**Proceedings of the 17th World Congress**
**The International Federation of Automatic Control**
**Seoul, Korea, July 6-11, 2008**

# Natural Environment Modeling & Fault-Diagnosis for Automated Agricultural Vehicle

**M. R. Blas** * **M. Blanke** *

* *Technical University of Denmark, Department of Electrical Engineering, Elektrovej Build. 326, DK (e-mail: mrb@elektro.dtu.dk,mb@elektro.dtu.dk)*

**Abstract:** This paper presents results for an automatic navigation system for agricultural vehicles. The system uses stereo-vision, inertial sensors and GPS. Special emphasis has been placed on modeling the natural environment in conjunction with a fault-tolerant navigation system. The results are exemplified by an agricultural vehicle following cut grass (swath). It is demonstrated how faults in the system can be detected and diagnosed using state of the art techniques from fault-tolerant literature. Results in performing fault-diagnosis and fault accomodation are presented using real data. Copyright ©2008 IFAC

## 1. INTRODUCTION

Agricultural machinery is increasingly getting automated. For example, agricultural vehicles have seen a revolution in automation by adoption of GPS for automatic steering. A number of technical limiting factors of these GPS systems however do exist. Some of these are faults inherent to GPS receivers, such as those induced by satellite occlusions and multipath errors. Others include that the GPS systems require a detailed navigation plan as they themselves cannot "see" features in the field that want to be followed. It is thus of interest to find a solution to these limitations.

Earlier research results can be split into 3 broad categories: work in fault-tolerance, stereo-vision, and position estimation for navigation in agriculture. Stereo-vision is an active topic in agriculture. It has been shown that stereo-vision can be used for navigation by finding the relative position of a vehicle to a variety of agricultural structures: Rovira-Más et al. (2007), Kise et al. (2005). Relative position estimation has been fused in Rovira-Más and Han (2006). Visual odometry has been fused in Andersen et al. (2007). Sensor faults in these articles are typically treated in simple fashion by gating on, for example, the innovation in a kalman filter. Literature however exists to treat fault detection systematically, Blanke et al. (1997), Blanke et al. (2006), by using systematic fault-tolerant design tools. They have been demonstrated to work in practice on for example ships: Blanke (2006). How the systematic methods as presented in Blanke et al. (2006) can be applied to navigation in agriculture and especially with vision-based navigation sensors has not previously been demonstrated.

This paper will deal with a specific field operation that involves the agricultural vehicle to follow cut grass (swath) in order to pick it up with a baler (see fig. 1). The system to be analyzed is equipped with stereo-vision, a single antenna GPS and an IMU, in one configuration. GPS positions of the vehicle that formed the swath are known. The combination of stereo-vision and GPS allows the system both to "see" the swath but also navigate based on the given map. This creates system redundancy that is essential for achieving fault-tolerance. A visual odometry algorithm on the stereo-camera allows for the relative positionment of the vehicle without GPS or IMU. The GPS receiver used was a high-end EGNOS receiver and the IMU was a tactical grade (low accuracy) MEMS based unit.

Fig. 1. This is an example of a swath that the navigation system should follow. The vision system locks on and tracks the swath most central in the image.

The two main ideas presented here is first a behavior model for representing the natural environment, namely the swath. Secondly, it is shown how parts of this model (the swath location) can be used in conjunction with sensor inputs to create a fault tolerant sensor fusion system. The fault diagnosis is illustrated using real data. Combining the sensor information optimally for state estimation beyond estimating faults is not delved into.

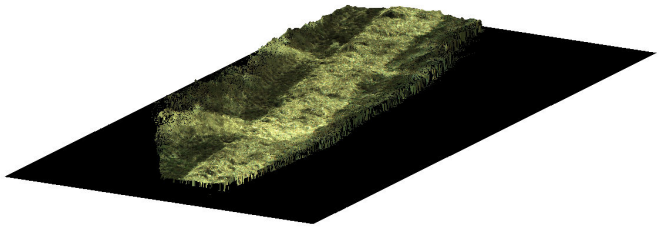

Fig. 2. Using position information and 3D data from the stereo-camera the entire field (if necessary) can be reconstructed in 3D. This is a 15 m section of a swath.

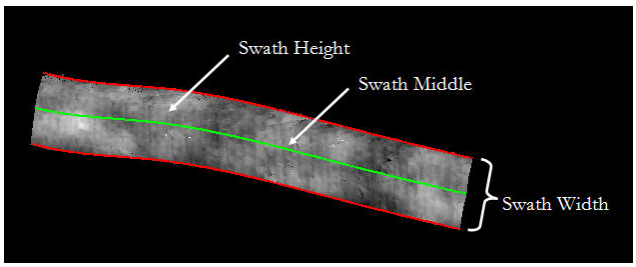


Fig. 3. The middle of the swath can be represented by a cubic spline with a number of knots. This is accompanied by the width of the swath at each location along the spline as well as the height of the material within the bounds of the swath. The height of the swath is shown as different intensities. The illustration is from a top-down view.

## 2. SWATH MODEL

A model of the swath requires extracting the salient features of the environment required for the field operation and storing them in the model representation. The salient features are the location of the swath and the distribution of the swath material across the swath. The location of the swath is modeled as a cubic spline with the number of knots being fixed for a certain length of swath. The spline is positioned along the middle of the swath. The distribution of the swath material is first modeled by defining the swaths width at any given point along the spline. Secondly, points inside the 2D volume enclosed by the swath width are assigned a value pertaining to the height of the swath at this point relative to the ground. The swath height is represented by a grid map of resolution 2 cm for each grid point. The model is illustrated in fig. 3. The concept behind the swath location, width, and height is now explained in more detail.

### *2.1 Swath Location*

The swath location is defined as being in a 2D coordinate system on the ground plane. A function $f$ represents the lines down the middle of the swaths. Given coordinate pairs $(x, y)$ then $f$ is:

$$y = f(x) \tag{1}$$

The model of the swath location is then $s(x)$ with $s \in \mathcal{S}_3(\boldsymbol{k}_{0:n})$, where $\boldsymbol{k}_{0:n}$ are the spline knots and the spline coefficients and $\mathcal{S}_3$ is the cubic spline domain. Then the model is equal to the swath location plus the approximation error $\epsilon_a$ of fitting a spline to $f$:

$$s(x) = f(x) + \epsilon_a \tag{2}$$

Based upon the concept of having a controller that allows the vehicle to follow the swath location, the x-track error $\varepsilon_x$ (signed shortest distance from the control point to the spline) can be found as a function of the tractor position and the spline. Defining the spline $s_b$ in body coordinates a function $\mathcal{E}$ can be set to find the x-track error:

$$\varepsilon_x = \mathcal{E}(s_b) \tag{3}$$

### *2.2 Swath Width*

For each point on $s$ the swath width is defined by the two points on either side of $s$ that are orthogonal to $s$ at these points and which lie a distance $d_w$ away from $s$ at each point. The function $g$ defines $d_w$ as a function of $x$ and the model $s$. The tangent of $s$ at a specific point is given by differentiation of $s$. This allows for the swath width to change along the length of the swath.

$$d_w = g(x, s) \tag{4}$$

### *2.3 Swath Height*

The swath height is the mean swath height inside each grid square in the grid map bounded by the swath sides as defined by the swath width and location. Given the grid map coordinates $(x, y)$ the swath height $z_m$ at this point is defined by the function $h$:

$$z_m = h(x, y) \tag{5}$$

### *2.4 Swath Sets*

The above model is only for a single swath. Each swath in the field is then described by a set of the above functions with $\mathcal{M}$ being the set of such swaths and $n$ being the number of swaths:

$$\mathcal{M} = \{s_i, g_i, h_i\}, i = 1...n$$

## 3. BEHAVIOR MODELS

The following sections outline how the behaviour-based model of the swath can combined with models of GPS, stereo-vision sensor and odometry data. The purpose is to arrive at a set of constraints that can be used for analysis of system structure and subsequent generation of residuals for fault diagnosis. This idea was brought into the field of fault diagnosis by Staroswiecki and Declerck (1989) and later expanded, see Staroswiecki and Comet-Varga (2001) and Blanke et al. (2006). The advantage of this approach over classical methods, Tzafestas and Watanabe (1990), include the ability to use a formulation of behaviors at high level of abstraction.

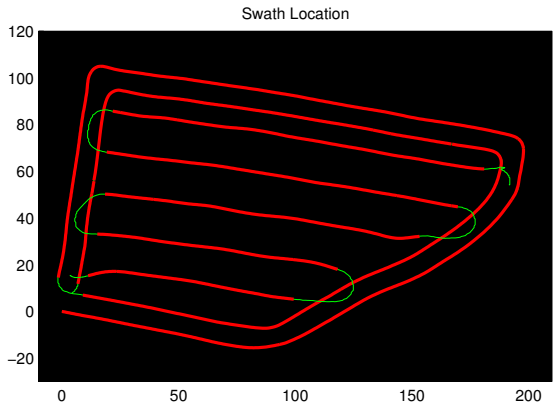


Fig. 4. The location of the swath as recorded by the GPS mounted on the vehicle that formed the swath. It is assumed the start and end points of the GPS trajectory have been recorded. The swath location is shown in red and has been fitted with a spline. Green indicates the trajectory followed by the vehicle where it was not forming swath - raw GPS readings. All units are in meters in the NED coordinate system.

### *3.1 GPS*

The GPS positions of the vehicle that formed the swath is fitted with the spline (see fig. 4). The distribution of the swath cannot be measured by the GPS receiver. Knowledge of the machine settings used to construct the swath are assumed known giving an approximate height $h_{est}$ and width $d_{w,est}$ of the swath. These settings can be used as estimates for the swath model fitting. Given the measured GPS positions $\hat{\mathbf{p}}_2^n$ of the vehicle and the attitudes of the vehicle $\hat{\mathbf{\Theta}}_2$ the position of swath formation behind the vehicle can be calculated. The path formed by these positions is then fitted with the spline model using the function $k$ to provide an estimate of the swath location. The information in the GPS map is then, in an abstract formulation:

$$s_g = k(\hat{\mathbf{p}}^n, \hat{\mathbf{\Theta}}_2) \tag{6}$$

### *3.2 Stereo-Camera*

A stereo algorithm is used to find the correspondence between features in the left and right image sensor $(i_l, i_r)$. The position of the features relative to the stereo-camera can then be inferred in 3D. Modern vision algorithms then exist to register 3D models with the 3D point cloud provided by the stereo-camera: Goshtasby (2005). An algorithm has been constructed that allows such registration between the swath model and the 3D points. The stereo-algorithm and registration will be denoted by the function $a_{reg}$. Thus, given the two images a measurement of the swath location $s_c$, width $g_c$, and height $h_c$ can be computed for the part of the swath in the image.

$$\begin{bmatrix} s_c \\ g_c \\ h_c \end{bmatrix} = a_{reg}(i_l, i_r) \tag{7}$$

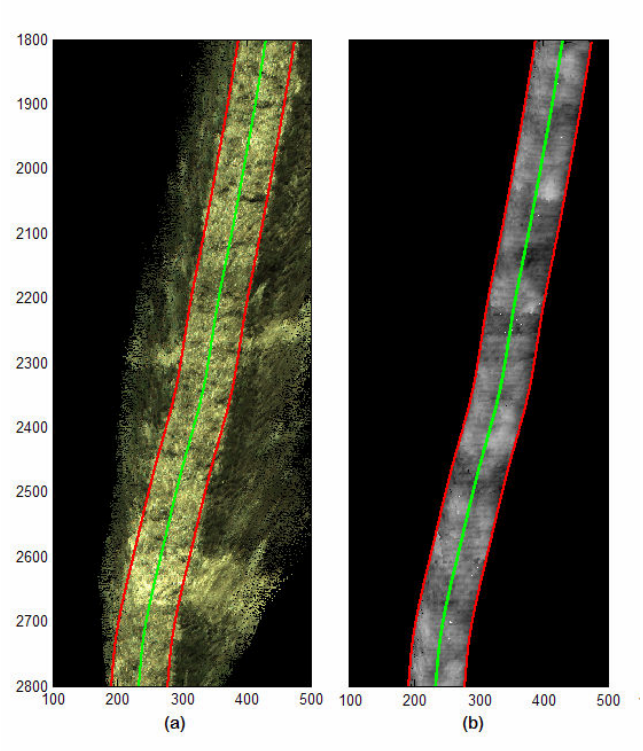


Fig. 5. (a) RGB topdown-view of the swath with swath location and swath width superimposed on the swath. (b) The swath model showing both swath location, width and height by image feature extraction.

These measurements are stored in a map representation for an individual swath. $s_m$ is the spline formed by combining previous measurements.

## 4. STRUCTURAL MODEL

The structural model describes the behavior of variables in the normal, fault-free system using a behavior based approach. A violation of a behavior indicates a fault in the system. Given the current setup a fault could for example be in a sensor, an algorithm, and/or an assumption about the environment. The structural approach treats faults unambiguously.

The constraints are composed of a number of measurement $(m)$, differential $(d)$, and system constraints $(c)$. The variables in the constraints are likewise composed of two groups: the subset of known variables $\mathcal{K}$ and the subset of unknown variables $\mathcal{X}$. The constraints are listed in Eq. 9.

$$\begin{aligned} K &= \{\mathbf{v}^b, \mathbf{a}^b, \mathbf{p}_1^n, s_g, s_c, s_m, \mathbf{R}_b^n(\Theta, \lambda)\} \\ X &= \{\mathbf{p}^n, \dot{\mathbf{p}}^n, s^b, s, \varepsilon_x\} \end{aligned} \tag{8}$$

Table 1. Incidence Matrix.

| | known | | | | | | unknown | | |
|---|---|---|---|---|---|---|---|---|---|
| \ | $s_g$ | $s_c$ | $s_m$ | $\mathbf{v}^b$ | $\mathbf{a}^b$ | $p_1$ | $\mathbf{p}$ | $\dot{\mathbf{p}}$ | $\varepsilon_x$ |
| $m_1$ | | | | 1 | | | | 1 | |
| $m_2$ | | | | | 1 | | | 1 | |
| $m_3$ | | | | | | 1 | 1 | | |
| $m_4$ | 1 | | | | | | 1 | | 1 |
| $m_5$ | | | 1 | | | | | | 1 |
| $m_6$ | | 1 | | | | | | | 1 |

Table 2. Dependability Matrix.

| | constraints | | | | | | |
|---|---|---|---|---|---|---|---|
| \ | $d_1$ | $m_1$ | $m_2$ | $m_3$ | $m_4$ | $m_5$ | $m_6$ |
| $p_1$ | 1 | 1 | | 1 | | | |
| $p_2$ | | 1 | 1 | | | | |
| $p_3$ | | | | 1 | 1 | 1 | |
| $p_4$ | | | | | | 1 | 1 |

$$
\begin{aligned}
c_1 &: s^b = \mathbf{R}_b^n(\Theta,\lambda)s + \mathbf{p}^n \\
c_2 &: \varepsilon_x = \mathcal{E}(s^b) \\
d_1 &: \dot{\mathbf{p}}^n = \frac{d}{dt}\mathbf{p}^n \\
m_1 &: \mathbf{v}^b = \mathbf{R}_b^n(\Theta,\lambda)\dot{\mathbf{p}}^n \\
m_2 &: \mathbf{a}^b = \frac{d}{dt}\mathbf{R}_b^n(\Theta,\lambda)\dot{\mathbf{p}}^n \\
m_3 &: \mathbf{p}_1^n = \mathbf{p}^n \\
m_4 &: s_g = s \\
m_5 &: s_c = s^b \\
m_6 &: s_m = s^b
\end{aligned}
\tag{9}
$$

Following the notation in Fossen (2002): $\mathbf{v}^b$ is the tractor's velocity vector over ground seen in body coordinates as measured by visual odometry; $\mathbf{a}^b$ is the acceleration vector given by the IMU; $\mathbf{p}^n$ the position in (North, East) coordinates with $\mathbf{p}_1^n$ being the position measurement from the GPS; $\mathbf{R}_b^n$ is the rotation matrix from body to navigation frame, which is a function of $\Theta$, the attitude vector (Euler angles roll, pitch and yaw) and of $\lambda$, the latitude;. $s^b$ is the spline body coordinates. In this analysis, the $\mathbf{R}_b^n$ matrix is assumed to be known.

## 5. STRUCTURAL ANALYSIS

As described in Blanke (2006) a structural analysis is then performed on the structural model. The constraint $d_1$ is a differential constraint and as such cannot fail. To simplify the structural analysis the constraints $\{c_1, c_2\}$ are pulled into the measurement constraints $\{m_4, m_5, m_6\}$ and faults in them are treated as subsystem faults in the respective measurement constraints instead. The incidence matrix can be seen in table 1. A number of parity relations are then found as shown in Table 2.

Based on the dependability matrix in Table 2, the parity relations are derived in analytical form and used as the basis for the residuals (Eq. 10). As all the columns of Table 2 are different, it follows that all faults should be structurally detectable and isolable as the faults will have a unique signature in these residuals.

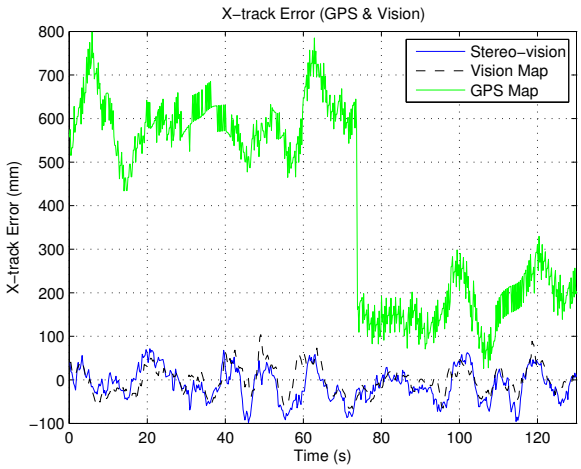


Fig. 6. The vehicle was driven manually over a swath. The driver centered the vehicle over the middle of the swath and drove for 2 min while maintaining this centered position. The x-track errors from the subsystems were recorded.

$$
\begin{aligned}
r_1 &: \mathbf{v}^b - \mathbf{R}_b^n(\Theta,\lambda)\frac{d}{dt}\mathbf{p}_1^n = 0 \\
r_2 &: \frac{d}{dt}\mathbf{v}^b - \mathbf{a}^b = 0 \\
r_3 &: \mathcal{E}(s_c) - \mathcal{E}(\mathbf{R}_b^n(\Theta,\lambda)s_g + \mathbf{p}_1^n) = 0 \\
r_4 &: \mathcal{E}(s_c) - \mathcal{E}(s_m) = 0
\end{aligned}
\tag{10}
$$

## 6. FIELD TESTS

The properties of residuals were investigated based on recorded data. The data stems from the test field run illustrated in fig. 4. The position of the swath was first logged by following the middle of the swath manually - emulating the vehicle forming the swath. This was repeated for a second pass emulating the vehicle that should pick up the swath. This provides some form of limited ground truth. The position error of the driver is bounded between the runs as he constantly steers relative to the swath. Experience with driving with balers puts the error associated with not driving exactly over the center of the swath to under ±0.2 m as this is required to pick up the swath successfully. In the data examined the GPS has a false offset in the second pass relative to the first pass of approximately 0.6 m for the first approx. 70 s before it corrects its position estimate to bring it to about 0.15 m of the swath location. This second offset is acceptable for normal operation. Field tests enabled calculation of residuals $r_1, r_3$ and $r_4$ as an instrumentation issue prevented data reception from the IMU. The field test is hence representing a case of one permanent failure and an additional fault occurring.

### *6.1 Detailed Design of Residual Generators*

The parity relations Eq. 10 are now further scrutinized as a basis for change detector and hypothesis evaluation design. A common assumption for readily available change detection algorithms is that of a Gaussian amplitude distribution. A required property of residuals for average run length calculations is whiteness. Fig. 7 shows the

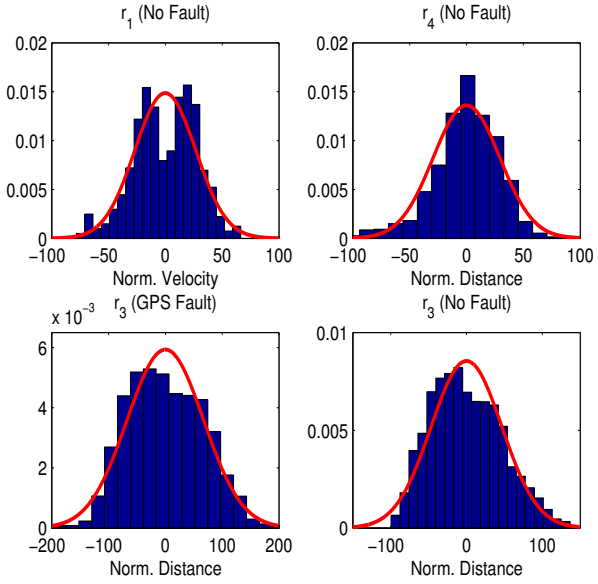


Fig. 7. Normalized histograms for residuals $r_1, r_3, r_4$ in the faultless scenario along with $r_3$ in the case of the GPS fault. Fitted gaussian distributions are shown on top of the histograms.

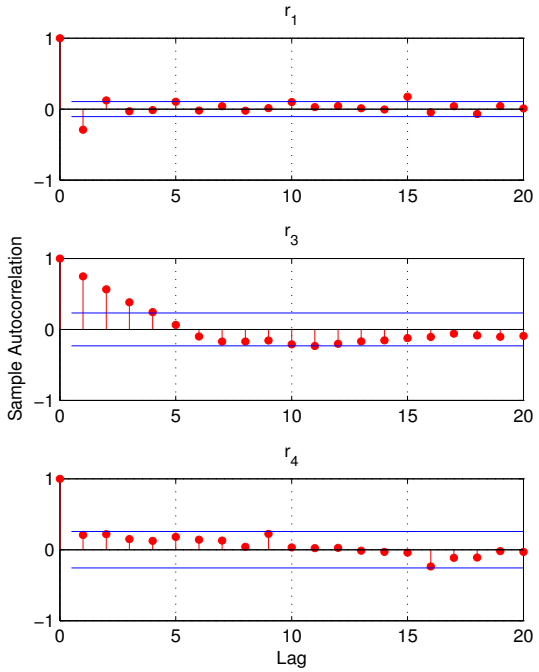


Fig. 8. Autocorrelation function for residuals $r_1, r_3, r_4$ in the faultless case. The blue lines indicate the bounds of the 95% confidence intervals.

histograms of residuals. The figure also shows Gaussian distributions with mean and variance as observed. Residual 1 suffers from a deadband in the calculation of velocity from the stereo images, hence, the distribution is not Gaussian. The cause of the deadband must be investigated further. Residual 3 appears to follow a shifted Rayleigh distribution.

Statistical change detection will ideally be based on a log-likelihood ratio test

$$s_i = \ln(p_{\theta_1}(r_i) - \ln(p_{\theta_0}(r_i)) \tag{11}$$

where the probability densities from the observed distributions should be used, $p_{\theta_1}$ for the case of a fault, $p_{\theta_0}$ for the normal case, respectively. With the shifted Rayleigh shape distribution of $r_3$ in the no-GPS-fault case, a change detection of CUSUM or GLR type is straight forward to compute, Basseville and Nikiforov (1993). The Rayleigh distribution gives, however, some computational burden over the tests when Gaussian distributions are assumed. The testing was therefore conducted using a Gaussian assumption. In the above analysis, it is also noted that a GPS fault is strongly detectable in residual $r_3$ while only weakly detectable in $r_1$ due to the position differentiation in the parity relation of $r_1$. A low-pass filter is hence applied on $r_1$.

With respect to testing whiteness of the residuals, Fig. 8 shows the autocorrelation functions for the three available residuals. Residual $r_3$ is seen to comprise some correlation due to filtering within the algorithm that determines the spline approximation to the swath. Whiteness is particularly important to reach design conclusions about average run length.

### *6.2 Change Detection and Hypothesis Evaluation*

Change detection on the residual vector could be made using a vector-based approach, where a known signature $\rho(\tau) = [\rho_1(\tau), \rho_3(\tau), \rho_4(\tau)]^T$ is sought for in $\mathbf{r}(t) = [r_1(t), r_3(t), r_4(t)]^T$. The computational burden is, however, larger than by applying a simple CUSUM test for change in mean on $r_3$ and $r_4$ and make a threshold test on $r_1$. With adequate logics used for hypothesis testing, this allows for isolation of faults. The result for $r_3$ is shown in fig. 10. The CUSUM test is given a mean value to test for so that an erroneous x-track error of up to 0.2 m is permissible. From the system is started it takes 1 s to isolate a fault. The CUSUM test for residual $r_4$ stays 0 for the data indicating that $s_c$ and $s_m$ agree, and that any variation between them is due to noise.

Isolation of the fault to the GPS subsystem (GPS sensor and GPS Map) can, as minimum be achieved by considering the residual vector $[r_3, r_4]^T$. Stronger isolation is achieved by considering the full residual vector $[r_1, r_2, r_3, r_4]^T$ where the output of the GPS subsystem is also compared to the visual odometry which in turn is compared to the IMU output. The GPS jump is clearly seen as a spike in residual $r_1$ as shown in Fig. 11.

## 7. FAULT HANDLING

There are a number of approaches to fault accomodation. In the present case, a GPS fault can be isolated, and the magnitude of the fault can be estimated. It is then a straight forward task to reconfigure the navigation controller either to avoid using the faulty sensor or to attempt to compensate for the fault by compensating the

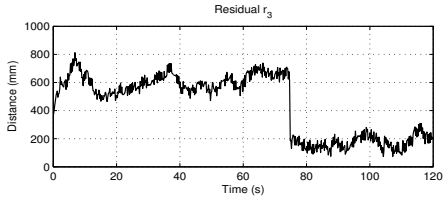


Fig. 9. Residual $r_3$.

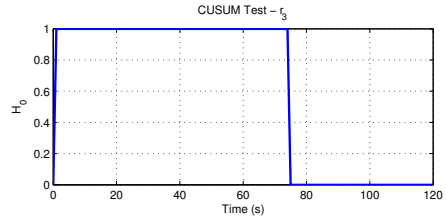


Fig. 10. CUSUM test for residual $r_3$. $\mathcal{H}_0 = 0$ indicates no fault.

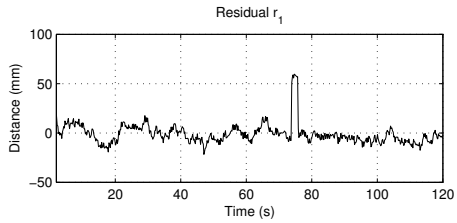


Fig. 11. Low-pass filtered residual $r_1$.

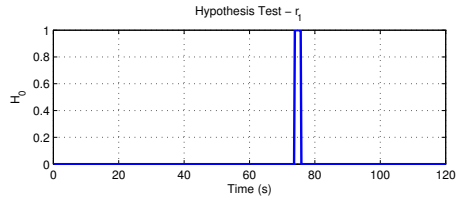


Fig. 12. Hypothesis test by thresholding for residual $r_1$.

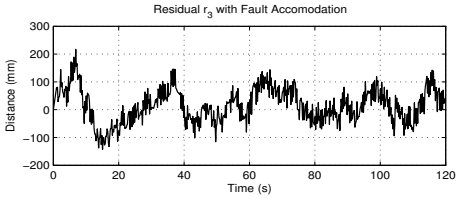


Fig. 13. By estimating the magnitude of the fault in the GPS sensor it was possible to do fault accomodation.

sensor readings. Given that the residuals have detected the fault in the GPS x-track signal it is indeed possible to handle the fault automatically. The approach adopted here was to do this is by estimating the GPS fault magnitude using a Kalman filter and the vision x-track signal. The estimated mean difference is used as the correction estimate. Threshold detection on $r_1$is used to prevent invalid signals from being passed to the controller. The response of the thus corrected position estimate is shown in Fig. 13.

It can be seen that discrepencies in the GPS and vision x-track signals quickly converge (the two signals in the example are roughly within $\pm 10$ cm of each other). This approach only works for correcting faults which are stationary errors to the input signal. It is stipulated that most other GPS faults will be caught by the visual odometry or IMU.

## 8. CONCLUSION

It has been demonstrated both using principles from fault diagnosis and from fault-tolerant sensor fusion theory, as well as experimentally, that an agricultural vehicle equipped with GPS, IMU and stereo-vision can be made fault-tolerant to sensor faults. The results presented here were done offline and work will be pursued to demonstrate them online. Real-time implementations are available for the image processing algorithms. The computational requirements for the fault-tolerant sensor fusion framework is insignificant compared to the image processing. The system should thus have a very good chance of working online.

## REFERENCES

H. J. Andersen, T. Bak, and M. Christensen. Fusion of gps and visual motion estimates for robust outdoor open field localization. *Second International Conference on Computer Vision Theory and Applications, VISAPP. INSTICC.*, pages 413–418, 2007.

M. Basseville and I. Nikiforov. Detection of abrupt changes: Theory and applications. *Prentice-Hall*, 1993.

M. Blanke. Fault-tolerant sensor fusion for marine navigation. *Proc. 7th IFAC Conf. on Manoeuvring and Control of Marine Craft*, Elsevier IFAC, sep 2006.

M. Blanke, R. Izadi-Zamanabadi, S. A. Bøgh, and C. P. Lunau. Fault-tolerant control systems - a holistic view. *Control Engineering Practice*, 5(5):693–702, 1997.

M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki. *Diagnosis and Fault-Tolerant Control 2nd Edition.* Springer, 2006.

T. I. Fossen. *Marine Control Systems.* Marine Cybernetics, 2002.

A. A. Goshtasby. *2-D and 3-D Image Registration: for Medical, Remote Sensing, and Industrial Applications.* Wiley, 2005.

M. Kise, Q. Zhang, and F. Rovira-Más. A stereovision-based crop row detection method for tractor-automated guidance. *Biosystems Engineering*, pages 357–367, 2005.

F. Rovira-Más and S. Han. Kalman filter for sensor fusion of gps and machine vision. Technical report, 2006 ASABE Meeting Presentation - Paper Number 063034, 2006.

F. Rovira-Más, S. Han, J. Wei, and J. F. Reid. Autonomous guidance of a corn harvester using stereo vision. *Agricultural Engineering International: the CIGR Ejournal.*, 9, 2007.

M. Staroswiecki and G. Comet-Varga. Analytical redundancy relations for fault detection and isolation in algebraic dynamic systems. *Automatica*, 37(5):687–699, 2001.

M. Staroswiecki and P. Declerck. Analytical redundancy in nonlinear interconnected systems by means of structural analysis. In *Proc. IFAC AIPAC'89 Symposium.*, volume 2, pages 23–27. Elsevier - IFAC, 1989.

S. Tzafestas and K. Watanabe. Modern approaches to system/sensor fault detection and diagnosis. *Journal A.*, 31(4):42–57, 1990.