DISCRETE WAVELET TRANSFORM IN LINEAR SYSTEM IDENTIFICATION

Zdeněk Váňa

Doctoral Thesis Statement

Department of Control Engineering
Faculty of Electrical Engineering
Czech Technical University in Prague

Prague, February 2014
Zdeněk Váňa: Discrete Wavelet Transform in Linear System Identification
Doctoral Thesis Statement © 2014

PhD Program:
Electrical Engineering and Information Technology

Branch of Study:
Control Engineering and Robotics

Supervisor:
Prof. Ing. Michael Šebek, DrSc.
Department of Control Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague

Supervisor-specialist:
Doc. Ing. Lukáš Ferkl, Ph.D.
Department of Control Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague
Declaration

This doctoral thesis statement is submitted in partial fulfillment of the requirements for the degree of doctor (Ph.D.). The work submitted in this dissertation is the result of my own investigation, except where otherwise stated.

I declare that I worked out the thesis independently and I quoted all used sources of information in accord with Methodical instructions about ethical principles for writing academic thesis. Moreover I declare that it has not already been accepted for any degree and is also not being concurrently submitted for any other degree.

Prague, February 2014

----------------------------------
Zdeněk Váňa
Mathematical modeling of systems and processes is an essential tool used in science and engineering from its very beginning. Newton’s laws are an example for a model. They describe the motion of bodies such as the various bodies in the solar system: the Sun, the planets and their moons. Models are used extensively in all branches of science and engineering, which raised to the term model-based engineering. The fitting of the model to the actual physical system takes a central position, which systems engineers refer to as system identification, and it gained a lot of attention until quite recently in control engineering. The development of model-based control is originated in 1960’s and it is one of the milestones in control. Before that time, most of control design strategies had been based either on heuristic methods (Ziegler-Nichols method for tuning of PID controller parameters) or on system properties obtained via simulation or measurement (impulse response, Bode graph). Since 1960’s, a huge development of control techniques employing the model of the controlled process has been recorded and hand in hand with this development, system identification became a more important part of control theory. Nowadays, system identification is an indispensable part of modern control theory and covers a wide range of problems and approaches to their solution. Moreover, the progress in system identification is still under pressure, since the growing requirements on control naturally reflect into the growing demands on the model. Details about the history of system identification can be found in Gevers [2006].

Model is an image of the real process that represents its properties essential to the application and that behaves within a given context similar to the real process. There is, however, a plenty of choices for the modeling itself (a description of the model, a model structure or a model order), which affect the
resulting model and its properties. Therefore, there can be several models of the same process, each describing the process in a different way, hence the model of the process is by no means unique. As the model is used for a subsequent control, it is obvious that it should reflect only those properties of the system, which are important from the control point of view. In addition, the model should also be of a sufficient quality to be suitable for a control. The model parameters estimate is computed from measured input and output data and since these measurements record both wanted and unwanted process behavior, it is therefore very suitable to analyse the data and extract the appropriate portion of information from them.

One of the most used signal analysis had been introduced by French mathematician and physicist Jean Baptiste Joseph Fourier (*1768; † 1830) and, in his honor, has been called Fourier Transform (FT). Fourier analysis reflects the time domain into the frequency domain and results in frequency spectra of the time-varying signal. The natural disadvantage of this transform is the fact, that one domain disables detection of important phenomena in the second domain and vice versa - one is not possible to detect important time instants from the frequency spectra as well as to specify main frequencies (or harmonic functions) from the time domain. Many years after Fourier, Hungarian electrical engineer and physicist Denis Gabor (*1900; † 1979) comes up with Short-Time Fourier Transform (STFT), where the FT of the signal is applied within the given time-window only. STFT is thus trade-off between the time and the frequency signal description, however, its limiting factor is type and mainly fixed size of time-window. As a next logical step, STFT with time-varying window has been inspected what laid the fundamental idea of Wavelet Transform (WT). WT analyses the signal from both time and frequency points of view. In principle, WT is similar to STFT. The main difference is that the time-window is not fixed, but scaled across the levels of WT.

1.1 Motivation

Wavelet transform, from its pure essence, brought a lot of possibilities into the science. As every new mathematical tool, it introduced a new way of describing of some parts of the nature and enabled to mathematically formulate some scientific problems. As was already mentioned, the main advantage of WT is its possibility to analyse the signal in both time and frequency domains. This
is, actually, not only the advantage, but simultaneously quite important feature and it should be understood correctly. Wavelet analysis arose from STFT by generalization of the time-window onto wavelets, which is well localised both in time and frequency. It consequently means that instead of the signal being described as a function of either time or frequency, it can be viewed in both time and frequency simultaneously. Analysis in time proceeds by shifting the wavelet along the time axis and analysis in frequency proceeds by scaling the wavelet.

This dual time-frequency approach to signal analysis is, of course, counter-balanced by more complicated mathematical background, however, just thanks to this duality WT found a way to plenty of real applications as one of the most convenient mathematical instrument. As WT serves mainly as a tool for signal analysis, typical ways of its utilization belong into the signal processing, namely i) detection of signal discontinuities, ii) trend detection, iii) detection of self-similarities, iv) particular frequency detection, v) signal suppression, vi) signal denoising and vii) data compression. Although these above-mentioned ways are widely known and used, applications of WT are not limited to them at all. the analysis itself is not the only new product of wavelet research, but the wavelet functions as well. There are many kinds of wavelet functions each of them possess several desired and useful properties. Thus the signal analysis does not have to be the only purpose of utilizing of wavelets. Additionally, as time went on, it showed up that one can look at the wavelets from several different points of view. While classical approaches to wavelets are via functional analysis or vector spaces theory, other concepts were proposed, e.g. through theory of frames in vector spaces, set theory or theory of finite elements.

As was mentioned above, wavelets with their characteristic properties have number of possible applications in many diverse fields, usually those closely related to the scientific research and development. Some obvious applications are biological signals analysis (EEG, EKG), analysis of seismic activity and prediction of earthquakes, analysis of sounds, multidimensional signal analysis together with data compression often applied to image processing as specific representative, analysis and attempts to predict the behavior of stock market, financial data analysis, etc. This all denote the capabilities of WT and wavelets themselves and predict them great importance in the future. However, despite the apparent strength of WT, it is still not used as much and as frequently as it can be and as it deserves.
Until recently, wavelet theory has not been extensively applied to many theoretical and practical problems belonging to the system identification so much. However, there are several papers which deal with application of wavelets in this field. Almost all of that applications exploit the superior properties of wavelet analysis or are at least based on some specific property of the particularly used wavelet function - both cases have already been discussed in Chapter 1. Moreover, absolute majority of that applications deal with a specific wavelet function only and do not consider general wavelet function.

2.1 Wavelet transform in linear system identification

2.1.1 2nd order systems identification

Starting with simple Linear Time-Invariant (LTI) systems, one of the first attempts was to estimate oscillatory properties of a system like natural frequency, damping and stiffness. Ruzzene et al. [1997] proposed a wavelet based estimation method for the system of several interconnected oscillatory 2nd order systems. For the estimation, the Morlet’s wavelets were used because of their advantage of simple mathematical formulation.

Next, Boltezar and Slavic [2004] being inspired by the previous paper solved a similar problem, but with the help of parametrized Gaussian windows instead of Morlet’s wavelet. The idea of wavelet edges showing the importance of particular wavelet coefficients was utilized, however, usual method based on ordi-
nary edge-effect (Staszewski [1998]) was shown to be unsuitable. The authors thus came up with three new approaches to the edge-effect, all improving the proportionality between the wavelet coefficients and the analysed signal.

2.1.2 Wavelet analysis of system relevant signals

Another method of employing wavelets for system identification is to apply the WT to the system relevant signals. Such a general characterization is used intentionally, since there are several different utilizations of wavelet analysis for system inputs, outputs or states across the literature.

Luk and Damper [2006] exploited one of the fundamental properties of wavelets - their mutual orthogonality - to design a suitable system input. It yields the inverse WT of the system’s impulse response. The mutual orthogonality of wavelet functions plays an important role also in the Serban [2007], where the authors took advantage of this property to decompose the input and output of the system into more signals, each lying in specific frequency range. Several models thus could be identified each describing a distinct part of the overall dynamics. Morlet’s wavelets were used.

The ability of removing noise has also found attention: Wang et al. [2010] denoised system input and output to suppress the high frequency content of data to improve the accuracy of parameter estimation.

At last, a general framework of applying the WT on system inputs and outputs should be mentioned, see e.g. Erlicher and Argoul [2007]. This framework does not, actually, contain any specific procedure at all, but represents the natural way of utilization of WT. It is used across different applications of WT no matter the problem specification or complexity. A quite surprising fact is that not only are advantages of wavelets the reason for using the WT for improving either the identification method or the resulted model, but also endeavour or enthusiasm are other frequently occurred reasons for using of WT. In general, the later reasons are usual for investigating a new way or trying to look at the problem from a different point of view. Indeed, those reasons are not explicitly mentioned in any paper, however it is important human nature which forces us to do so even without any well-founded reason.
2.1.3 Wavelet as modulating functions

When studying the WT applications for linear systems identification one should not forget to pay attention to the relationship of wavelets and modulating functions. The basic principle of applying modulating function for system identification is well established topic and firstly was suggested by Shinbrot [1954]. A great summarizing study about the usage of the modulating functions on the system identification has been performed in triplet Preisig and Rippin [1993a,b,c].

A few years later, one of the first attempts to employ wavelet function as a modulating function was doubtless Kosanovich et al. [1995]. The Poisson wavelet was used here, which is derived from the Poisson probability density function being a kernel of Poisson transform. Note that Poisson wavelet is neither compactly supported nor orthogonal, thus can not be considered as a real wavelet or modulating function, although the previous related work of Kosanovich had shown that Poisson wavelet complies any of wavelet properties. However, when approximating the exponential, one obtain a function which satisfies desired properties. Consequential work Ramarathnam and Tangirala [2009] analyses the use of Poisson wavelet in more details. Preisig came back to wavelets as modulating functions in Preisig [2010], where he discussed the suitability of input signal for further purpose of system identification. There is however no general interconnection of wavelets and modulating functions, but the utilizing multi-wavelets as modulating functions only.

2.1.4 Linear time-varying system identification

When going through the number of publications concerning this theme, for example Doroslovacki et al. [1998], we can get the impression of existence the only one way of exploiting wavelets for LTV system identification. Namely, varying parameters are generally considered as $n$-dimensional functions so they can be expressed as a linear combination of suitable $n$-dimensional basis functions - wavelet functions in this case. Wavelets are used because of their superior selectivity in frequency domain, because of their inherent orthogonality and because of very large foundation of wavelet functions. Such an approach transforms time-varying system parameters into time-invariant coefficients of parameters approximation via wavelets. Therefore the whole problem becomes time-invariant as well and is solvable as a classical linear system identification.
problem. This idea could be improved by using a shift-scale plane analysis Staszewski [1998], wherein the problem of the proper wavelet selection is addressed.

At first sight, the method could seem as quite simple compared to the obtained result. However, there is one big obstacle which the user should take care when employing this technique. The wavelet functions are square-integrable functions, therefore in discrete-time space even compactly supported functions. It consequently means that any their linear combination has also compact support, so the linear approximation of varying parameter as well. This is probably the biggest defect of the method and its main limiting factor for being used for prediction of evolution of system parameters. On the other hand, parameters approximation via wavelets provides us with an insight into the time-frequency parameters behavior and further analysis can disclose partial relations among parameters development and adjoining events. The discussed method is demonstrated mainly on simulation case studies across the publications, but there are few of them applying it in practise, for instance in biomedical engineering for EEG analysis Wei et al. [2008] or on analysis of hysteretic behavior Chang and Shi [2010].

2.2 Wavelet transform in nonlinear system identification

The procedures of using wavelets on linear processes treated in the previous section are applicable for the identification of non-linear processes as well. There are few different methods of utilizing wavelets which, though regarded as general methods, belong rather into the non-linear section Sjoberg et al. [1995].

One of them is applying wavelet functions for a support vector machine algorithm as admissible support vectors. Wen et al. [2005] proposes a wavelet support vector machine with reproducing wavelet kernel especially for the identification of nonlinear dynamic or approximating a non-linear function. The main advantages of using wavelets here are their compactness, orthogonality and a good reproducibility of wavelet kernel. Moreover, wavelet kernel usually performs much faster learning in comparison to standard neural networks or fuzzy logic Li and Liu [2006].

Another approach employs wavelet function as a sigmoid function within
a neural network, then called wavelet network Zhang and Benveniste [1992]. A lot of publications have been published on this theme, for more details, the reader is referred to Ghanem and Romeo [2001].

2.3 Extensions of wavelet transform

2.3.1 Advanced wavelets

As time went on, the wavelet theory recorded further development and extension. For instance, new “families” of wavelet functions have been discovered and the wavelet theory has been established for non-orthogonal wavelet functions. Therefore, apart from the simple wavelet analysis, the bi-orthogonal wavelets Ho and Blunt [2003], wavelet frames Sureshbabu and Farrell [1999] or multi-wavelets Strang and Strela [1994] can be possibly used for system identification.

As the reader probably mentioned, many wavelet applications have exploited wavelets with an explicit mathematical description. Using of just these wavelets is not surprising. Since it is a common practise to use impulse or step functions (or their combination, e.g. pseudo-random binary signal) as a system input, the wavelet analysis of such signals can be mathematically derived only in case of wavelets with any mathematical expression. One can have a different point of view as well - when the user has a possibility to design the system input (e.g. performing a simulation), exploiting these wavelets can be advantageous due to the results in the form of direct mathematical formulae. The benefit is apparent: case study independent direct equations for parameters estimation. On the other hand, many models are still computed as case specific or under process. Therefore it is not desired to derive any general results at all, but rather to sketch a basic invention and derive its elementary properties like asymptotic properties, reproducibility, etc.
3  |  Aims of the Thesis

The thesis is entitled Discrete Wavelet Transform in Linear System Identification and since it is a very general topic, it had to be studied before the work on thesis started. The theory of linear system identification became very important when the model-based control theory has arisen and is now widely used in practise. Next, a search for a good models that are suitable for control applications yields the fact that both academicians and engineers develop or adapt the identification methods for their specific application. An alternative to developing new theories is to combine seemingly unrelated theories thereby adding new components to the identification procedures and adding new views on the involved theories.

Therefore, the main theme of the thesis lies in the interconnection of wavelet theory and theory of linear system identification. The goals are split into the following subjects:

1. To perform a comprehensive survey of the methods of exploiting the wavelet transform for system identification.

2. To find and describe a suitable way of incorporation of wavelet transform into the problem of general single-input single-output linear system identification. Analyse the method and demonstrate it on a suitable example.

3. To extend the method to multivariable systems. Analyse the method and demonstrate it on a suitable example.

4. To investigate and find the utilization of wavelet transform within the continuous-time linear system identification. The discussion on implementation issues must be included.
Here, the basic idea of incorporation of the wavelet transform into the system identification is introduced. For more details, please, see the thesis itself.

### 4.1 Incorporation of wavelets into system identification

The discrete wavelet transform can be understood as a frequency filtering. Moreover, there is the proposition, which gives the conditions for the filtering to be used directly on input-output data while also directly affecting the prediction error. This proposition however required the predictor to be linear in parameters, therefore a well-known ARX (Auto-Regressive with eXternal input) structure has been chosen. The predictor for ARX model structure is of the form

\[ \hat{y}(t|t-1, \theta) = -\sum_{k=1}^{n_a} a_k q^{-k} y(t) + \sum_{k=0}^{n_b} b_k q^{-k} u(t) \]  

(4.1)

and is linear in unknown parameters \( a_k, b_k \). This equation can be rewritten as

\[ \hat{y}(t|t-1, \theta) = z^T(t) \theta, \]

where \( z(t) = [-y(t-1), \ldots, -y(t-n_a), u(t), \ldots, u(t-n_b)]^T \) is the vector of data measured up to time \( t \) and \( \theta = [a_1, \ldots, a_{n_a}, b_0, \ldots, b_{n_b}]^T \) is the vector of unknown parameters. Expressing this equation for all measured data and concatenating them, we obtain the set of equations \( \hat{Y}(t, \theta) = Z(t) \theta \). The
solution $\hat{\theta}_N$ of the equation $\hat{Y}(t, \theta) = Z(t)\theta$ can be easily obtained as a solution to the optimization problem

$$
\hat{\theta}_N = \arg \min_{\theta} \frac{1}{N} \sum_{t=1}^{N} \frac{1}{2} \left[ Y(t) - \hat{Y}(t, \theta) \right]^2 = \left[ Z^T(t)Z(t) \right]^{-1} Z^T(t)Y(t). \quad (4.2)
$$

$Y(t)$ is vector of measured outputs compound in the same way as $\hat{Y}(t, \theta)$, see e.g. Ljung [1999].

The wavelet coefficients are to be evaluated as an inner product of the time signal and even shifts of the wavelet filters. On the $\ell^2(\mathbb{Z}_N)$ space, the inner product can be written as a vector multiplication. If $z$ and $R_k \varphi$ are vectors, then for the coefficients of the wavelet transform of the signal $z$ the following holds

$$
z \ast \tilde{\varphi} = \langle z, R_k \varphi \rangle = \langle R_k \varphi, z \rangle = (R_k \varphi)^Tz = (R_k \varphi)^Tz. \quad (4.3)
$$

For simpler notation, let us consider real valued wavelet filters, what gives us $(R_k z)^T = (R_k z)^T$. The obtained form can be directly used for incorporation of wavelet transform into the system identification (SID) problem. Note that the complex conjugate is applied on a wavelet filter. Specifically, the columns of the data matrices $Y$, $Z$ represent mutually shifted input and output signals. Each of them can be transformed by wavelets. Since the wavelet transform is linear, the problem can be written as a multiplication with an appropriate matrix $T$. The problem is thus transformed onto the problem of minimizing $[TY - TZ\theta]^2$, where the matrix $T = T(\varphi, \psi, \mathcal{P})$ contains all possible shifts of wavelet filters $\varphi, \psi$ at all applied levels $\mathcal{P} \subset \{1, \ldots, p\}$ and we will call it “wavelet matrix”. By adding some user defined weighting matrix $W$, the final optimization problem is

$$
\hat{\theta}_N = \arg \min_{\theta} \frac{1}{N} \sum_{t=1}^{N} \frac{1}{2} \left[ WTY - WTZ\theta \right]^2 \quad (4.4)
$$

and is still solvable via ordinary least squares. When orthogonal wavelets are used, the energy of the transformed signal is preserved in both time and frequency domains by virtue of Parseval’s theorem. Hence error minimization in either domain would give the solution of the same quality Mukhopadhyay et al. [2010].

Notice that transformed prediction error $TY - TZ\theta$ can be understood in 2 possible ways, which both are mathematically identical. The first point of view
is transforming data stored within $Y$, $Z$, thus analysing their time-frequency properties. The second point of view is transforming just the prediction error as $T(Y - Z\theta)$, hence analysing the time-frequency properties of the prediction error with no notion of the data themselves.

Also, at first look, there seems to be a better way of input-output data filtering (transforming), namely to transform input-output data first and then use it for creating of the matrices $Y, Z$. However, when transforming the data first, we obtain wavelet coefficients corresponding to wavelet filters shifted by even number of samples and they do not preserve the time-structure of the data anymore. Therefore, the consequently created matrices $Y, Z$ would have a spoiled structure.

4.2 Wavelet matrix $T$

Let the measured data be of length $D$ and $n_a$ be an order of an estimated model. From the structure of $Z$ it turns out that the length of input-output data for analysis is equal to the dimension of the column space of $Z$, that is $N = D - n_a + 1$, what is desired to be the power of 2. There are two possible points of view on the basic principle of wavelet analysis:

1. Both the approximations and the details of the analysed data are kept of length $N$ (by making use of the upsampling operator) and scaled wavelet filters
   \[ \varphi_j = 2^{j/2} \varphi_1(2^j t), \quad \psi_j = 2^{j/2} \psi_1(2^j t), \quad t \in \mathbb{N} \]
   are used. Then the length of wavelet filters at the $j^{th}$ level is $L_j = 2^{j-1}L_1$ and the shift at this level is $s_j = 2^{j-1}s_1 = 2^j$.

2. On the other hand, the lengths of both approximations and details at all levels are decreased and the analysis is always performed by the same filters $\varphi_1, \psi_1$. Consequently, the length at each level is $L$ and shift $S$ (generally $S = s_1 = 2$).

Since the analysed data need not be of length $2^p$ for any $p \in \mathbb{N}$, the latter approach is more convenient for further use. Based on that, data of length $N$ is decomposed into the approximations and the details, both of length $N/2$. As we do not assume periodically extended vectors, the dimension of the subspaces
will be in our applications always less than \( N/2 \). The dimension depends not only on the data length but on the length of basic wavelet filter \( L \) as well, since \( N \geq L + 2k \) has to hold, where \( k \in \mathbb{N}_0 \) is maximum possible number of the shifts of the wavelet filter of the length \( L \) at the particular analysis level.

Under these assumptions, the data of length \( N_1 = N \) is decomposed into the approximations and the details, both of length \( N_2 = 1 + \left\lfloor \frac{N_1 - L}{S} \right\rfloor \), where \( \lfloor z \rfloor \) denotes the integer part of \( z \). This formula can be written recursively as

\[
N_{j+1} = 1 + \left\lfloor \frac{N_j - L}{S} \right\rfloor \quad (4.5)
\]
as long as the data are long enough for analysis at the next level. Then the number of iterations is maximum level \( p \) of wavelet analysis. With the knowledge of \( p \) and individual lengths \( N_j, j = 1, \ldots, p \) the wavelet matrix \( T \) can be computed as (see the thesis)

\[
T_j = \begin{bmatrix} I & 0 \\ 0 & T_j \end{bmatrix}, \quad \text{where} \quad T_j = \begin{bmatrix} T_{j,D}(\psi) \\ T_{j,A}(\varphi) \end{bmatrix} = \begin{bmatrix} \psi \\ R_2 \psi \\ \vdots \\ R_{2(N_j+1-1)} \psi \\ \varphi \\ R_2 \varphi \\ \vdots \\ R_{2(N_j+1-1)} \varphi \end{bmatrix} \quad (4.6)
\]

It is obvious that the wavelet filters have to be replenished onto the length of \( N_j \). The unit matrix \( I \) has now the size \( \sum_{i=2}^{j} N_i \times \sum_{i=2}^{j} N_i \), matrix \( T_j \) has size \( 2N_{j+1} \times N_j \) and finally, the whole wavelet transform matrix is

\[
T = \prod_{j=1}^{p} T_j = T_p \cdots T_1 \quad (4.7)
\]
4.3 Weighting matrix $W$

The weighting matrix $W$ is a user-defined, diagonal matrix with its elements as weighting coefficients for the particular wavelet filters given by their shift and scale. Several approaches to the weighting of wavelet functions were discussed within the thesis.

4.4 Asymptotic properties of the estimate

The convergence of the estimate of the ARX model parameters via wavelets and its quality are investigated in the thesis.

Since the noise $e(t)$ is correlated with regressor $Z$ (measured data) for the ARX model structure, the estimate of parameters is biased and the following convergence limit holds:

$$
\lim_{N \to \infty} \hat{\theta}_N = \theta^* + \lim_{N \to \infty} E \left\{ (Z^T T W^2 T Z)^{-1} (W T Z)^T e \right\}.
$$ (4.8)

Concerning the quality of the parameters estimate, the variance of the frequency function estimate (in case of an open-loop) at certain frequency was in the thesis proved to be

$$
\text{Var}G(\omega, \hat{\theta}_N) \approx \frac{n}{N} \frac{\Phi_v(\omega)}{\Phi_u(\omega)} \left( \sum_{j=1}^{p+1} |V'(j) \hat{w}_j(\omega)|^2 \right)^{-1},
$$ (4.9)

where $v(t) = H(q, \hat{\theta}_N)e(t)$ is filtered noise, where (power) spectra of the response $y(t) = G(z)u(t)$ of the system $G(z)$ on the input $u(t)$ is $\Phi_y(\omega) = \Phi_u(\omega)|G(e^{i\omega})|^2$ and where

$$
V'(j) = \frac{V(j)}{\max_k \{V(k)\}}
$$ (4.10)

is the normalized weight for the $j^{th}$ level analysis, $V(k)$ is $k^{th}$ element of the vector of weights from which the matrix $W$ is constructed.
4.5 Wavelets for multivariable systems

Although there are number of possible descriptions of LTI systems, two of them are much more frequently used then the others: a state space system description and a transfer function description. The choice of description can be largely assumed as a part of choice of appropriate model structure. There are also some typical SID methods for usual multivariable model structures: the already mentioned prediction error method for the transfer function description and Subspace State Space System IDentification (4SID) method (sometimes just called “subspace” method) for the system in a state space description. Moreover, these methods can be extended to be able to handle the constraints on unknown parameters Prívara et al. [2010, 2012].

In the thesis, possible incorporation of wavelet transform into the identification of multivariable systems is investigated. A condition on system identification which enables the incorporation of the discrete WT (DWT) into it method was already revealed. That condition is a suitable expression of the identification problem so that a time-structure of measured data is preserved so that the DWT! (DWT) can be applied. Realize that when this structure is broken, the DWT can still be applied, but the meaning of filtering as well as time and frequency localisation are completely lost and the DWT becomes just a transformation tool with no interpretation.
The thesis presented a new approaches to utilization of the wavelet transform in the field of system identification. Since the wavelet transform as a mathematical tool serves mainly for signal analysis both in time and frequency domains, the algorithm introduced in the thesis represents a natural way of interconnection of discrete wavelet theory and theory of system identification. This approach is based mainly on favourable properties of wavelet basis functions and brings several advantages:

1. The set of wavelet basis functions at all possible levels (i.e. all scales of one particular couple of wavelet basis functions - father and mother wavelets) forms a set of filters. This set then covers the whole frequency range determined by the properties of analysed signal, more specifically, by its length and a sampling time.

2. Due to theoretical restrictions, orthogonal wavelets have compact support thus simple structure. Consequently, all convolutions are exactly computable, thus no information carried in signal is lost.

3. Some wavelet filters are orthogonal in time and complementary in frequency domains, therefore each filter extracts the specific portion of information from the signals without any duplicity. This fact also contributes to numerical conditioning of the identification algorithm.

4. Moreover, there are several kinds of wavelet basis functions (wavelet families) with different time or frequency properties. Most of them satisfy the necessary conditions given by the wavelet theory, however, there are also some kinds of wavelet functions, which do not. These exceptions
are covered by theory of wavelet frames, which is generalization of the wavelet theory and is built on the theory of Riesz’s basis.

5. Once implemented, this method is quite generic and intuitive, while the design of appropriate linear filters could be quite time consuming. In addition, this provides the user with a big advantage in real problems, where the frequency characteristics of the system to be identified is not known a priori. The satisfactory results could be acquired by tuning of weights only, which corresponds to some knowledge of the system. On the other side, the implementation itself requires deeper understanding to wavelet theory.

At the beginning, the thesis discusses the possibilities of incorporation of wavelet transform into the system identification in the form of a comprehensive study. Then, a general concept of the incorporation of wavelets was introduced. At first, the method was derived for simple SISO systems as well as its asymptotic properties were discussed. Regarding the asymptotic properties of the proposed methods, the results were derived based on properties of the PEM for identification of ARX model. Further on in the thesis, the proposed concept was extended on firstly for MISO and then for MIMO systems. The algorithms for all parts were implemented and their performance was demonstrated on case studies at the end of appropriate chapters. Finally, a detailed analysis of the utilizing the wavelets as modulating functions were elaborated.

Although the wavelets are used, in principle it still is proper selection and (pre)filtering of data with all its pros and cons. There are mainly 2 ways where the methods from the thesis can be employed with advantages:

a) A sufficiently accurate model is required, well describing the system’s behaviour at particular frequency ranges.

b) The only submodel is identified (the lowest order model as possible) which takes into account behaviour on required frequency ranges. It includes ability to identify slow or fast subsystems of singularly perturbed system as well as to do model reduction for identification for control.

Indeed, both ways can be linked together.
5.1 Fulfillment of the objectives

Here a short note on fulfilment of the aims from Chapter 2 is provided.

√ To perform a comprehensive survey of the methods of exploiting the wavelet transform for system identification. → This objective was completed and described in Chapter 3.

√ To find and describe a suitable way of incorporation of wavelet transform into the problem of general single-input single-output linear system identification. Analyse the method and demonstrate it on a suitable example. → This objective was satisfied by development and description of the algorithm, its implementation and demonstration on an example. All is stated within Chapter 6. This is the main part of the author’s publications Váňa and Preisig [2012]; Váňa et al. [2011].

√ To extend the method to multivariable systems. Analyse the method and demonstrate it on a suitable example. → This objective was handled in several points of view on multivariable systems description. The ways of incorporation of wavelet transform into multivariable system identification was described and demonstrated within the Chapter 7. This objective was also partially described in Váňa and Preisig [2012]; Váňa et al. [2011], partially since at that time, the work on the objective was still in progress.

√ To investigate and find the utilization of wavelet transform within the continuous-time linear system identification. The discussion on implementation issues must be included. → This objective was satisfied by the Chapter 8, where the modulating function method primarily designed for continuous-time system identification were shown to be a great alternative for utilization of wavelets within system identification.

All the algorithms implemented within this work are available on the enclosed CD. Moreover, the reader can find there those examples used within case studies, which are not adherent to any coordination with some commercial company.
5.2 Notes on benefits and usefulness of the thesis

The thesis provides several points of view on wavelets, what enables the reader to understand both the wavelets and mutual consequences between wavelets and system identification more deeply. However, for a really deep understanding to both theories, the reader is referred to an appropriate literature, since the thesis states the necessary basics only. A lot of consequences of theories of wavelet transform and of system identification are very intuitive, hence are simply applicable in different fields.

Except the algorithms described in the thesis, the interconnection of theories of both wavelet transform and system identification can be considered as an indisputable contribution of the thesis. Nowadays, there are lot of distinct (not only mathematical) tools serving to some purpose, which are well-known and widely used only in some particular branch, however very seldom elsewhere. It is therefore very important not only to develop some new methods (approaches, algorithms, etc.), but also to look for them within different fields. This is also a strong reason for doing it even in situations where no new results can be obtained, since using of tools which are “new” in particular field can always show new analogies.


Author’s reviewed conference papers related to the thesis – WoS

Author’s reviewed conference papers not related to the thesis – WoS


References


The thesis presents several approaches to system identification in which wavelet transform is employed for both single and multivariable system identification enabling selection of the particular frequency range of interest. We will show the use of wavelet filters with a property of superior selectivity in frequency domain and having compact support in time domain, which, in turn, influences an accurate implementation. These properties provide the user with possibility of measured data analysis in frequency domain without any loss of information. Consequently, selection of a proper filter allows the user to identify the system on a desired frequency range or to identify a number of systems for distinct frequency ranges. This is specifically convenient for the systems with dominant modes, such as singularly perturbed systems. The possibility of selection of a specific frequency range can be utilized for application-based identification and consequent control, when only a limited frequency range is required. Moreover, the thesis treats the possibility of applying the wavelets within continuous-time system identification.

Next, the thesis provides several points of view on wavelets, what enables the reader to understand both the wavelets and mutual consequences between wavelets and system identification more deeply. A lot of consequences of theories of wavelet transform and of system identification are very intuitive, hence are simply applicable in different fields.

Except of the algorithms described in the thesis, the interconnection of theories of both wavelet transform and system identification is the main contribution of the thesis. Nowadays, there are lot of distinct highly professional
tools, which are well-known and widely used by people from some particular branch only. It is therefore very important not only to develop new methods, but also to look for suitable methods across different scientific fields. There is a strong reason for doing it even in situations where no new results can be obtained, since using of tools which are “new” in particular field can always show new analogies and links.
Předložená disertační práce pojednává o přístupech k modelování a identifikaci jedno- i víceozměrných systémů, při nichž se dá s výhodami využít vlnkové transformace jakožto nástroje, jenž díky svým vlastnostem mimo jiné uživatele umožňuje zaměřit se na konkrétní frekvenční vlastnosti naměřených dat. V práci je ukázáno použití vlnkových filtrů, jenž mají velmi dobré selekční vlastnosti ve frekvenční oblasti a konečný definiční obor v časové oblasti, který přímo ovlivňuje výpočetní přesnost algoritmu. Obě zmíněné vlastnosti umožňují uživateli analyzovat naměřená data ve frekvenční oblasti beze ztráty informačního obsahu dat. Důsledkem toho je, že výběrem vhodných vlnkových filtrů může uživatel identifikovat parametry modelu pouze na základě vybraných frekvenčních vlastností naměřených dat nebo také vytvořit více dílčích modelů, každý popisující modelovaný systém na jiném frekvenčním rozsahu. Takovýto přístup k modelování je vhodný např. pro modelování systémů s několika významnými módy. Samotná možnost volby specifického frekvenčního rozsahu využitím pro modelování je s výhodami použitelná při identifikaci systémů v inženýrské praxi, kde model dobře popisující pouze významné mody systému může být jednoduchým a zároveň naprosto dostatečným pro následné řízení.

Tato disertační práce navíc ukazuje vlnkovou transformaci, její bázové funkce a vlastnosti z několika úhlů pohledu, což čtenáři umožňuje detailní pochopení nejen samotné vlnkové transformace, ale i jejích souvislostí s teorií identifikace systémů. Mnoho souvislostí těchto dvou teorií jsou ve své podstatě velmi intuitivní a mohou tak být jednoduše přenositelné mezi různými vědeckými obory.
Nejenom algoritmy popsané v této práci, ale i vzájemné propojení obou teorií vlnkové transformace a identifikace systémů je jedním z hlavních přínosů této disertační práce. V dnešní době existuje velmi mnoho různých profesionálních výpočetních programů, jenž jsou velmi známé a široce používané lidmi napříč vědeckými obory a dokonce i mimo akademický svět. Bohužel, mnoho velmi silných matematických nástrojů (metod, postupů, přístupů, algoritmů apod.) v těchto programech implementovaných je dobře známo a široce používáno pouze ve specifické komunitě lidí sdílejících příbuzný obor. Mimo tuto komunitu je použití výše zmíněných nástrojů spíše výjimkou. Je proto velmi důležité nejen vyvíjet nové metody, ale i hledat vhodné, již vyvinuté, metody napříč vědeckými obory. Pro takové hledání je velmi silný důvod i v případech, kdy z principu nemohou z nového spojení teorií vzejít lepší výsledky. Tímto důvodem jsou zřejmě nové pohledy na příslušné teorie a popis nových analogií mezi nimi. I tyto malé střípky mohou být inspirující a vést k větším a dalekosáhlejším výsledkům.