Hand Detection Algorithm: Pre-processing Stage

Raissa Likhonina
Department of Signal Processing, Czech Academy of Sciences, Institute of Information Theory and Automation, Pod Vodárenskou věží 4, Prague 8, Czech Republic

Keywords: QRD RLS Lattice Algorithm, Hypothesis Testing, System Identification, Exponential Forgetting, Hand Detection, Ultrasound.

Abstract: The present work describes a new approach to hand detection based on QRD Recursive Least Squares (RLS) Lattice algorithm and probabilistic approach to system identification. The described method is supposed to be used as a pre-processing stage for a hand gesture recognition application based on ultrasound technology. The approach includes a noise cancellation concept and uses linear Finite Impulse Response (FIR) based regression models in MATLAB environment. Within the algorithm the hypothesis testing technique is implemented. The work shows the results of computation using real data from an ultrasound device. The final version of the algorithm is supposed to be implemented on the embedded Xilinx Zynq device.

1 INTRODUCTION

Hand gesture recognition is a technique, which becomes very popular in many fields including automobile industry, smart home/buildings, applications for disabled people, etc. Various technologies used for collecting raw data from hand movement are available nowadays. They include wired gloves, stereo cameras, depth-aware cameras, thermal cameras and radar. The advantages and limitations of hand gesture recognition applications are well described in (Pradipa & Kavith, 2014; Premaratne, 2014).

In this work a different approach to a hand recognition problem is offered. This approach is based on ultrasound technology, which promises to ensure low cost, low power consumption and friendly human-machine interface (HMI).

The approach is supposed to be based on adaptive Recursive Least Squares (RLS) algorithms (Moonen, 1999), which are popular in digital signal processing applications.

In many works devoted to RLS algorithms it was pointed out that a certain difficulty in implementation of the algorithms on hardware platforms exists. It is caused by high computational complexity of RLS algorithms and certain problems with numerical stability (Kadlec, 1985; Kadlec, 1986; Moonen, 1999). Therefore, a large number of investigations were made to handle these issues. To decrease computational complexity, fast versions of RLS algorithms and several approaches for their optimization were developed (Cioffi & Kailath, 1984; Constantinides et al., 2004; Diniz, 2002; Fang et al., 2002; Lee et al., 1981; Moonen, 1999). To make the algorithms numerically robust, a so-called QR decomposition (QRD) of the information matrix was proposed (Bottomley & Alexander, 1991; Moonen, 1999; Regalia, 1993).

This work provides a new approach to hand presence detection. The method is based on QRD RLS Lattice algorithm with exponential forgetting (EF) and double precision arithmetic (Kadlec & Likhonina, 2016) and extended with hypothesis probability estimation of the model order (Kadlec, 1985; Peterka, 1981). The approach is supposed to serve as a pre-processing stage in a hand gesture recognition application based on ultrasound technology.

2 HAND DETECTION APPROACH

This section describes the algorithmic technique as well as the basic concept of the algorithm working with real ultrasound data, which were obtained from

Likhonina, R.
Hand Detection Algorithm: Pre-processing Stage.
DOI: 10.5220/0009885206950701
In *Proceedings of the 17th International Conference on Informatics in Control, Automation and Robotics (ICINCO 2020)*, pages 695-701
ISBN: 978-989-758-442-8

the device equipped with microphones and the ultrasound speaker.

2.1 Algorithmic Technique

In this work a noise cancellation technique well described in (Moonen, 1999) is used. The basic concept is: there are two signals – the desired signal and the reference ultrasound source signal. The desired signal consists of reflections both from the hand and from the environment. Using the reference ultrasound signal, the algorithm removes/reduces responses coming from the environment. The assumption is made that the signals are uncorrelated.

The noise cancellation problem here is solved using the QRD RLS Lattice in the error feedback form with double precision arithmetic and EF. The preference is given to this type of the algorithm due to its convenience in implementation and its computational speed. The algorithm can be derived from the QRD algorithm (Kadlec, 1985). It uses QR decomposition of the information matrix, which ensures its numerical robustness (Kadlec, 1985; Kadlec, 1986; Moonen, 1999). More detailed information about the algorithm derivation can be found in (Kadlec, 1985). However, a short description of the most important equations is provided here.

The derivation of Lattice algorithm is based on the following equations multiplied with a data vector (Kadlec, 1985):

$$\hat{\theta}_f(N+1|n) = \begin{bmatrix} \hat{\theta}_f(N+1|n-1) \\ 0 \end{bmatrix} + \begin{bmatrix} -\theta_{bz}(N+1|n) \\ 1 \end{bmatrix} \cdot \Lambda_z^{-1}(N+1|n) \cdot K(N+1|n) \quad (1)$$

$$\theta_b(N+1|n) = \begin{bmatrix} 0 \\ \theta_{bz}(N+1|n) \end{bmatrix} + \begin{bmatrix} 1 \\ -\hat{\theta}_f(N+1|n-1) \end{bmatrix} \cdot \Lambda_z^{-1}(N+1|n) \cdot K(N+1|n) \quad (2)$$

$$\bar{\theta}_{bz}(N+1|n) = \theta_b(N+1|n-1) \quad for\ n = 2, \dots, N \quad (3)$$

Note that $\hat{\theta}_f(N+1|n)$ is a vector of autoregression coefficients in forward direction;

$\theta_b(N+1|n)$ is a vector of autoregression coefficients in backward direction calculated from decomposition of matrix *V(n+1)*;

$\theta_{bz}(N+1|n)$ is a vector of autoregression coefficients in backward direction calculated from decomposition of matrix *V(n)*;

$\Lambda_z\ (N|n)$ are diagonal elements of matrix D in matrix decomposition $V^{-1}(n) = U(n) \cdot D(N) \cdot U^T(N)$, where U is an upper triangular matrix with units on the main diagonal, D is a diagonal matrix with positive elements on the main diagonal.

$\Lambda\ (N+1|n)$ are scalars in N different decompositions of $V(n+1) = L(n+1) \cdot D(n+1) \cdot L^T(n+1)$, where L is a low triangular matrix with units on the main diagonal.

$K(N+1|n)$ are coefficients of response.

After this step, relations between prediction (resp. filtration) errors in forward and backward directions can be obtained. On the basis of these variables, updating of parameters $\Lambda_z(N+1|n)$, $\Lambda(N+1|n)$, $K(N+1|n)$ is provided (Kadlec, 1985).

The relations for updates of the parameters $\Lambda_z(N+1|n), \Lambda(N+1|n)$, $K(N+1|n)$ are the following (Kadlec, 1985):

$$\bar{K}(N+1|n) = K(N+1|n) + \frac{h_z(N+1|n) \cdot e(N+1|n-1)}{1+\xi(n-1)} \quad (4)$$

$$\bar{K}(N+1|n) = \alpha^2 \cdot K(N+1|n) \quad for\ n = 1, 2, \dots, N \quad (5)$$

$$\bar{\Lambda}_z(N+1|n) = \Lambda_z(N+1|n) + \frac{h_z^2(N+1|n)}{1+\xi(n-1)} \quad (6)$$

$$\bar{\Lambda}(N+1|n) = \Lambda(N+1|n) + \frac{e^2(N+1|n)}{1+\xi(n)} \quad (7)$$

$$\bar{\Lambda}(N+1|n) = \alpha^2 \cdot \Lambda(N+1|n) \quad for\ n = 0, 1, \dots, N \quad (8)$$

$h_z(N+1|n)$ are prediction errors computed in backward direction from decomposition of matrix *V(n)*;

$e(N+1|n)$ are prediction errors in forward direction.

Let us remind also that for initialization step the following equations are valid (Kadlec, 1985):

$$h_1 = y_{t-1}; e_0 = y_t;\ \xi_0 = 0 \quad (9)$$

To update parameters (4)-(8), it is necessary to know forward prediction errors $e(N+1|n)$ and backward prediction errors $h_z(N+1|n)$. They are defined as follows (Kadlec, 1985):

$$e(N+1|n) = y_t - \hat{\theta}_f(N+1|n) \cdot Z(n) \quad (10)$$

$$h_z(N+1|n) = [-\theta_{bz}^T(N+1|n), \quad 1] \cdot Z(n) = y_{t-n} - \theta_{bz}^T(N+1|n) \cdot Z(n-1) \quad (11)$$

The interaction between forward and backward prediction errors of order *n-1* and *n* without using vectors $\hat{\theta}_f(N+1|n)$, $\theta_{bz}(N+1|n)$ is given by equations (Kadlec, 1985):

$$e(N+1|n) = y_t - \hat{\theta}_f(N+1|n) \cdot Z(n) = \\ = e(N+1|n-1) - K(N+1|n) \cdot \Lambda_z^{-1}(N+1|n) \\ \cdot h_z(N+1|n) \;\; for\; n = 1, \dots, N \tag{12}$$

where $e_0 = y_t$.

$$h_z(N+1|n) = [-\theta_{bz}^T(N+1|n), \quad 1] \cdot Z(n) = \\ = y_{t-n} - \theta_{bz}^T(N+1|n) \cdot Z(n-1) = \\ = h(N+1|n-2) \;\; for\; n = 1, \dots, N \tag{13}$$

where $h(N+1|n)$ are backward prediction errors computed from decomposition of matrix *V(n+1)* and

$h_0 = y_{t-1}$.

$$h(N+1|n) = h_z(N+1|n) - K(N+1|n) \cdot \\ \Lambda^{-1}(N+1|n-1) \cdot e(N+1|n-1) \, , n \\ = 1, \dots, N \tag{14}$$

where $e_0 = y_t$.

Thus, equations (12)-(14) define prediction errors $e(N+1|n)$, $h_z(N+1|n)$, which are necessary for updating parameters $\Lambda_z(N+1|n), \Lambda(N+1|n)$, $K(N+1|n)$ according to equations (4)-(8). It should be also noted that in this case there is no need to update matrix *V(n)* (Kadlec, 1985).

It is worth mentioning that the algorithm computes all necessary parameters for calculating probability density function $p(y_t|D(t-1); n_t)$ (Kadlec, 1985).

The structure of QRD RLS Lattice algorithm can be shown in Fig. 1 (Kadlec, 1985), where means multiplication with a constant and stands for "memory cell", with which help the following equation is performed: $h_z(N+1|n) == h\;\;(N+1|n-1)$, where $h\;\;(N+1|n-1)$ is from the previous time step.

The previously described Lattice algorithm can be changed to obtain filtration errors instead of prediction errors. This change allows deriving the normalized forms of the algorithms performing computation in range *(-1; 1)* (Kadlec, 1985), which decreases computation complexity and ensures numerical stability of the algorithm. The more detailed information see in (Kadlec, 1985).

The equations, which define relations between filtration and prediction errors, are as follows (Kadlec, 1985):

$$\bar{e}(N+1|n) = e(N+1|n) \cdot (1 - \bar{\xi}(n)) \tag{15}$$

$$\bar{h}_z(N+1|n) = h_z(N+1|n) \cdot (1 - \bar{\xi}(n-1)) \tag{16}$$

It is also valid:

$$1 + \xi(n) = \frac{1}{1 - \bar{\xi}(n)} \tag{17}$$

The detailed description of derivation of the equations can be found in (Kadlec, 1985).

It should also be mentioned that the QRD RLS Lattice algorithm is suitable for incorporation of hypothesis estimation. This is due to its modular structure, where each module can perform the order update. In its turn it allows obtaining estimations of all orders during computation (Pohl et al., 2008).

The hypothesis probability updates are based on the probabilistic theory and made as follows (Kadlec, 1985; Kadlec, 1986; Peterka, 1981):

$$p(H_n|D(t)) = \\ = \frac{p_n(y_t|u_t, D(t-1))}{p(y_t|u_t, D(t-1))} p(H_n|D(t-1)), \tag{18}$$

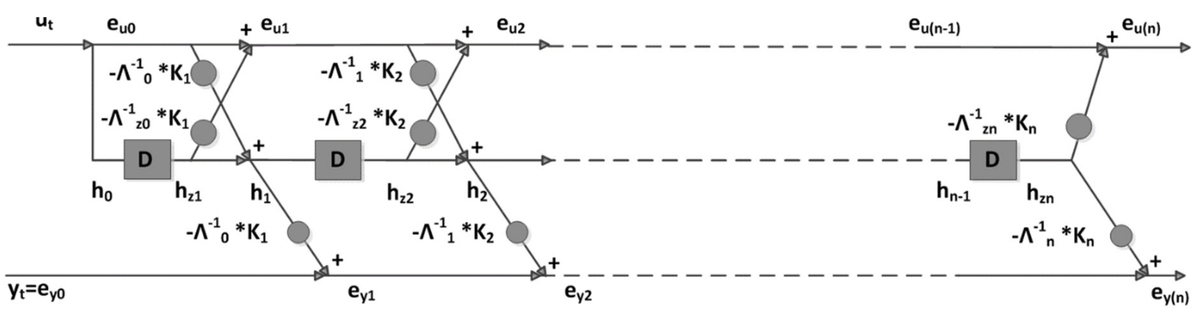

Figure 1: RLS QRD Lattice algorithm (Kadlec, 1985).

where y_t are data observed at time t, D(t) and D(t-1) are data previously observed from y_0 through y_t and y_{t-1} respectively, u_t is input, H_n is a hypothesis of a model order n. The term $p_n(y_t|u_t, D(t-1))$ is the probabilistic description of the modelled system with order given by hypothesis H_n.

From equation (18) it is clear that there are two stages of probability estimation. The first stage, which is presented by the numerator in (18), computes the order update. It means that the old probability estimates are updated by new data during the first stage. In the second stage, which is presented by the denominator of (18), the normalization of the updated order estimates is performed. The second stage fulfils the forgetting of hypothesis probability density function (Pohl et al., 2008). These stages can be incorporated into the QRD RLS Lattice algorithm.

It is obvious that to compute probability estimates of the hypotheses, it is necessary to know $p(y_t|D(t-1), H_n)$, which is calculated as follows:

$$p(y_t|D(t-1);\, H_n) = \\ = \pi^{-\frac{1}{2}} \cdot \frac{\Lambda^{\left(\frac{\vartheta-n}{2}+1\right)}}{\bar{\Lambda}^{\left(\frac{\bar{\vartheta}-n}{2}+1\right)}} \cdot \frac{|V(n)|^{\frac{1}{2}}}{|\bar{V}(n)|^{\frac{1}{2}}} \cdot \\ \cdot \frac{\Gamma((\bar{\vartheta}-n)/2+1)}{\Gamma((\vartheta-n)/2+1)} \tag{19}$$

where Λ is the optimal solution error of the model for hypothesis H, $V(n)$ is an autocorrelation matrix, Γ is a gamma function, ϑ is the amount of data accumulated in matrix $V(n)$.

2.2 Real Data Experimental Approach

The block diagram presenting the basic concept of the experiments with real ultrasound data is shown in Fig. 2.

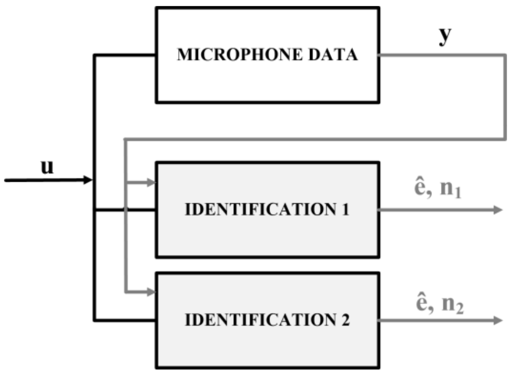

Figure 2: Block diagram.

The block diagram consists of three main parts:

- the block representing real data coming from the microphones,
- two identification blocks standing for regression models of two different orders.

The identification blocks are based on linear finite impulse response (FIR) models (Moonen, 1999). They represent regression models of two different orders, which allow hypothesis testing implementation. All three parts have common input u, which is modelled according to certain parameters described in the next section. The identification blocks estimate parameters of the hand model and computes prediction and filtration errors. Moreover, they estimate the order probability, which allows detecting the hand appearance.

Thus, if there is no hand in front of the device, the identification model with a smaller order will have a higher probability; otherwise, the identification model of a higher order will suit better for the estimation process. Based on the outputs of the order probability estimation, the hand presence can be detected.

The ultrasound data applied in the experiments were obtained from the ultrasound device developed in UTIA. The device comprises three basic components of the hardware platform represented by TE0820 FPGA SoM module, TE0706 carrier board and UTIA evBoard v1.7. The detailed information about the device can be found in (Pohl & Kohout, 2019).

3 RESULTS

A series of experiments have been performed to show functioning of the described approach. The results of the experiments are thoroughly discussed in this section.

As it was pointed out in the previous section, the real data used in the experiments are real ultrasound data obtained from the ultrasound device. The input signal provided by the ultrasound speaker is a set of chirps represented by a sinusoid wave with a period of 5 samples, a sampling frequency 192 kHz, and 880 samples space between them. For testing purposes, the input signal was simulated considering all above mentioned parameters and is shown in Fig. 3.

The output signal is obtained from 31 microphones situated on the ultrasound device. The ultrasound speaker sends the signal 500 times on the 40 kHz frequency. Each signal has 880 samples. The example of the output from one of the microphones is shown in Fig. 4.

The highest peaks on the graph represent the responses from the hand, while smaller peaks are signals coming back from other objects in the environment. The goal is to identify the hand with the help of hypothesis testing algorithm described in the previous sections.

As the first step, it is necessary to set orders for two identification blocks as well as to choose the optimal value of the EF factor.

After fulfilling a set of experiments, it was proved that the optimal orders for two regression models are n_1=2000 and n_2=100. The EF factor is set equal to 0.9999, the time period is N=305500. The results of identification process are shown in Fig. 5.

Figure 5 shows two graphs, which represent the results of hypothesis testing algorithm in different periods of identification process. The upper graph shows the beginning of identification process, while the graph on the bottom of the picture presents the results, when the algorithm has already learnt to identify parameters and order of the system correctly.

As it is clear from the picture, the red line stands for the high order of the regression model, i.e. n1=2000, while the green one is for the order n2=100. In the beginning of identification process the algorithm needs some time to start estimating parameters in a right way, but after appx. 9000 samples the estimation process converges to correct values and the algorithm functions precisely enough. It should be noticed that when the hand appears, the higher order system is considered to be more appropriate for identification process, while in the situation when there is no hand, the smaller order system is preferred, which is represented by a green line.

Based on the experiments, it can be concluded that the algorithm using hypothesis testing functions reliable and precisely enough and can be used for solving a noise cancellation problem.

The time of computation considering N=305500 and the highest order n_1=2000 is 81.0914 seconds, which is not satisfactory if there is a need for a quick real time data processing. Therefore, the algorithm needs accelerating.

4 DISCUSSION

The experiments prove that the proposed approach to system identification using QRD RLS Lattice algorithm is promising and provide precise outputs. One of the advantages of the approach is that the optimal value of the EF factor can be chosen and set permanently and it will not influence the results of hand presence detection. However, to set the orders of identification models, certain assumptions about hand distance from the device should be made, which can be taken as a limitation of the approach.

Based on the hypothesis order probability estimation, the system can identify the hand presence precisely. On the other hand, the identification process needs a certain amount of time to learn to converge to the right values, which is the other limitation of the approach.

However, the results of the experiments are promising and the further work is supposed to elaborate the ways of accelerating the algorithm and the ways of shortening its computation time.

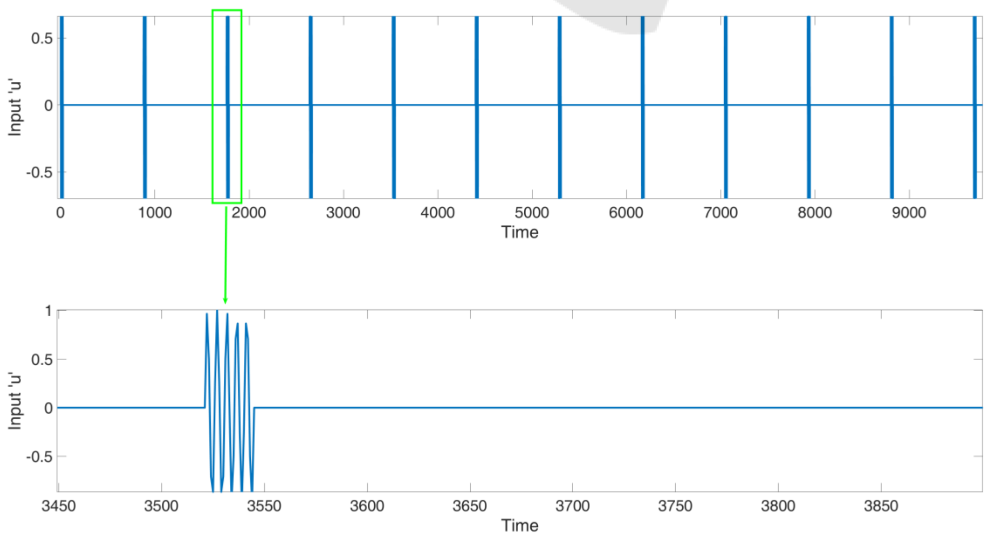

Figure 3: Input signal.

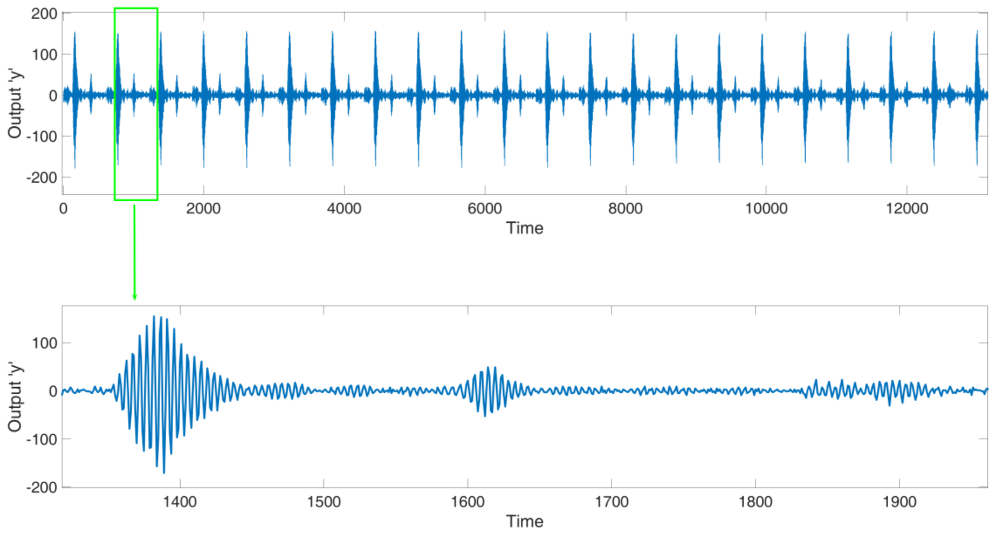

Figure 4: Output signal.

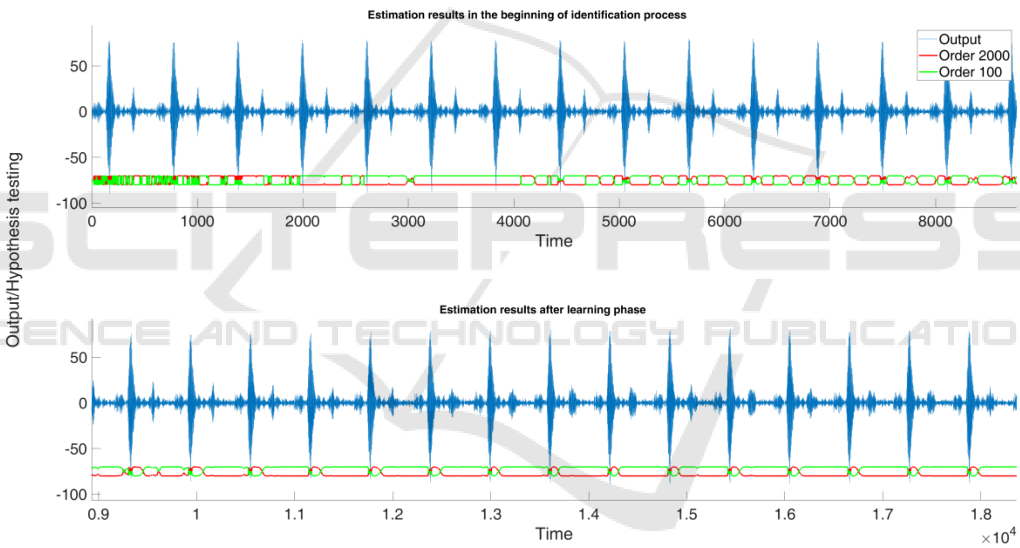

Figure 5: Hand detection.

5 CONCLUSION

The present work presents a new approach to a hand detection technique based on ultrasound technology. The approach describes a pre-processing stage used for a hand recognition algorithm. The algorithm used in the work is QRD RLS Lattice algorithm with double precision arithmetic and EF. The algorithm is extended with the hypothesis order probability estimation.

A series of experiments was performed with real data from the ultrasound device. The experiments show that the hand detection process is not very sensitive on a choice of the EF factor. The results are precise enough; however, to process data in real time, the algorithm needs accelerating.

The final goal is the application for hand detection, where the hand appears for a short period of time at a certain distance from the ultrasound source. The final version of the algorithm is supposed to be implemented on Xilinx Zynq devices operating in real time with a microphone and ultrasound transducers. The final implementation is supposed to benefit from FPGA structure and pipeling technique to accelerate the computation process.

ACKNOWLEDGEMENTS

This work has been supported by the ECSEL JU project SILENSE "(Ultra)Sound Interfaces and Low Energy iNtegrated Sensors" Project No.: ECSEL JU 737487-8 and MSMT 8A17006 (Ministry of Education Youth and Sports of the Czech Republic).

REFERENCES

Bottomley, G. E., Alexander, S. T. A novel approach for stabilizing recursive least squares filters. *IEEE Transactions on Signal Processing,* vol. 39, pp. 1770-1779, 1991.

Cioffi, J., Kailath, T. Fast recursive-least-squares transversal filters for adaptive filtering. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 2, pp. 304-337, 1984

Constantinides, G., Cheung, P. Y. K., Luk, W. Synthesis and optimization of DSP algorithms. Kluwer, Dordrecht, 2004.

Diniz, P. S. R. Adaptive filtering: algorithms and practical implementation. Second edition, Kluwer Academic Press, Norwell, MA, 2002.

Fang, F., Chen, T., Rutebnbar, R. A. Floating point bit-width optimization for low-power signal processing applications. *Proc. Int. Conf. on Acoustics, Speech and Signal Proc.*, vol. 3, pp. 3208-3211, 2002.

Kadlec, J. Continuous probabilistic identification of autoregression model with unknown order. In: *Analýza, syntéza a rozpoznávání řeči*, ČSVTS, Prague, 1985.

Kadlec, J. Probabilistic identification of regression model in fixed point. Ph.D. thesis, UTIA CAS, Czech Republic, September 1986.

Kadlec, J., Likhonina, R. Adaptive RLS algorithms reference implementations (application note). UTIA, 2016.

Lee, D., Morf, M., Fridlander, B. Recursive least-squares ladder estimation algorithms. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 3, pp. 627-641, 1981.

Moonen, M. Introduction to adaptive signal processing. K. U. Leuven, Leuven, Belgium, 1999.

Peterka, V. Bayesian approach to system identification. In: *Eykhoff, P. (Ed.), Trends and Progress in System Identification*. Pergamon Press, Oxford, pp. 239-304, 1981.

Pohl, Z., Tichy, M., Kadlec, J. Implementation of the least-squares Lattice with order and forgetting factor estimation for FPGA. *EURASIP Journal on Advances in Signal Processing*, pp. 1-11, 2008.

Pohl, Z., Kohout, L. UTIA evaluation board v1.7-v1.8. Beamforming demo (application note), UTIA AV CR, v.v.i., 2019.

Pradipa, R., Kavith, S. Hand gesture recognition – analysis of various techniques, methods and other algorithms. *International Conference on Innovations in Engineering and Technology (ICIET' 14).* Available at http://www.rroij.com/open-access/hand-gesture-recognition--analysis-ofvarious-techniques-methods-and-theiralgorithms.pdf

Premaratne, P. Human computer interaction using hand gestures, cognitive science and technology. DOI: 10.1007/978-981-4585-69-9_2, Springer Science+Business Media Singapore, 2014

Regalia, P. A. Numerical stability properties of a QR-based fast least squares algorithm. *IEEE Transactions on Signal Processing*, vol. 41, no. 6, pp. 2096-2109, 1993.