Detecting External Measurement Disturbances Based on Statistical Analysis for Smart Sensors

André Dietrich, Sebastian Zug, Jörg Kaiser
Department of Distributed Systems
Otto-von-Guericke-Universität Magdeburg
{dietrich,zug,kaiser}@ivs.cs.uni-magdeburg.de

***Abstract*—The transducer process of a sensor is interference-prone to environmental conditions or external disturbances depending on sensor type, measurement procedure etc. Dependable sensors are characterized by a broad independence of those factors or/and they can both detect situations that make a correct measurement impossible and validate the measurement result.**

In this paper we describe a statistical approach for the detection of faulty measurements caused by external disturbances. Our fault detection algorithm is based on a comparison of faultless reference measurements with current sensing values. Using this enhancement, a sensor becomes a real smart sensing device and supplies an additional validity estimation of each measurement. The approach was implemented and validated in a demonstration setup that integrates an infrared sensor array disturbed by a strong extraneous light.

***Index Terms*—disturbed sensor measurements, fault detection, non-parametric statistics, smart sensor**

I. Introduction

Sensor systems in industrial applications have to cope with harsh environmental conditions. Varying influences like dust, fog of oil or steam interfere the measurement process and produce an ageing of the sensor unit. The results are well known: faulty measurements or crashed sensors that significantly disturb an application up to complete system failure.

Hence, system designers endeavor to integrate fault tolerance mechanisms for sensor applications that detect faulty components and validate each measurement. For this purpose we enhance the idea of smart sensors developed in [1] to fault tolerant purposes. Our smart fault tolerant sensor combines different methods of fault detection and merges the results to a validity estimation of the measurement [2]. The general interface of the smart sensor offers additional information, the quality of a measurement now. This approach aims to detect an individual abnormal sensor behavior that observes each sensing unit. This information supplied by a fault tolerant smart sensor supports decision about forwarding of a respective message in the network and can be used for the selection and weighting of the measurement in a fusion engine. This concept reduces the communication and calculation effort in a sensor network.

Which sensor faults can be identified in a smart sensor application? Faulty sensor measurements occur due to external and internal reasons. Fig. 1 shows the three elemental components of a smart sensor – transducer, processing unit and network interface – and assigns the following fault types:

1) External faults are all parameters d_E that influence the environmental conditions in such a way that the measurement process is disturbed. An example for incorrect results caused by external changes is an ultra sonic distance measurement during variable pressures. Thus, over time, the acoustic velocity varies and the sensor produces an untrustworthy distance measurement.
2) The transducer maps a physical value on an electronic signal. The translation can be affected by electronic or mechanic faults d_T of the sensor system, for instance a broken power supply.
3) The third category, processing faults d_P, based on soft or hardware faults inside the smart sensor processing unit. This can be a complex bug in filter software or a hardware failure.
4) The last barrier is the communication via a network. The sensor measurements can be partially or completely lost if a crashed node jams the wireless sensor network for instance.

All four fault types are usually modeled with an additive or multiplicative influence to the correct measurement signal and with an abrupt, increasing or intermittent temporal behavior. The polymorphism of the impact enhances the fault detection to a complex challenge.

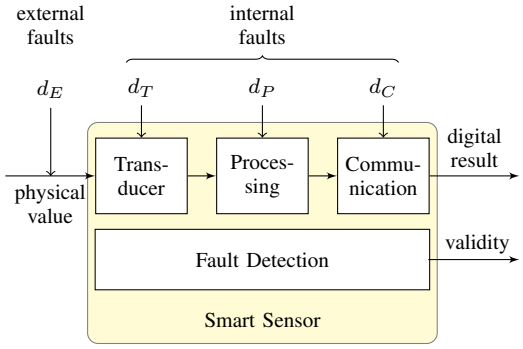


Fig. 1. Influence of faults on a smart sensor measurement process

In this paper we describe our approach for a detection of varying environmental conditions that have a negative influence on the measurement process. Those effects are called disturbances of measurement if the quality of the result decreases significantly. Complex and reliable sensor systems attempt to obtain all relevant environment information, for instance an ultra sonic sensor that is equipped with temperature, humidity and pressure auxiliary sensors. Based on whose additional measurements the influence of the disturbance can be eliminated. In many applications such advanced sensor systems are not available or the important environment variables are not known yet. In this case a statistical validation of the measurements offers a robust and inexpensive way to minimize the influence of improper measurements. The low cost infrared sensor we used for the validation allows distance measurements that can be disturbed by extraneous light. The influence level to the measurement depends on the light frequency, power, orientation related to the sensor and the reflectance of the environment. A coherent additional measurement of all these parameters is not realizable. However fault detection is necessary, for illustration we compare undisturbed and disturbed measurement series in Fig. 2. The sensor output voltage is depicted for different distances of an object (10-80 cm). An undisturbed measurement

DOI: 10.1109/ISIE.2010.5637754 [bibtex]

produces an output of 0.25 V for objects near to 80 cm, in case of a disturbance an obstacle 60 cm away from the sensor generates the same result. The displacement of the disturbed measurement ranges from 8 cm to −18 cm.

We tested different distance sensor based on laser systems (Wenglor YT87MGV80 [3]) or infrared light (Sharp Series) due to their susceptibility to faults caused by extraneous light. All sensors were affected by this disturbance but with different strength. We chose the sensor type with the highest measurement displacement, to illustrate the potential of our method to increase the fault tolerance aspect of a cheap sensor.

Due to missing information about the environmental conditions of each measurement we address the detection of external faults using statistical non-parametric test methods in this paper. A short state of the art in Sec. II introduces fault detection methods and afterward Sec. III points out different statistical test methods. Sec. IV presents our demonstration scenario, examines specific properties of the infrared sensors and illustrates the influence of extraneous light. In Sec. V we describe the selection and optimized parameterization of a test method and a validation of our approach using an experimental setup.

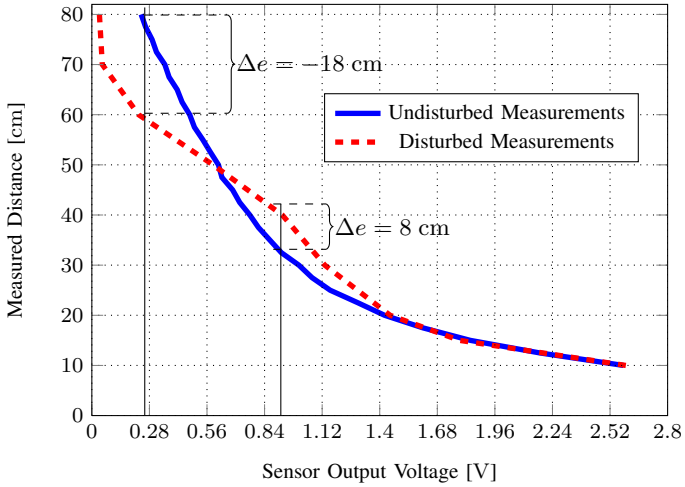


Fig. 2. Characteristic line of an infrared sensor with/without disturbance of extraneous light

II. Related Work

Sensor fault detection is based on a comparison of redundant information. This redundancy in measurement system can be obtained by three ways:

a) Hardware Redundancy: Hardware redundancy is used for safety critical applications in different ways [4]. Redundant, heterogeneous or homogeneous sensors measure the same or related values and observe a common area. Faulty sensors are identified by scanning all measurement deviations from mean, median etc. The criterion for an acceptance can be defined for instance as a simple threshold related to the measurement uncertainty or based on statistical knowledge like x-percentile, etc. [5]. Such methods discard a deviating minority as faulty measurements like a k-from-n voter. This means, the maximum number of simultaneous faulty sensors which can be detected is defined by the number of redundant measurements. If a general disturbance manipulates the majority or all measurements, like extraneous light in our scenario, those fault detection methods fail.

b) Model Based Redundancy: Another approach is to "simulate" a redundancy with a mathematical model of the observed system, for example [6]. The response to known inputs is calculated using this process model and compared to the reaction of the real system. If a state vector is assumed, the resulting residua can be classified by rule based systems, neuronal networks, fuzzy sets [7].

c) Signal Analysis: Model based redundancy uses the knowledge of the observed system to derive a validity of the measurement. Signal analysis monitors the parameters of the measurement process and models the transducer behavior. This is more robust in case of an uncertain behavior of the controlled system. Signal noise, frequency response, velocity of amplitude change, etc. are known as parameter of a measurement for an undisturbed system, a large overview of different methods for fault detection as well as diagnosis is given in [8] or [9]. In contrast to the most methods like parametric ARMA [10], spectrum/wavelett analysis [11], etc., our method does not require a specific model. For our individual signal analysis we compare single time series of measurements with faultless references samples (sensor's fingerprint) what can be done by statistical methods.

III. Parametric/Non-parametric Test Methods

Statistical tests can be divided into parametric and non-parametric methods [12]. Parametric test make specific assumptions with regard to one or more population parameters that characterize the underlying distribution(s) the test is employed for. In general parametric tests provide a more powerful analysis than analogue non-parametric tests but this advantage can be negated if one or more of its assumptions are violated.

Since we are not interested in dealing with different probability distributions but rather in comparing independent samples of an unknown distribution, non-parametric tests are always a good choice. One sample from an undisturbed measurement is used as a reference sample. We compare online a window of the last n measured samples with this reference. The test methods based on the assumption of a null hypothesis, for instance the equality of the distribution of the two samples. The output of the test is a p-value that indicates the probability of a valid null hypothesis. The tests we selected for further investigations in fault detection are:

1) *Ansari-Bradley Test*: tests the null hypothesis that the scale parameter of the distributions from which two samples were drawn are equal
2) *Fligner's Test*: tests the null hypothesis that all input samples are from populations with equal variances [13]
3) *Kolmogorov-Smirnov two-sided Test*: test for the null hypothesis that two independent samples are drawn from the same continuous distribution.
4) *Mann-Whitney U Test*: tests the null hypothesis for the equality of medians in two samples

An important advantage of the mentioned tests we used is that they allow to compare samples of different sizes. In the following sections we describe the selection and optimization process of a statistical test based on a careful analysis of the sensor system.

IV. Sensor System and Experimental Setup

A. Setup

Infrared distance sensors of the Sharp GP series are commonly used in robotic and mechatronic applications for environment detection due to the small measurement beams and low energy consumption. Sensors of the GP series determine the distance of an

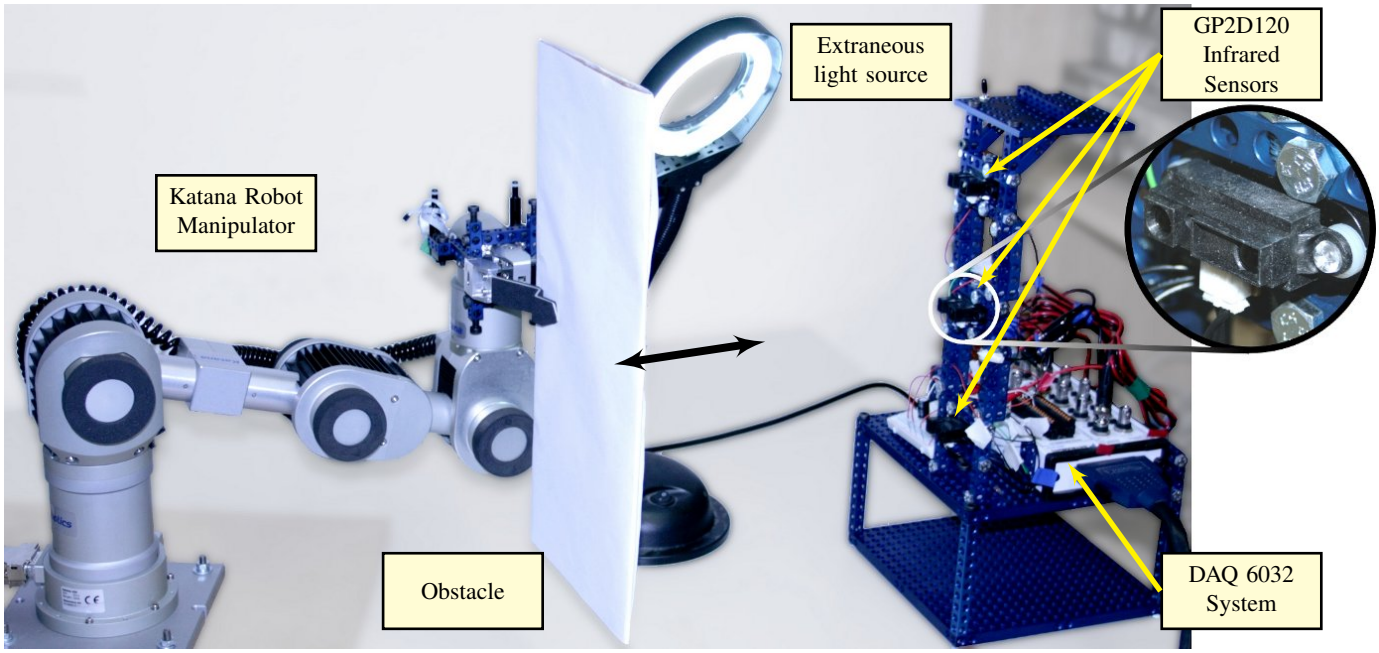


Fig. 3. Measurement Setup containing a robot manipulator, an extraneous light source, and the distance sensor system

object up to 500 cm by triangulation and output an analogue voltage level or a digital switching output. The relation between distance and voltage is ambiguous (for distances smaller than 10 cm) and non-linear. Disadvantages of the sensor systems are the interference of the measurement by extraneous light and a dependency on a very stable power supply.

Fig. 4. GP2D120 Infrared sensor, left-hand light emitter, right-hand light detector

The used GP2D120 sensors [14] depicted in Fig. 4 possess a refresh rate of approx. 40 ms and monitors distances from 10 cm to 80 cm with a resulting sensor output from 2.7 to 0.1 V. For our validation we combined three GP2D120 sensors which are assembled on a solid structure and connected to a Data Acquisition System. It consisted of a PCMCIA measurement card DAQ 6032 and a connection board BNC 2110 [15] produced by National Instruments. The system included a 16-bit analogue-digital converter. The fault detection was implemented in a Python script. For a reproducible obstacle movement a robot manipulator moved as shown in Fig. 3 a white board into the sensor beam. The measurement was disturbed by fluorescent tube which position was constant for all measurements.

B. Sensor Reference Measurements

Based on this experimental setup we analyzed the distribution of the measurements for constant distances. Fig. 5 shows the influence of the obstacle distance on the distribution of 10000 measurements by different colors. The contained smaller diagram points out the distribution for a distance of 15 cm and presents it in a histogram manner. Obviously the noise level, represented by the width of the distribution, increases with higher distances. The distribution shape varies from asymmetrical to a symmetrical distribution.

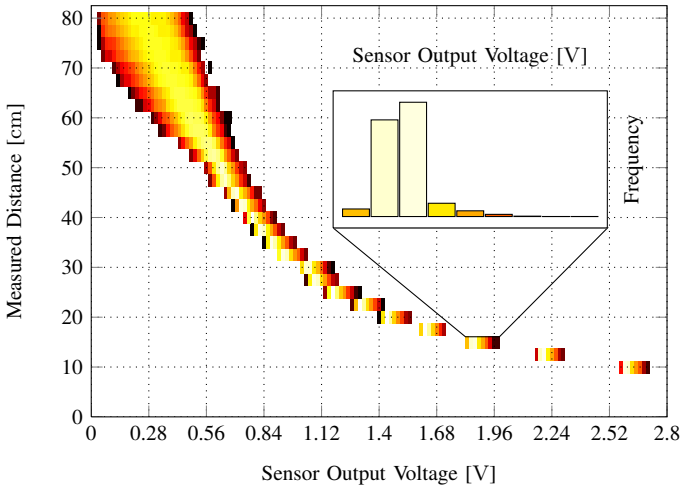


Fig. 5. Distribution of a GP2D120 sensor outputs for different object distances

The disturbances due to extraneous light are depicted in Fig. 2. The blue, solid reference line corresponds to the median of all distributions in Fig. 5. For simplification, this characteristic lines based on the median $x_{.50}$, the upper $x_{.75}$ and lower quartile $x_{.25}$ in each point were approximated by exponential polynomial functions $\{f_{.25}(v), f_{.50}(v), f_{.75}(v)\}$, while v means the voltage value. The divergence of approximation and reference measurements was smaller than 1.8 mm.

C. Determining the Probability Distribution

Related to our assumption, that the measurement noise cannot be characterized by a common probability distribution function, we checked a huge amount of data samples against different probability distributions. Modern computation makes it quite simple to do this. For the sake of generality we checked our measurement samples with the one-sample Kolmogorov-Smirnov test against all 81 continuous distributions contained in the Python SciPy.stats package [16]. Besides Gaussian, Exponential, Log-Normal or the Weibull distribution it also contains not very common ones such as the Nakagami or Rayleigh distribution which are used in communication theory. For a sufficient statement about the distributions of our measurements we checked more than 40 samples with at least 10000 values and afterwards averaged the results. The top five of those averaged results are presented with D and p-values in Tab. I. The D value denotes for the Kolmogorov-Smirnov test the largest absolute difference between the cumulative observed distribution and the cumulative distribution expected on the basis of the hypothesized distribution. As we can easily see, even if we take a significance criterion with $\alpha = .01$ into account, we have to reject all of the tested distributions.

TABLE I
AVERAGED RESULTS OF THE TOP FIVE TESTED PROBABILITY DISTRIBUTIONS

Range	Distribution	p-value	D
1.	Fisk	.008225	.030894
2.	Frechet Right Sided	.007979	.100676
3.	Johnson SB	.007874	.450270
4.	Generalized gamma	.005787	.637735
5.	Minimum Weibull	.005725	.102094

For future work it is noticeable, that the measurement noise cannot be modeled by a normal distribution. Smoothing or the very common prediction algorithms for fault tolerance purposes based on a Kalman-Filter are not applicable here.

In this case the best and even more general approach is to utilize non-parametric tests to be able to compare the equality of a measurement sample with a reference sample.

V. SELECTION AND OPTIMISATION OF A NON-PARAMETRIC TEST

Like mentioned in Sec. III, non-parametric tests have a lot of advantages and like for the most statistical tests a null hypothesis is expected, which has to be rejected if the calculated p-value is under some given bound. In our application the p-value can be reinterpreted as an indicator of disturbance, a high p-value means no or less disturbance in contrast to a lower value.

The properties an ideal test should own are enumerated below:

1) Reliability: small amount of false positives and false negatives
2) Low computational costs: a small reference sample and small sliding window
3) Short response time: short delay between the occurrence of a fault and its detection
4) Sharp distinction: huge distance between the p-value of disturbed and undisturbed measurements

In the following section we give an explanation on how we created a standardized reference sample on the basis of our measurements. Afterwards we examine the behavior of different non-parametric tests for varying sizes of the sliding window and the reference sample. Those values will be determined in the following evaluation step as well as the classification quality and speed.

A. Determining a common Reference Sample

To be independent from any probability distribution, we worked only with samples. The difficulty now was to determine the appropriate reference sample for all sensor measurements. As shown in Fig. 5, with increasing distance the voltage values decrease and the variance grows.

Because of small variations of the signal noise in our sensor setup, we generated a separate reference sample for every sensor. This may be traced to the fact that each sensor is characterized by individual noise behavior and limited precision of the measurement setup especially the analogue-digital converter.

As illustrated in Fig. 5 we took $u = 29$ undisturbed measurement samples with different distances. Related to standardizing of a Gaussian distribution [12] we shifted each sample by subtracting the median $x_{.50}$ and divided it by the interquartile range (lower quartile $x_{.25}$, upper quartile $x_{.75}$). In some literature the measurement range is used as a scale parameter instead of the interquartile range [17]. But in some of our measurements there were extreme outliers so that in this case it was better to use the more stable interquartile range:

$$Z_i = \frac{X_i - x_{.50i}}{x_{.75i} - x_{.25i}}, i = 1 \dots u$$

All u standardised samples were joined per sensor to a common sample Z. For a calculation of an optimized reference sample Z_{ref} all elements of the common sample Z were sorted and divided into m groups with equal length (like a histogram generation). Afterwards the medians of every group were used as the standardized reference sample. By means of obtained equations from the polynomial approximation for the median and quartiles for different voltage values (see Sec. IV-B) we were able to move and spread our new reference sample according to every measured voltage (v). By this way we overcome the discrete distributions for u distance/voltage values.

$$X_v = Z_{Ref} \cdot (f_{.75}(v) - f_{.25}(v)) + f_{.50}(v)$$

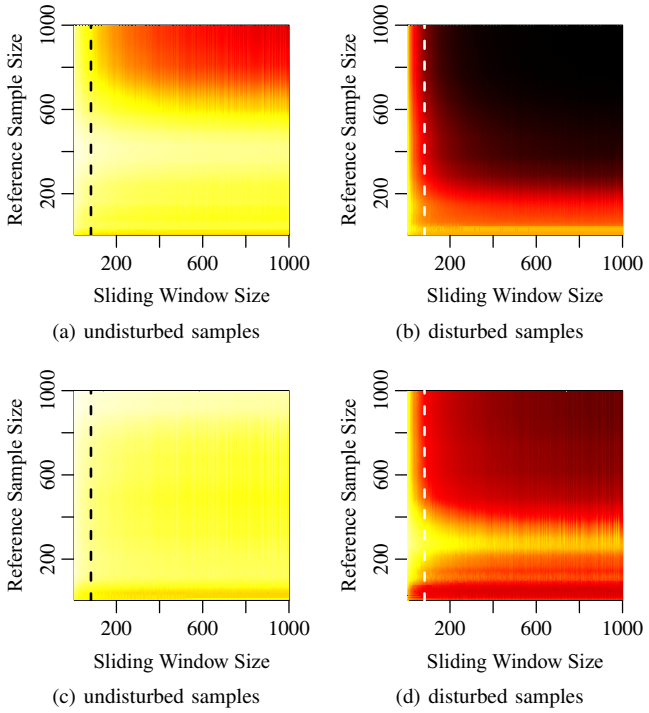
The idea of this approach was to get somehow closer to one-sample tests, because our cleaned reference sample should highly correspond to the unknown reference probability distribution and should therefore produce better results.

B. Sliding Window Size

In a first examination of the four previously mentioned non-parametric tests we had to make a decision about the size of the sliding window (number of the last n measured values). Hence, for every test a varying sliding window ($5 \leq n \leq 1000$) of measurements was compared offline with different sizes of reference samples ($5 \leq m \leq 1000$). Every non-parametric test was performed 500 times with every size configuration, once with clean and once with disturbed measurements, and for different distances between sensors and obstacle, all p-values were recorded.

Fig. 6 shows an example of the average p-values of disturbed and undisturbed measurements calculated with the Ansari-Bradley test and the Kolmogorov-Smirnov test for measurements with a distance of 60 cm between sensor and obstacle. The axis of ordinates defines the granularity of the reference sample whereas the abscissa determines the size of the sliding window. Bright areas mark a probability of an acceptance of the null hypothesis, while dark regions mean a rejection of the assumption.

It is easy to see that these tests show a different reliability in classification for different size configurations. According to the graphical analysis we decided to use a threshold value of .33 for the



(a) undisturbed samples (b) disturbed samples

(c) undisturbed samples (d) disturbed samples

Fig. 6. The Heatmaps show the averaged p-values at different size configurations (bright areas: close to 1., dark areas: p-values lower than .33) for the Ansari-Bradley test: (a) and (b); Kolmogorov-Smirnov test: (c) and (d)

classification that means a p-value lower than this threshold indicates a disturbance and a higher one indicates an undisturbed measurement. In general it is more common to use a threshold like .05 or even lower to reject the null hypothesis, but with the decision to use .33 as threshold we could even decrease the size of the sliding window to 75 values for all tests. For all tests this value was sufficient enough to classify disturbed from undisturbed measurements with an appropriate reference sample size. A bigger window and a lower threshold may increase the quality of classification but would also decrease the speed. For a sliding window size of 75 values in our system setup, disturbed measurements should be identified within a maximum of 3 s (one measurement per 40 ms).

C. Determining Reference Sample Size using Quality Criteria

During this evaluation step we tried to determine the optimum reference sample size for every test and also tried to compare the quality of classification between disturbed and undisturbed measurements. Therefore we used three common measures for the performance of classification [18]. In the following tp denotes the amount of true positives, fn the amount of false negatives and fp means for the false positives.

1) *Precision* can be seen as a measure of exactness or fidelity,

$$Precision = \frac{tp}{tp + fn}$$

2) whereas *Recall* is a measure of completeness

$$Recall = \frac{tp}{tp + fp}$$

3) both measures can be combined with a harmonic mean, which is called the *F-Score*

$$F\text{-}Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}$$

For the classification task we randomly generated a test sequence out of disturbed and undisturbed measurements, those samples were combined alternatingly, whereas the length of every subsequence amounts to 75 measurement values (big enough to completely fill the sliding window). In total the test sequence consisted of 5000 randomly selected samples. Every non-parametric test was applied to this generated sequence with an increasing reference sample size, whereby a calculated p-value lower than or equal to .33 indicated a disturbed measurement and a higher p-value indicated no disturbance. The number of correct and incorrect classifications was recorded, so that the measures described above could be computed. Regarding the measured *F-Scores* for every test and the reference sample size in Fig. 7 it is obvious that the Fligner's test with a reference sample size of 90 measurement values showed the best classification quality which unfortunately decreases with a growing reference sample size. In contrast to this the *F-Scores* of the Ansari-Bradley test and the Mann-Whitney U test remained constant after a certain value, in case of the Ansari-Bradley test after 140 and in case of the Mann-Whitney U test after 180 reference values. The Kolmogorov-Smirnov test initially produced the slowest but also a continuous increase of the *F-Score*. After a reference sample size of 1600 this trend changed so that no *F-Score* higher than 0.579 was produced. An overview of

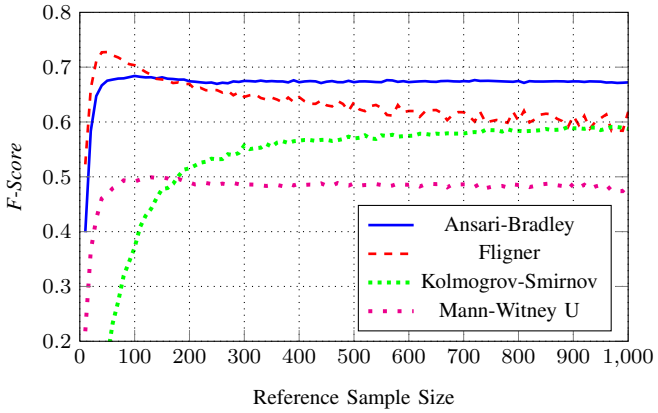


Fig. 7. Measured *F-Scores* of the non-parametric tests at different reference sample sizes

the best results (highest *F-Score*) of every test is given in Tab. II including the reference sample size (Ref.) the averaged speed which is necessary for a test to identify disturbance ($Speed_1$ stands for the averaged amount of values and $Speed_2$ for the averaged time).

As we can see in Tab. II, the Fligner's test was the "overall" winner in our setup. This test did not only reach the highest *F-Score*, *Recall* and *Precision* but also required the smallest amount of reference values, only 90. In addition, this test reacted faster to disturbance than the others, with an average response time of 1.8 s.

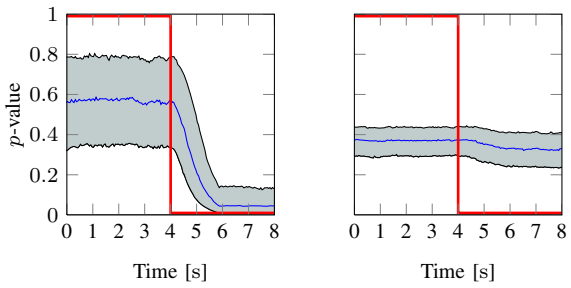
D. Response Behavior

A better view on different response behavior is given in Fig. 8 where the Fligner's test and the Mann-Whitney U test can be

TABLE II
EVALUATION RESULTS FOR THE NON-PARAMETRIC TESTS

Test	Ref.	*Recall*	*Precision*	*F-Score*	Speed_1	Speed_2
Ansari-Bradley	140	.726	.636	.678	58	2.32 s
Fligner's	90	.752	.707	.729	45	1.80 s
Kolmo.-Smirnov	1600	.607	.554	.579	70	2.80 s
Mann-Whitney U	180	.529	.458	.491	*n.d.*	*n.d.*

compared. For this purpose, we run different calculations according to the previous evaluation, but with test sequences consisting of longer successions of disturbed and undisturbed measurement values, actually 125 values (or 5.0 s). The size of the sliding window was 75 and the sizes of the reference sample for every test were the same as the one determined (see Tab. II).



(a) Fligner's test (for measurements with a distance of 10 cm)

(b) Mann-Whitney U test (for measurements with a distance of 60 cm)

Fig. 8. Response (upper, lower quartile and median of p-values) of the best and worst non-parametric test for different distances between sensor and obstacle (see Fig. 2), with starting disturbance at 5 s. The red line marks the behavior of an ideal test, while the dashed line marks the limit p-value of .33.

Fig. 8(a) and 8(b) show the typical response (with upper, lower quartile and median p-value) of the Fligner's test for disturbed and undisturbed measurements at a distance of 10 cm and of the Mann-Whitney U test at 60 cm (compare with Fig. 2). It shows that the Fligner's test was able to clearly distinguish between disturbed and undisturbed measurements, even if there was no real faulty measurement at 10 cm. In contrast to this, the Mann-Whitney U test showed only a slight difference between the calculated p-values for disturbed and undisturbed measurement values, unfortunately this difference seems to be too small for classification, despite of the huge measurement fault at 60 cm. In contrast the Fligner's test was able to detect faulty measurement at 60 cm with a high confidence.

VI. RESULTS

In this paper we presented a general concept of fault detection for external disturbances. We used an exemplary scenario with extraneous light disturbance for a prototypical implementation and identification of an appropriate statistical test. According to the enumeration in Sec. V the Fligner's test showed the best response behavior and detection quality according to an optimized parameter set. The generality of this approach can be adapted to other sources of interference too. Moreover, due to usage of non-parametric statistical tests the presented method is not bound to a specific probability distribution, the only thing that is required is a sample from an undisturbed measurement. Another benefit of the method we presented is that it is possible to detect disturbed measurements even under a slow and continuous increase of external disturbances.

Currently we are working on the extension of the presented fault detection method in the following directions:

- A more detailed classification scheme that differs between typical noise behavior related to smoke, fire, fog, temperature, drops on the lens, etc. and offers an extend perception of the environment. In this way a simple distance sensor becomes a multi-modal sensor.
- Our implementation has problems to cope fast movements of the obstacle. The integration of a mathematical model and a tracking algorithm should help to handle fast moving objects.
- For a "real" smart sensor we want to implement our methods to an 8-bit micro-controller system.
- For a detection of processing and communication faults we have to design a concept that integrates multi-sensor measurements. For this purpose, model based detection methods should be used in connection with classification and decision units.

ACKNOWLEDGEMENT

This work is partly funded by the German Ministry of Education and Science (BMBF) within the project "Virtual and Augmented Reality for Highly Safety and Reliable Embedded Systems" ((ViERforES) - no. 01IM08003C).

REFERENCES

[1] R. Breckenridge and S. Katzberg, "Smart sensors for the 80's- The status of smart sensors," in *Sensor Systems for the 80's Conference, Colorado Springs, Colo*, 1980, pp. 41–46.
[2] S. Zug and J. Kaiser, "An approach towards smart fault-tolerant sensors," in *Procedings of the International Workshop on Robotics and Sensors Environments (ROSE 2009)*, Lecco, Italy, November 2009.
[3] Wenglor Sensoric , *Data sheet YT87MGV80*, 2003. [Online]. Available: http://www.q-tech.hu/pdf/Wenglor/Photoelectric20sensors/YT87MGV80.pdf
[4] V. Nelson, "Fault-tolerant computing: Fundamental concepts," *Computer*, vol. 23, no. 7, pp. 19–25, 1990.
[5] B. Hardekopf, K. Kwiat, and S. Upadhyaya, "Secure and fault-tolerant voting in distributed systems," in *Aerospace Conference, 2001, IEEE Proceedings.*, vol. 3, 2001.
[6] R. Isermann, "Model-based fault-detection and diagnosis–status and applications," *Annual Reviews in control*, vol. 29, no. 1, pp. 71–85, 2005.
[7] P. Frank and B. Köppen-Seliger, "Fuzzy logic and neural network applications to fault diagnosis," *International Journal of Approximate Reasoning*, vol. 16, no. 1, pp. 67–88, 1997.
[8] M. Basseville and I. Nikiforov, *Detection of abrupt changes: theory and application*. Citeseer, 1993.
[9] R. Isermann, *Fault-diagnosis systems*. Springer, 2006.
[10] M. Niedzwiecki and K. Cisowski, "Adaptive scheme for elimination of broadband noise and impulsivedisturbances from AR and ARMA signals," *IEEE Transactions on Signal Processing*, vol. 44, no. 3, pp. 528–537, 1996.
[11] Z. Peng and F. Chu, "Application of the wavelet transform in machine condition monitoring and fault diagnostics: a review with bibliography," *Mechanical Systems and Signal Processing*, vol. 18, no. 2, pp. 199–221, 2004.
[12] D. Sheskin, *Handbook of parametric and nonparametric statistical procedures*, 3rd ed. Chapman & Hall/CRC, Aug. 2003.
[13] M. J. Crawley, *The R Book*, 1st ed. Wiley & Sons, Apr. 2007.
[14] Sharp Cooperation, *GP2Y0A02YK Data Sheet*, 2007. [Online]. Available: http://sharp-world.com/products/device/lineup/data/pdf/datasheet/gp2y0a21yk_e.pdf
[15] National Instruments, *NI DAQCard-6036E Data Sheet*, 2005. [Online]. Available: http://sine.ni.com/nips/cds/print/p/lang/de/nid/11914
[16] SciPy Project Page, 2009. [Online]. Available: http://www.scipy.org
[17] R. Gentleman, V. Carey, W. Huber, S. Dudoit, and R. Irizarry, *Bioinformatics and computational biology solutions using R and Bioconductor*. Springer Verlag, Sep. 2005.
[18] C. Goutte and E. Gaussier, "A probabilistic interpretation of precision, recall and f-score, with implication for evaluation," in *Proceedings of the 27th European Conference on Information Retrieval*, 2005, pp. 345–359.