Fault Detection and Localization on Underwater Vehicle Propulsion Systems Using Principal Component Analysis

N. Mišković* and M. Barišić*

* Faculty of Electrical Engineering and Computing/Dept. for Control Engineering, Zagreb, Croatia
e-mail: nikola.miskovic@fer.hr, matko.barisic@fer.hr

Abstract—**In order to achieve full and viable autonomy of an underwater vehicle, it is of high importance that the vehicle can operate in at least a foreseeable, if not an unhindered way, under conditions of a technical failure or breakdown. Thus, a technique for dealing with a broad range of possible faults and failures is needed. Every such technique can be methodologically separated into a sub-technique for fault detection and an adaptive mechanism on the control law or behavior. This paper concerns itself with the problem area of fault detection and localization. It describes the mathematical method of Principal Component Analysis and provides guidelines for its use on fault detection and localization. The benchmark example used herein is a DC motor as a simplified model of a propulsion subsystem of an AUV. The particular algorithm in the PCA class that will be elaborated upon in this paper is the Partial Principal Component Analysis. The paper gives results of simulation in Simulink, comments on the successes and shortcomings of the method, and contains proposals for future research.**

I. Introduction

The subject of the paper is fault detection and localization in an underwater vehicle propulsion system. Fault detection and localization is one of the two principal parts of a system that enables fault tolerant operation of an automated system. The other part is a hybrid or deliberative policy or an adaptation mechanism that changes the control law for a particular failure occurrence. This area of research is of critical importance in the control engineering of underwater vehicles, especially of the semi- or fully autonomous regime. An autonomous underwater vehicle operates fully immersed in an isolated environment. This environment is ripe with physical effects and challenging impositions to control law synthesis that are to the most part counterintuitive and outside of standard experience of the human crew aboard the base vessel. The environment is relatively poorly explored, human presence is minimal, and cost-per-mission of any human endeavor within it is prohibitive. Furthermore, the physical properties of the permeating medium are such that both sensing and communication with the base vessel is difficult and mounts its own set of problems. Therefore, the ability of any human-originated instrument, such as an AUV, to continue to operate in spite of technical faults is paramount. This is achieved by robustly and correctly implementing a fault tolerant control mechanism. The actuators of the AUV must be guided to the greatest possible efficiency and a mission profiler must be aware of the occurrence of a fault in order to make a decision to return to base or position itself in a location, orientation or position which minimizes the cost of the reclamation and salvage mission. All of this depends critically on the correctness and speed of the failure detection and localization. This paper will present one of the techniques for achieving a robust, trustworthy detection of a technical fault on the actuator subsystem of an AUV, the Principal Component Analysis approach. In section II, it will give explanation and mathematical foundation for the fault detection and localization algorithm, as well as lay out anticipated modifications that open up a plethora of additional possibilities of this approach. In section III, the algorithm laid out in section II will be tested and the trustworthiness and speed of fault detection and localization determined. Parameters leftover from the synthesis in section II to be determined empirically and boundary conditions will be looked into and robust values assigned. For the purposes of experimental verification of the results of synthesis in section II, the propulsion system will be described as a simple linear, time-invariant DC motor.

II. Synthesis Of The Fault Detection And Localization Algorithm

A. *Principal Component Analysis*

Principal component analysis (PCA) is a linear method which transforms a multidimensional space to a space with fewer dimensions. This is due to the fact that the greater the correlation between the data, the greater their redundancy. If that is the case, the data can be compressed in a set of fewer dimensions. Here, compressed assumes that a proper qualitative understanding of the system's characteristics is maintained while the number of dimensions decreases. Since the number of dimensions is a strong measure of how much computing power is needed aboard an AUV, this is an alluring possibility. Computing power translates either into the power rating of the onboard electronics, and thus indirectly into the rating, weight and capacity of the batteries, or the time necessary to perform calculations, which destabilizes and places constraints on the control law. Both of these manifestations of the necessity for higher computing power impinge upon the autonomy and base-craft independence which are defining qualities of any viable AUV.

The main objective of the PCA is to define a subspace of the problem space such that a mapping of all vectors to the subspace maintains a high level of problem-relevant information. For the problem of fault detection, problem-relevant information is any such that contributes to the ability of the on-board software to detect and localize an occurrence of a technical failure. This subspace is defined by l orthogonal vectors in the problem space, where l is any number smaller than the dimension of the problem space. The constraint on the smallness of l is the necessity of information conservation. In mathematical terms, the subspace, i.e. the linear combination of l orthogonal vectors, must have the greatest possible variance.

Let **X** be a matrix of measured data in the form of

$$\mathbf{X}=\begin{bmatrix} x_1(1) & x_2(1) & \cdots & x_m(1) \\ x_1(2) & x_2(2) & \cdots & x_m(2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(n) & x_2(n) & \cdots & x_m(n) \end{bmatrix} \quad (1.1)$$

where m is the number of observed variables and n the number of time samples. According to [2], [3], the matrix **X** can be written in the form of

$$\mathbf{X}=\mathbf{T}\mathbf{P}^T \quad (1.2)$$

where **T** is the matrix of projections (scores) and **P** the matrix of principal components.

Matrix **P** has to consist of such vectors so that the variance of the projection **T** on the vector is maximal. In order to fulfill the previous condition, it can be shown that matrix **P** has to consist of eigenvectors of the covariance matrix $\mathbf{\Sigma}=\mathrm{cov}(\mathbf{X})\triangleq\mathbf{X}^T\mathbf{X}$. The eigenvectors have to be sorted respectively to the descending eigenvalues.

There are several criteria by which the number of principal components can be determined. The one that will be used in this work is that the number of chosen principal components must describe at least 80% of the total variance. One of the ways to determine the percentage of the total variance is using a cumulative sum. The cumulative sum gives the percentage of the described variance using eigenvalues of each principal component. It can be calculated using (1.3)

$$s_{i,cum}\left[\%\right]=\frac{\sum_{j=1}^{i}\psi_j}{\sum_{k=1}^{r}\psi_k}\times 100, \quad (1.3)$$

where ψ_i is the eigenvalue, r the total number of eigenvalues and i the number of chosen principal components [2].

Once the number of principal components (l) is determined (normally it is $l<m$), the initial matrix **X** can be described with l greatest eigenvalues (eigenvectors) of the covariance matrix **Σ**. In other words, matrix **X** can be expressed with (1.4) and (1.5)

$$\mathbf{X}=\mathbf{T}\mathbf{P}^T+\mathbf{E} \quad (1.4)$$

$$\mathbf{E}=\tilde{\mathbf{T}}\tilde{\mathbf{P}}^T, \quad (1.5)$$

where $\mathbf{T}=\begin{bmatrix}\mathbf{t}_1 & \mathbf{t}_2 & \cdots & \mathbf{t}_l\end{bmatrix}$ and $\mathbf{P}=\begin{bmatrix}\mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_l\end{bmatrix}$ represent principal components, and $\tilde{\mathbf{T}}=\begin{bmatrix}\tilde{\mathbf{t}}_{l+1} & \tilde{\mathbf{t}}_{l+2} & \cdots & \tilde{\mathbf{t}}_m\end{bmatrix}$ and $\tilde{\mathbf{P}}=\begin{bmatrix}\tilde{\mathbf{p}}_{l+1} & \tilde{\mathbf{p}}_{l+2} & \cdots & \tilde{\mathbf{p}}_m\end{bmatrix}$ represent the space of noise (residual). With this procedure, the space dimension is reduced from m to l and a statistical model of the process is created.

B. Hotelling's T^2 Statistics

In order to be able to construct control laws that stratify abnormal behaviors of the AUV, it is necessary to develop an algorithm which produces a definite decision whether, and if so, which kind of failure has occurred. Since in subsection A we produced a statistical model of the system, this algorithm has to give a definitive and illuminating analysis of un-modeled (residual) dynamics. The results of this analysis must point to the answer whether the difference between the measured variables of the realized, factual process and our statistical model occur due to the dimensional compression or a fundamental model mismatch. The latter answer actually assumes that a technical fault has occurred such that it has, as a result, "spoilt" the dynamics of the factual, realized process in relation to the nominal process used for the extraction of the sub dimensional statistical model. Since the extracted sub dimensional model is of statistical nature (including a residual value), it is clear that a statistical measure is needed. Using the statistical model produced according to subsection A, it is possible to make conclusions on fault appearance by observing the statistical measure of the residual that covers the un-modeled dynamics. There are many types of statistical measures but the one that we have decided to use is the Hotelling's T^2 statistics.

T^2 statistics or Hotelling's distance can be defined as:

$$d_i=(\mathbf{x}_i-\mathbf{m})(\mathbf{P}\mathbf{D}\mathbf{P}^T)^{-1}(\mathbf{x}_i-\mathbf{m})^T\frac{I(I-R)}{R(I^2-I)} \quad (1.6)$$

where x_i is the i-th row in matrix **X** and **m** is vector of mean values of referent data [10], [11]. Matrix **D** is a diagonal matrix containing eigenvalues of **Σ**. R is a number of principal components of the referent model and I number of charge modeled in the referent model. Calculated value d_i becomes the actual residual signal, whose value is monitored during the AUV mission and figures in the fault-detection logic.

If the process is working in a fault-free regime, the calculated statistic measure should not vary considerably. On the other hand, if the fault is injected, this measure should show deviation from its fault-free value.

C. Creating a statistical model

If an underwater vehicle is performing a maneuver, then the set point value for at least some, and in the case of a detailed dynamical model (which models DOF coupling), most or all of the 6 DOF, varies in time. I.e, the AUV control block is performing in a *tracking*, rather than *regulation* mode. Therefore, it is not possible to arrive at a statistical model because the behavior of the system is not predictable, in a simple, algebraic sense. This points to the constraint of fault detection methods relying on the usage of the statistical model of the AUV. It follows that such methods, among which is the PCA approach, are limited to regimes of AUV navigation and operation where the set point signal is constant for a reasonably long period of time. In other words, set points for the measured variables must have definitive and unchanging values on a timescale much greater than the transients of the control loop.

A statistical model is based on finding mean values $\overline{x}_i$ and deviations σ_{x_i} of measured states in a fault-free regime. This has to be done so the data acquired later would be normalized correctly using (1.7).

$$\begin{bmatrix} x_{10}(k) \\ x_{20}(k) \\ \vdots \\ x_{m0}(k) \end{bmatrix} = \left(\begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_m(k) \end{bmatrix} - \begin{bmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \vdots \\ \overline{x}_m \end{bmatrix} \right) \begin{bmatrix} \frac{1}{\sigma_{x_1}} \\ \frac{1}{\sigma_{x_2}} \\ \vdots \\ \frac{1}{\sigma_{x_m}} \end{bmatrix} \tag{1.7}$$

This acquisition can be performed either offline or online. The offline method is based on building a database with mean values and deviations of measured data for every possible value of a set point signal (say using test runs in a controlled safe environment or by running simulations on a detailed, first-principles model). This approach requires significant memory space but offers the advantage of fault detection being possible immediately after the transient is over. The online method requires a finite fault-free period in which the mean values and deviations can be calculated based on the data acquired in this period of time. The greatest advantage of this approach is that there is no need for great memory stores but the downside is the requirement of a finite fault-free period. Weighing the relative merits of the two approaches, it is easily discernible that the online approach is much better suited for use aboard AUVs. Following this line of reasoning, we have elected to use the online approach in the experiments.

D. *Partial PCA (PPCA)*

Up to this point it was assumed that the data matrix **X** consists of all measured output states. This approach can works for fault *detection*. However, it is not possible to make distinction between different faults i.e. perform fault *localization* using this approach. This is the reason why partial principal components analysis (PPCA) is introduced.

PPCA takes reduced sets of measured data into consideration. The objective is to generate many sets of data and calculate PCA and Hotelling's statistics for each. These calculations are the same as in PCA apart from matrix **X** being of smaller dimensions (the number of columns is smaller due to smaller number of measured data in each set).

III. EXPERIMENTAL VERIFICATION

A. *Process Description and Fault Scenario*

The system used to test the ability of the PCA method to identify the occurrence and nature of the faults is a common linear time-invariant DC motor in the propulsion unit of the AUV. Its parameters and performance indices do not influence the stability or effectiveness of the algorithm because the method is not model-based. The input signal into the system is set point speed ω_{SP}. Measured state variables are measured speed ω, measured armature current i_a and the output of the current PI controller u_a.

Faults whose effects and identification is explored with the experiment are:

Fault 1: Increase of the load inertia. In a normal work regime, a physical abruption of the motor can occur, e.g. a cord can get wound up around the shaft or a piece of underwater debris can get stuck in the propeller. This type of fault causes the inertia of the load to increase, which can eventually cause irreparable damage via inciting the DC motor to spark and short-circuit.

Fault 2: Drift of the speed sensor. Over time, marine speed sensors can accumulate a drift-type error. This occurs with the clogging and corrosion of the impeller and the shaft of the tachometer exposed to adverse environmental conditions, the oxidation of the windings within the tachometer and other effects of wear and tear. This fault usually occurs gradually, and the detection is possible once a certain value of the measurement drift is reached.

Fault 3: Drift of the armature current sensor. This fault can occur the same way as fault 2, over a prolonged period of time. The cause is the wear and tear of the current sensor i.e. aging-induced change of electrical characteristics of the measurement circuit and elements.

All of these supposed faults are additive and therefore are easy to model. The simulation scheme with the control loop of current and speed together with faults is shown in Figure 2.

The fault scenario is shown in Figure 1.

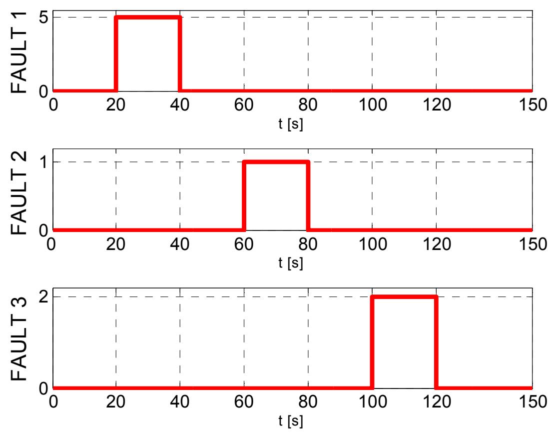

Figure 1: Fault scenario

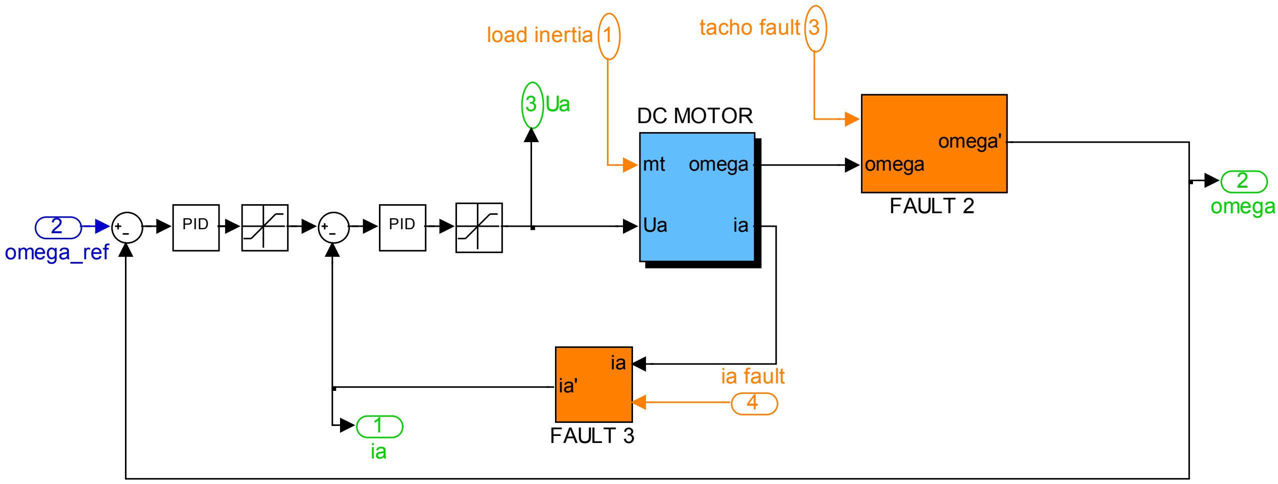

Figure 2: The detailed block diagram for the DC engine model with points and manner of additive fault injection

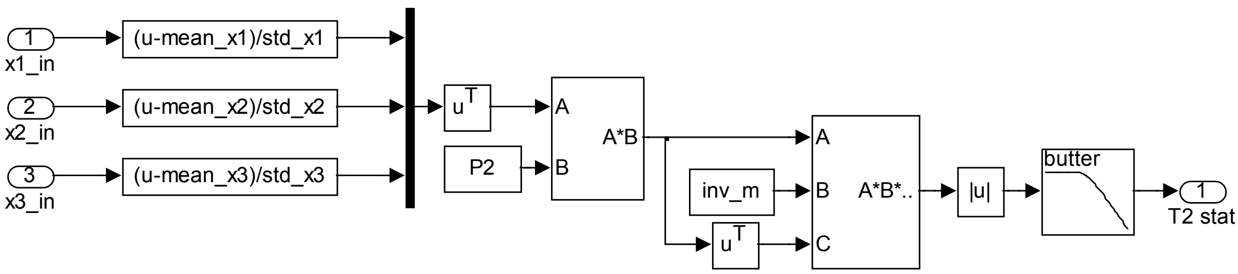

Figure 3: The detailed block diagram of the Hotelling's Statistic block in Figure 4

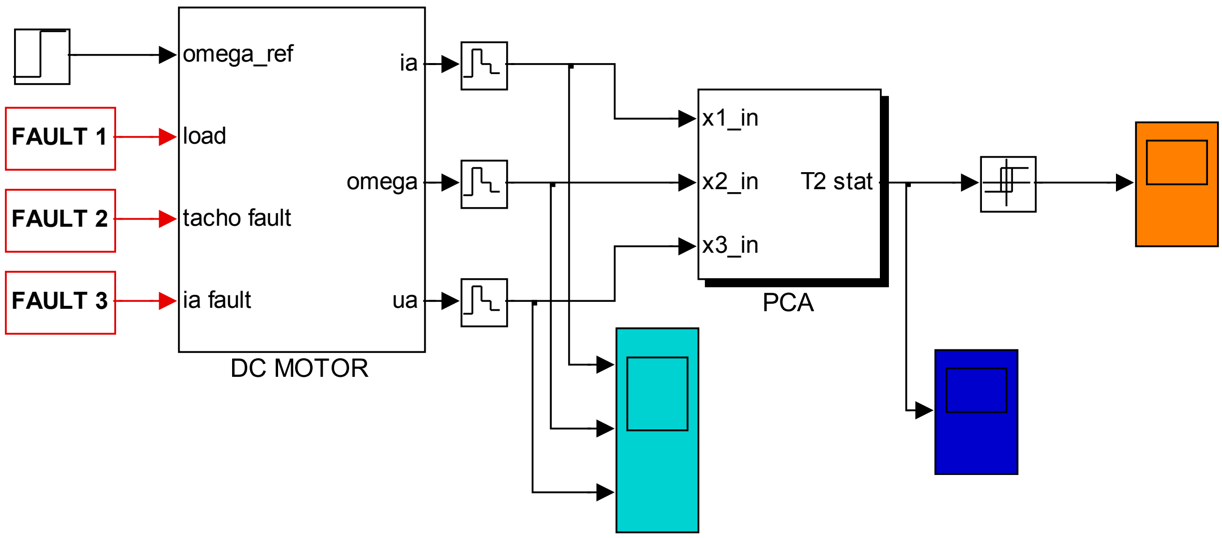

Figure 4: The total simulation block diagram for the PCA fault detection algorithm for a DC motor

B. Use of PCA on a DC Motor

Using the PCA method, the data matrix **X** has to consist of all measured outputs. The samples taken into consideration are those produced in the interval between 5s and 20s. This corresponds to the fault-free regime, according to Figure 1. The contribution of each principal component can be determined using (1.3). From Figure 5 it is obvious that one principal component is enough to describe the process in principal component subspace.

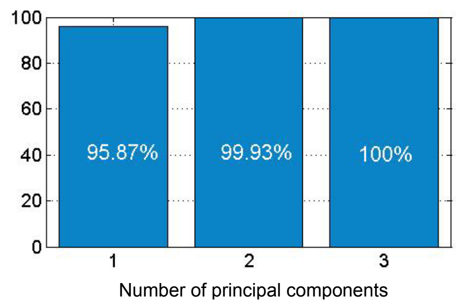

Figure 5: Level of contribution to absolute precision of the statistical model of the principal components in order of absolute eigenvalue

After calculating the matrix **P** and reducing it to the main eigenvalue (the matrix $\mathbf{P_2}$), we can produce a simulation scheme given in figure 3 which transforms the data to the reduced space and calculates Hotelling's statistics (in other words, creates a residual). The low-pass Butterworth filter at the end of the processing chain is used to smooth the residual. Its use is essential because with the online method Hotelling's statistics is calculated for every input sample so the residual signal can inherit significant noise from the state variables. This problem can be solved if Hotelling's statistic is calculated for an interval of input samples of the state variables, but this also corresponds to the usage of a low-pass filter albeit a moving average one.

Finally, in the residual generated with the PCA block is shown in figure 6.

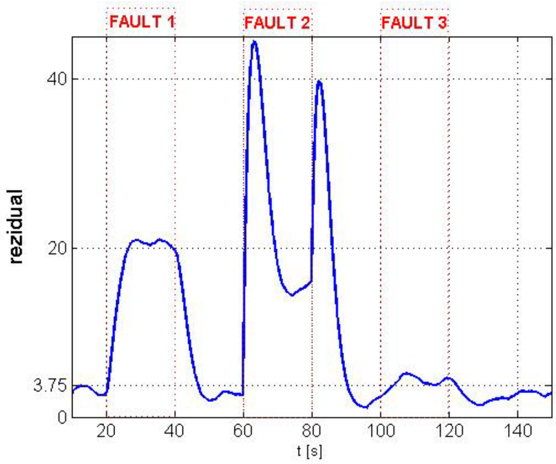

Figure 6: Tracked residual signal through the simulation with fault scenario in Figure 1

If a threshold is selected at the value of 3.75 (as is marked in figure 6) we can obtain fault detections as in figure 7. Times of detection are shown in **TABLE 1**. Note that the choice of the detection threshold influences both the measure of confidence and the speed of fault detection in a conflicting manner. Therefore, the best-case threshold should be the result of the solution to the optimization problem between these two conflicting criteria. However, these influences are very difficult to express analytically and consequently, an empirical setup of the threshold is preferred. The setup is a result of a set of borderline case studies using simulation software. With a threshold set up in this way, it is obvious that good results in fault *detection* are achieved. The problem with this approach is that the *localization* of the fault is impossible. The PCA block can detect faults on the basis of analyzing the overall residual signal over time. However, the amalgamated residual fails to be informative on which fault took place. This is the reason why Partial PCA is introduced in the next stage of the experiment. It is hoped that, in accordance with the theory, the separate residuals that it will generate will make *localization*, as well as *detection*, of the faults possible.

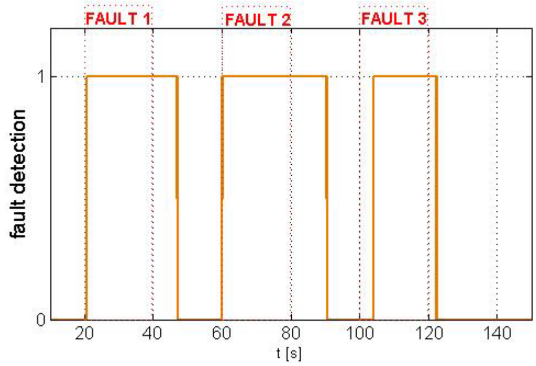

Figure 7: Logical determination of occurred fault based on residual levels in Figure 6

TABLE 1: OVERVIEW OF DETECTION SPEED FOR THREE SIMULATED TYPES OF FAULTS OCCURRING IN THE DC MOTOR FOR THE PCA ALGORITHM

times in [s]	START	DETECTED	END	DETECT. END
FAULT 1	20	20.65	40	46.95
FAULT 2	60	60.07	80	90.43
FAULT 3	100	102.67	104.07	122.33

C. *The Use of PPCA on a DC Motor*

In order to form a PPCA model it is necessary to group the three measured signals in all possible groups of two. That gives us three different PPCA blocks. The first takes ω and i_a, the second ω and u_a, and the third i_a and u_a. Using (1.3), we obtain contributions of each component as shown in figure 8.

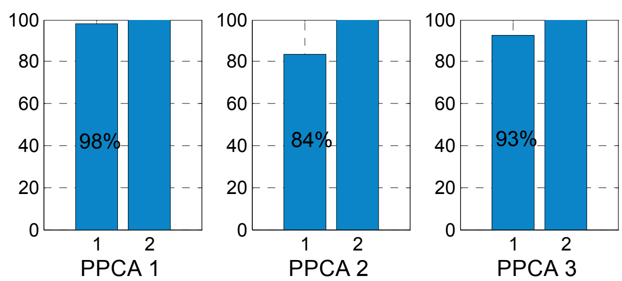

Figure 8: Level of contribution to absolute precision of the statistical model broken down by PC combinations for the PPCA localization in order of absolute eigenvalue

The PPCA blocks have the same form as the PCA block detailed in figure 3, with the exception that not all output signals are present. The total simulation scheme using PPCA method is shown in Figure 9.

The Block for detection and localization performs logical functions over the residuals "squared-off" via a relay set to the threshold value obtained empirically through simulation. Thus, the resultant residuals perform as logical functions with values exclusively 0 or 1, and can be subjected to logical functions. The full analysis of the system gives following connections between the residuals and the faults.

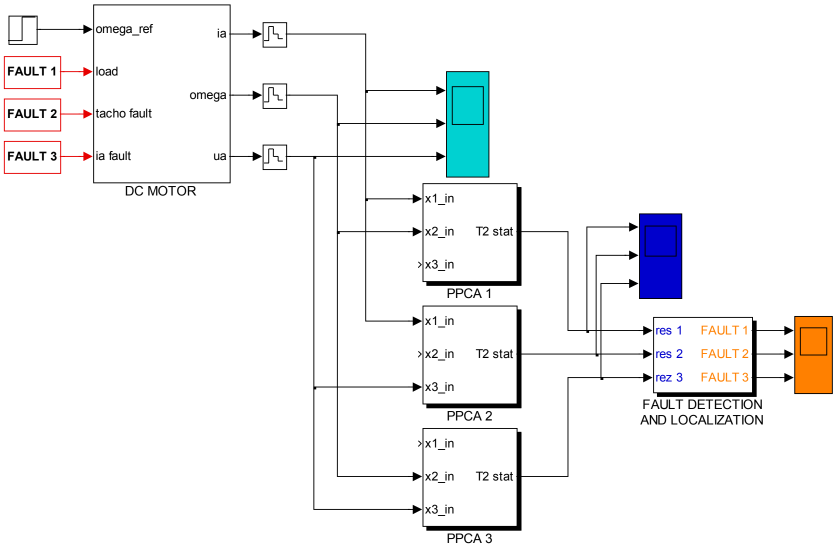

Figure 9: The complete block diagram for the PPCA fault detection and localization of faults on the DC motor

$$
\begin{aligned}
f_1 &= r_1 \wedge r_2 \wedge r_3 \\
f_2 &= \overline{r_1} \wedge r_2 \wedge r_3 \\
f_3 &= r_1 \wedge r_2 \wedge \overline{r_3}
\end{aligned}
\qquad (1.8)
$$

Using De Morgan's rules and the fact that complement equals $\overline{r} = 1 - r$, (1.8) can be written as a logical-algebraic equation in matrix form (1.9).

$$
\begin{bmatrix} \overline{f_1} \\ \overline{f_2} \\ \overline{f_3} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix} \begin{bmatrix} \overline{r_1} \\ \overline{r_2} \\ \overline{r_3} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \qquad (1.9)
$$

Input and output relays create a logic value. Threshold values, as was already described were chosen empirically on the basis of boundary-condition experiments. In this approach, the threshold values are constant, for the sake of simplicity and computational effectiveness. However, it is to be assumed that better performance would be achieved with an adaptive threshold. Such an algorithm for theoretically sound adaptation of the threshold value is a possible avenue for further research.

The residuals generated by PPCA blocks are shown in figure 10. The outputs from the block for fault detection and localization with the threshold of 0.1 are shown in figure 11. It is obvious that the detection and localization have been performed successfully. It can be seen that there are slight delays in fault detection. They are mainly a consequence of the way in which the residual generators are numerically set up. The dominant delaying influence in the signal processing chain of the residual is the low-pass filter. However, if there were no filter included, the output would have been noisy and the threshold values would therefore have been difficult to determine. With the low-pass filter included, the residual rises more slowly causing the late detection. However, the residual is much smoother and threshold is easier to determine. The graphs are also more intuitive in the case of graphical inspection. This could be exploited by having a fallback, last-ditch technique of eliminating potentially disastrous situations, using a human operator to decide the course of action, having conclusive information on the state of the AUV.

Table 2 includes times of detection.

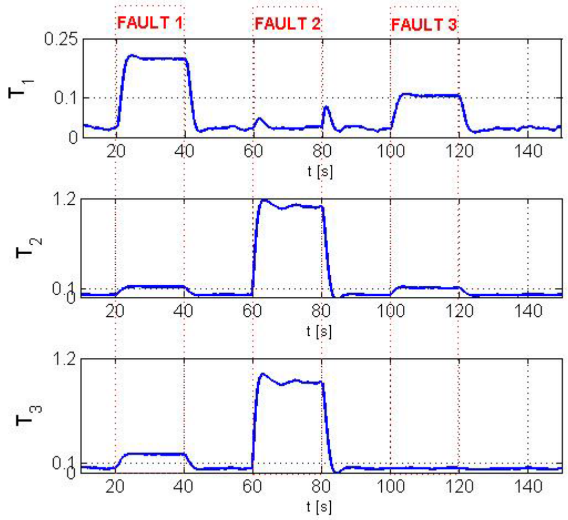

Figure 10: Tracked residuals generated by the PPCA algorithm which allows for fault localization

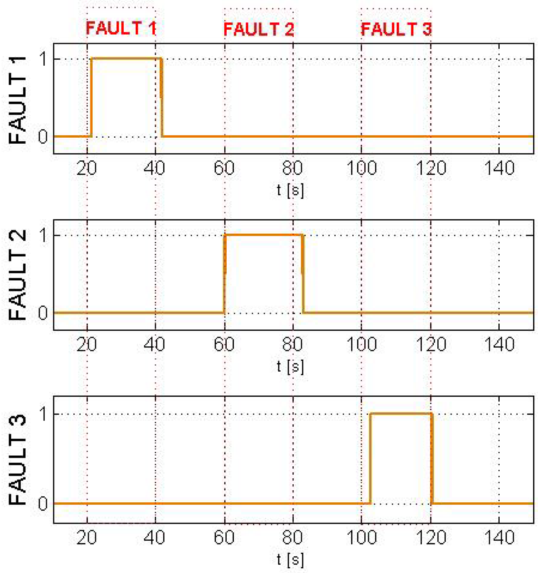

Figure 11: Resulting logical determination of fault occurrence as well as identification of the type of failure

TABLE 2: OVERVIEW OF DETECTION SPEED FOR THREE SIMULATED TYPES OF FAULTS OCCURRING IN THE DC MOTOR FOR THE PPCA ALGORITHM

times in [s]	START	DETECTED	END	DETECT. END
FAULT 1	20	21.3	40	41.67
FAULT 2	60	60.16	80	82.9
FAULT 3	100	102.67	104.07	120.33

Finally, a more realistic fault scenario is shown in figure 11 with gradual occurence of faults. Figure 11 also shows at which point the fault has been detected as well as the time when the detection detected the end of fault.

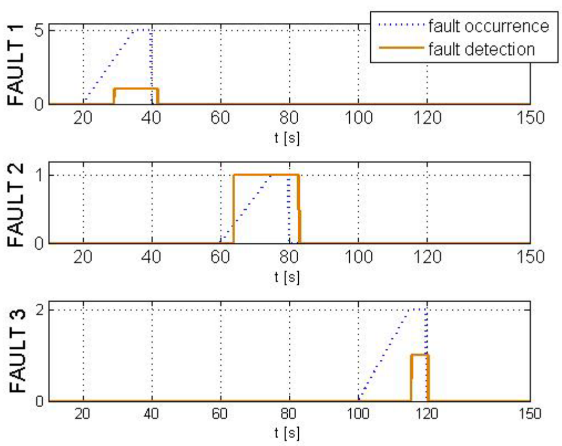

Figure 11: A realistic, linear introduction of additive faults into the system and time of its detection and localization by the PPCA method

This simulation is important because it shows at what value of fault the detection occurs. With the threshold determined earlier, fault 3 is detected when it gains the value of 2. Fault 2 is detected at about 0.3 while fault 1 is detected at about 2.5.

IV. CONCLUSION

PCA is a statistical method for fault detection and localization especially interesting to control engineers working with AUVs. The greatest advantage is that it is not necessary to know the model of the system in order to make a conclusion on fault appearance. This means that the method is appropriate for systems that cannot be easily or ever modeled, or for which the model is nonlinear, hybrid, or structurally ill-posed. Another advantage of the PCA is that, in addition to the additive faults, it is possible to detect multiplicative faults too.

The only requirement set on the method is the constructability of the statistical model. As was mentioned in subsection II.A, statistical model cannot be created for unpredictable or analytically challenging changes of the control signal. That is the greatest disadvantage of the method.

Further research might concentrate on improving detection dynamics and robustness. This could be achieved by following through on ideas tied to post-filtering the residuals through a nonlinear or adaptive filter of some sort, or by implementing an algorithm for the adaptation of the detection threshold itself.

REFERENCES

[1] Z. Vukic, H. Ozbolt and D. Pavlekovic ,*Improving Fault Handling in Marine Vehicle Course-Keeping Systems*, IEEE Robotics and Automation Magazine, June 1999.

[2] Jackson, J. E., *A User's Guide to Principal Components*, John Wiley and Sons, Inc. 1991. pp. 1-25.

[3] S. Persin, 'Construction of supervisory systems for industrial processes with subsystems for prompt fault detection', Ph.D. thesis, University of Maribor, Faculty of Electrotechnics, Computing and Informatics, Maribor, 2004. (in Slovenian).

[4] M. Blanke, R. Izadi-Zamanabadi, S.A. Bøgh and C.P. Lunau, "Fault-tolerant control systems – A holistic view," *Control Engineering Practice*, vol. 5, no. 5, pp. 693–702, 1997.

[5] J. H. Kim and G. O. Beale, "Fault detection and classification in underwater vehicles using the T^2 statistic," *Proc. of Mediterranean Control and Automation Conference*, Dubrovnik, Croatia, June 2001.

[6] J. H. Kim and G. O. Beale, "Fault detection and classification in underwater vehicles with quantification and contributing variables," *Proc. of Mediterranean Control and Automation Conference*, Lisabon, PT, July 2002.

[7] G. O. Beale and J. H. Kim, "Fisher discriminant analysis and the T^2 statistic for process fault detection and classification," *Proc. of IEEE IECON 2002*, Seville, Spain, 2002.

[8] Basseville, M. and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*, Prentice–Hall, 1993.

[9] R. J. Patton, P. M. Frank and R. N. Clark, Fault Diagnosis in Dynamic Systems: Theory and Applications, Prentice–Hall, Englewood Cliffs, NJ, 1989.

[10] H. Hotelling, *Multivariate Quality Control*. In C. Eisenhart, M. W. Hastay, and W. A. Wallis, eds. *Techniques of Statistical Analysis*. New York: McGraw-Hill, 1947.

[11] http://www.itl.nist.gov/div898/handbook/pmc/section5/pmc543.htm