise and deformation. This is due to some internal and external error sources like the motor eccentricity at low speed scanning, noise and deformation that could interfere with the wavelength of the light sources. Other phenomena that could also affect are related with the reflection, diffraction, absorption and refraction. In this work a minimization of these undesirable effects is proposed. This minimization is implemented by taking measurements in the energy centre of the signal, applying PSC calculation and a posteriorly digital rectification by machine learning model SVM.

A. Data Acquisition

Signals of different shapes are generated during an optical scanning process, depending on the type of light source, the sensor of the scanner and the whole scanning system characteristics. Some precision semiconductor optical like CCD, Complementary sensors Metal Oxide Semiconductor (CMOS) or PSD produce output currents related to the "centre of mass" of light incident on the device surface. It is well known that CCD, CMOS and PSD use the light quantity distribution of the entire beam spot. The light receiving element sets the beam spot centre or centroid and identifies this as the target position. Similar results are acquired from the photodiode used on the 45° sloping cylindrical mirror scanner. A Gaussian shaped signal is obtained; where its energy signal centroid corresponds to the target position. A hypothetical spot model attempts to explain how the signal is created by the photodiode on a scanner with a rotating mirror, has shown in Figure 4. The

Proceedings of the World Congress on Engineering and Computer Science 2013 Vol I WCECS 2013, 23-25 October, 2013, San Francisco, USA

signal created, as a similar Gaussian shaped, goes up (Figure 4, a) and falls down (Figure 4, e), and a fluctuating activity takes place around its maximum area see Figure 4 (b-d). However in experiments with the 45° sloping cylindrical mirror scanner, it has been observed that the OSS output is a Gaussian shaped signal with some noise and deformations as shown in Figure 5.



Fig 4. Principle of electrical signal formation during rotational scanning.



Fig. 5. OSS measurements at different light source, angledistance positions.

B. Power Spectrum Centroid

The centroid is an important feature of power spectrum. It can be seen as the gravity centre of the frequency components in a spectrum. The Power Spectrum Centroid has been experimentally observed in works related to digital signal processing on audio signals. In the method proposed in this work, it is intend to apply the PSC concept in order to find the energy signal centre by establishing a correlation between both concepts. The power spectrum is a positive real function of a frequency domain variable associated with a time domain function, whose dimensions are power per hertz or energy per hertz. This means that the power carried by the wave (signal) per unit frequency and its centroid is calculated in the frequency domain by (1) corresponding to the energy signal centre [5], where SC is the spectral centroid in frequency (Hertz). Xd[k] is the magnitude

ISBN: 978-988-19252-3-7 ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online) (voltage amplitude) corresponding to frequency bin k. k is the frequency bin (fs=N) in hertz and fs is the frequency sample. N is the length of the DFT (Discrete Fourier Transform).

$$SC_{HZ} = \frac{\sum_{k=1}^{N-1} k \cdot X^{d}[k]}{\sum_{k=1}^{N-1} X^{d}[k]}.$$
(1)

The first step to transit from time domain to frequency domain is to apply the Fourier transformation, which provides an alternative way of representing data, instead of representing the signal amplitude as a time domain function. Fourier series represent the signal by how much information (power) is contained at different frequencies and also allow to isolate certain frequency ranges (that could come from noise sources) if necessary. For OSS applications with a vector of data (finite series) the Fast Fourier Transform (FFT) is applied to compute from time domain to frequency domain. The Discrete Fourier transform (DFT) algorithm was used as a Fourier application for discrete data whose non-zero values have a limited (finite) duration.

The second step is to compute the power spectrum, which is the square of the absolute value of the FFT. Its result is considered as the power of the signal at each frequency.

The third step consists to calculate the power spectrum centroid by (1).

And finally, the centroid value is converted from frequency domain to time domain. The domain conversion is performed by a) a frequency bin correlation with time scale.

C. Support Vector Machines and Digital Rectification

Support Vector Machines (SVM) was developed to solve classification and regression problems in machine learning or pattern recognition area. The statistical learning theory goal in modeling is to choose a model from the hypothesis space, which is closest (with respect to some error measure) to the underlying function in the target space as described on [32, 33]. Support vector regression can be applied to: a. Linear regression by the introduction of an alternative loss

function (as Quadratic, Laplace, Huber and ε -insensitive);

b. Non-linear regression by a non-linear mapping of the data into a high dimensional feature space where linear regression is performed. SVM for Regression is an algorithm used in the forecasting of several kinds of applications as described on [6-13]. One of these examples is in the error prediction aimed to do some measurement corrections. In this paper our goal is predict the measurement error by means of a SVM regression to perform the digital rectification by measurement correction i.e. adding the predicted error to a measurement. Given the Energy Signal Centre measurements found in the Power Spectrum Centroid and the target angle value under measurement, the measurement error was calculated by (2). Where: E is the measurement error, representing how far the measurement from the real value is. \propto_R is the angle real value and \propto_M is the angle measured by the system.

 $E = |\alpha_R - \alpha_M|....(2)$

Proceedings of the World Congress on Engineering and Computer Science 2013 Vol I WCECS 2013, 23-25 October, 2013, San Francisco, USA

IV. EXPERIMENTAL WORK

Measurements were performed scanning from 45° to 135° , positioning the light emitter source at near distance to the optical aperture sensor and displacing it away from the optical aperture sensor through the same angle. Ten measurements were taken at each point e.g. angle, distance, repeating this measurement process every 5° . Figure 6 depicts this process. A total of 6020 measurements built the dataset that was used to train and to test the SVM algorithm in the error prediction. Figure 7 and 8, gives an error profile by position and angle. It is seen that the error does not follows a well-defined behavior or any known function. As far as the light emitter source is moved away from the optical aperture sensor, the error increments. Also, it was noticed that the best scanning window angle was close the 90°.

A. SVM training and testing

The following steps were used to train and to test the SVM regression algorithm:

- 1. After data was acquired, a scaling step on each column or feature took place.
- 2. In order to properly validate the SVM regression performance, the well-known k-fold cross-validation method was used.
- 3. SVM parameters were selected by means of a gridsearch strategy, i.e. for each parameter, a group of several parameters values were tested by k-fold cross-validation. The best parameters were the ones with the best performance measure.

The data was separated in two sets, training data set and test data set. Each instance in the training-test sets contains one target value (error measurement) and three attributes or features, angle, distance and frequency. In order to estimate how accurately the predictive model will perform in practice, the k-fold cross validation technique was used to randomly partition the training dataset into k equal size subsamples. The k-1 subsamples were used to train the SVM algorithm and last one was used to test the trained

SVM. This process is repeated k-times and rotating the subsamples. The SVM algorithm needs what is called a Kernel function, in order to map the original data into the feature space, on which the classification or regression task is performed.

Two kernel functions were tested: Polynomial:

$$K(x_i x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0.....(3)$$

Radial basis function (RBF):

$$K(x_i x_j) = exp(-\gamma x_i - x_j)^2, \gamma > 0.....(4)$$

Where, γ (gamma kernel), r (coefficient kernel) and d (degree of polynomial kernel) are kernel parameters.

All experiments were conducted in MATLAB 2011b (7.13.0.564) 32 bits version. The SVM regression algorithm from the PRTools Toolbox for MATLAB was used.

The original data was randomly partitioned into 2/3 for training data e.g. the whole k-fold cross validation and 1/3 to construct an independent test set. The entire process was repeated 200 times.

The following kernel parameters were analyzed –200 runs for each parameter combination:

- 1. Kernel type= [polinomial, radial]
- 2. NU-parameter = $[2E^{-8}, 2E^{-7}, 2E^{-6}, \dots 2E^{6}, 2E^{7}, 2E^{8}]$
- 3. Polynomial degree in Polynomial kernel = [1,2]
- 4. Kernel width in Radial kernel = $[2E^{-8}, 2E^{-7}, 2E^{-6}, ..., 2E^{-6}, 2E^{7}, 2E^{-8}]$
- 5. C-Complexity parameter= $[2E^{-8}, 2E^{-7}, 2E^{-6}, \dots 2E^{6}, 2E^{7}, 2E^{8}]$.

All the data sets values were linear scaling to the range [-1,1] to avoid attributes dominance and/or numerical difficulties during kernel calculations. Performance measure in the k-fold cross-validation was the Mean Square Error (MSE).



Fig. 6. Measurements angle and distance representation.



Fig. 7. Measurement Error Average by Position.



Fig. 8. Measurement Error Average by Angle.

Proceedings of the World Congress on Engineering and Computer Science 2013 Vol I WCECS 2013, 23-25 October, 2013, San Francisco, USA

V. RESULTS AND DISCUSSION

Experimental results show that, after testing several values in SVM regression training parameters, the configuration that yields to get the best MSE are:

NU-parameter =[2E⁻⁶] Kernel type= [Radial] C-Complexity Parameter=[2] With a MSE=0.016 as can be seen on Figure 9.

Educated practices in machine learning literature indicate, if possible, to run a second search or fine parameters search, by taking a narrow neighborhood around numerical ones. Hence NU was set at $2E^{-6}$ and the C parameter was varied from $1E^{-3}$ to $4E^{0}$ as describe on Figure 10.



Fig 9. First parameters value search.



Fig 10. Second parameters values search for $NU=2E^{-6}$.



Fig 11. Plot Regression

Using the final SVM regression model with best parameters values found, i.e. NU-parameter = $[2E^{-6}]$, Kernel type= [Radial], C-Complexity Parameter=[2], independent test set

was feed as input, and predicted values were obtained. Figure 11 shows a simple correlation between predicted values and target values. A R-correlation coefficient was calculated, achieving a value of 0.99556. As complement, Figure 12 shows test set target regression data and test set predicted regression data. It is seen the both charts profiles match with a highly degree of similarity.

Finally, adding the predicted error to the measurements as described by (5),

$$\alpha_{MC} = \alpha_M + E_p \qquad (5)$$

Where:

 α_{MC} is the angle measurement by the system corrected (digital rectified).

 α_M is the angle measured by the system.

 E_p is the predicted Error

Therefore, the measurements were digital rectified, increasing the accuracy of the optical scanning system. Hence, more than 98.3% of measurements with an error lower than 3° were achieved, as can be seen on Table 1.



Fig 12. Real Vs. Predicted Scaled Measurement Error



Table 1. Measurements Digital Rectified Errors Average Mapping.

Proceedings of the World Congress on Engineering and Computer Science 2013 Vol I WCECS 2013, 23-25 October, 2013, San Francisco, USA

VI. CONCLUSIONS

Power Spectrum Centroid method improvement by machine learning techniques was performed in optical scanning systems for accuracy enhancement with successfully results.

Measurements at power spectrum centroid before digital rectification presented errors from 0.00° at near distance e.g. angle = 130° , distance = 24 to 10.3° at long distance e.g. angle = 105° , distance = 53. A total of 51.3% of position measurement presented a lower error of 3° , and the remaining 48.7% showed errors from 3° to 10.3° , increasing the error at long distance due to incoherent light emitter source effects through distance, and others internal and external error sources previously described.

After digital rectification, errors shown values from 0.00° e.g. angle = 95°, distance = 23, @Table 1 to 4.87° e.g. angle = 100°, distance = 47, @Table 1. 98.3% of measurements digitally rectified presented an error value lower than 3°, and the remaining 1.7% from 3° to 4.87°. Described on Table 2.

Table 2. Digital Rectification Results Comparison

	Error Range	% of Error less than 3°	% of Error bigger than 3°	Position Example Angle/Distance
Before Digital Rectification	From 0° to 10.3°	51.3%	48.7%	130°/24 105°/53
After Digital Rectification	From 0°to 4.87°	98.3%	1.7%	95°/23 100°/47

Experimental results showed that digital rectification algorithms have a positive performance on measurement adjustments. Also, other improvements have been visualized for further research on the mechanical-electrical optical system, as in the energy signal centre digital processing and its rectification by the analysis of the motor rotation frequency effects.

Acknowledges

The authors would like to acknowledge Autonomous University of Baja California (UABC) and CONACyT who supported this research.

References

- [1] Y. Bao, H. Li, X. Sun, Y. Yu, J. Ou, Compressive sampling {based data loss recovery for wireless sensor networks used in civil structural health monitoring, Structural Health Monitoring 12 (1) (2013) 78-95.
- M. Rivas Lopez, O. Y. Sergiyenko, V. Tyrsa, W. Hernandez Perdomo, L. Devia Cruz, D. Hernandez Balbuena, L. Burtseva, and J. Nieto Hipolito, "Optoelectronic method for structural health monitoring," Structural Health Monitoring, vol. 9, no. 2, pp. 105–120, Mar. 2010. DOI. 10.1177/1475921709340975. [Online].
- http://shm.sagepub.com/cgi/doi/10.1177/1475921709340975
- [3] W. Flores Fuentes, M. Rivas Lopez, O. Sergiyenko, and J. Rivera Castillo, "Comparison of signal peak detection algorithms in the search of the signal energy center for measuring with optical scanning," in Proc. IEEE Section Mexico, IEEE ROC&C2011:XXII autumn international conference on communications, computer, electronics,

automation, robotics and industrial exposition, Guerrero, Mexico, Jan. 2011, CP10, PON15.

- [4] M. Rivas Lopez, W. Flores Fuentes, J. Rivera Castillo, O. Sergiyenko, and D. Hernandez Balbuena, "A Method and Electronic Device to Detect the Optoelectronic Scanning Signal Energy Centre". Croatia: In Tech, 2012, pp.1 30, (in Europe).
- [5] T. Park, Introduction to digital signal processing: Computer musically speaking, World Scientific Publishing Company Incorporated, 2010.
- [6] Gunn, Steve R. "Support vector machines for classification and regression." *ISIS technical report* 14 (1998).
- [7] J. Gascón-Moreno, E. Ortiz-García, S. Salcedo-Sanz, A. Paniagua-Tineo, B. Saavedra-Moreno, and J. Portilla-Figueras, "Multi-parametric Gaussian Kernel Function Optimization for e-SVMr Using a Genetic Algorithm," in Advances in Computational Intelligence, Ser. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, vol. 6692, chapter 15, pp. 113–120, 2011. DOI: 10.1007/978-3-642-21498-1_15. [Online]. Available: http://www.springerlink.com/index/10.1007/978-3-642-21498-1_15
- [8] Z. Zhang, "Some recent progresses in network error correction coding theory." IEEE, Jan 2008, pp. 1–5. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber =4476186
- [9] J. K. Choi, S. H. Park, D. J. Cho, and K. Y. Seo, "Correction error generation algorithm for differential positioning performance analysis of navigation equipment." IEEE, Oct 2008, pp. 1099–1103. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber =4694316
- [10] C.-W. Hsu, C.-C. Chang, and C.-J. Lin, "A Practical Guide to Support Vector Classification", Tech. Rep.,2003. DOI: Tech. Rep.
- [11] R.P.W. Duin, P. Juszczak, P. Paclik, et al. "PRTools4: A Matlab Toolbox for Pattern Recognition, Version 4.1", 2007.
- [12] F. Perez-Cruz, O. Bousquet, Kernel methods and their potential use in signal processing, Signal Processing Magazine, IEEE 21 (3) (2004), 57-65
- [13] K. Shinoda, N. Inoue, Reusing Speech Techniques for Video Semantic Indexing [Applications Corner], Signal Processing Magazine, IEEE, 30 (2) (2013) 118-122.