UDC 629.5.05:681.783.3

*Zdravko EŠKINJA*[1]
*Đula NAĐ*[2]
*Vladimir DJAPIĆ*[3]

# Multiple Autonomous Systems in Underwater Mine Countermeasures Mission Using Information Fusion as Navigation Aid

**Authors' addresses (Adrese autora):**
[1] *Brodarski Institute, Zagreb*
[2] *Faculty of Electrical Engineering and Computing, University of Zagreb*
[3] *NATO Center for Maritime Experimentation (CMRE)*
**Received (Primljeno):**2013-04-15
**Accepted (Prihvaćeno):** 2013-05-20
**Open for discussion (Otvoreno za raspravu):** 2014-06-30

Review paper

Autonomous bottom mine neutralization systems have a challenging task of mine reacquisition and navigation in the demanding underwater environment. Even after mine reacquisition, the neutralization payload has to be autonomously deployed near the mine, and before any action the verification (classification) of the existence of a mine has to be determined. The mine intervention vehicle can be an expendable (self-destroyed during the mine neutralization) or a vehicle that deploys the neutralization payload and it is retrieved at the end of the mission. Currently the systems developed by the research community are capable of remotely navigating a mine intervention underwater vehicle in the vicinity of the mine by using remote sonar aided navigation from a master vehicle. However, the task of successfully navigating the vehicle that carries the neutralization payload near the bottom and around the mine remains a challenge due to sea bottom clutter and the target signature interfering with the sonar detection. We seek a solution by introducing navigation via visual processing near the mine location. Using an onboard camera, the relative distance to the mine-like object can be estimated. This will improve the overall vehicle navigation and rate of successful payload delivery close to the mine. The paper presents the current navigation system of the mine intervention underwater vehicle and the newly developed visual processing for relative position estimation.

**Keywords**: *image processing, sensor fusion, sonar aided navigation, stereo vision, underwater photogrammatery, underwater vehicles*

## Korištenje fuzije podataka za ispomoć u navigaciji višestrukog autonomnog sustava u podvodnom protuminskom djelovanju

Pregledni znanstveni rad

Sustavi za autonomnu neutralizaciju podvodnih mina imaju izazovni zadatak pronalaženja mine i navigacije u zahtjevnom podvodnom okolišu. Nakon pronalaženja podvodne mine potrebno je autonomno položiti neutralizator pokraj mine i prije bilo kakve akcije potrebno je utvrditi postojanje (klasifikaciju) mine na lokaciji. Intervencijsko vozilo može biti potrošno (uništeno u procesu neutralizacije) ili vozilo koje se vraća nakon odlaganja neutralizatora. Trenutačno razvijani sustavi u mogućnosti su daljinski navesti neutralizacijsko podvodno vozilo u blizinu mine koristeći potpomognutu navigaciju putem sonara na glavnom vozilu. Ipak, zadatak uspješnog navođenja vozila blizu dna i oko mine ostaje izazov zbog toga što prirodni objekti na morskom dnu i sama mina utječu na kakvoću detekcije vozila u sonarskoj slici. Rješenje smo potražili uvođenjem obrade vizualnih podataka u područjima blizu mine. Korištenjem kamere moguće je estimirati relativni položaj u odnosu na minoliki objekt. Estimat će pridonijeti ukupnoj navigaciji vozila i uspješnosti ispostave neutralizatora u

neposrednu blizinu mine. Članak će prezentirati trenutačni navigacijski sustav na podvodnom vozilu i novorazvijeni sustav obrade slike za estimaciju relativne pozicije.

**Ključne riječi**: *fuzija senzora, obrada slike, podvodna fotogrametrija, podvodna vozila, potpomognuta sonarska navigacija, stereo vid*

# 1 Introduction

Underwater mine countermeasure missions require precise manipulation in unstructured ocean environments possibly at depths that are unsafe for human divers. In addition, handling of the neutralization payload near the active mine is itself a dangerous operation. Various types of underwater vehicles are used for ocean exploration. Manned submersibles offer human operators a direct presence in the minefield, but their usage is limited by safety considerations of being in a highly explosive zone. Remotely operated vehicles (ROVs) are an alternative for manned exploration. These vehicles are connected with the surface support vessel over a tether that provides high-bandwidth, two way communication as well as power for extended operations. The main drawbacks of ROVs are the existence of the tether and a need for a skilled pilot.
Un-tethered unmanned vehicles, generally called Autonomous Underwater Vehicles (AUV) enable missions where a tether is not feasible. Tele-operation of AUVs is limited due to the low bandwidth of acoustic communication. Therefore, AUVs have to implement all control and navigation on board, i.e. target detection, obstacle avoidance, robot localization, object mapping and manipulation, etc. To improve the solution cost effectiveness, the Autonomous Naval Mine Countermeasures Program at the Center for Maritime Experimentation (CMRE) proposes the mine intervention vehicle to be used as a part of a system of autonomous vehicles with different functionality [1]:

1. A mapping and searching AUV equipped with advanced navigation and a sidescan or synthetic aperture sonar, SAS (**Pogreška! Izvor reference nije pronađen.**);
2. A surface (or in the future an underwater) vehicle equipped with advanced navigation, multi-beam sonar and an Ultra-Short Baseline Positioning (USBL) acoustic localization system (Figure 2);
3. An expendable underwater vehicle with no positioning sensors for close range mine neutralization (Figure 3).

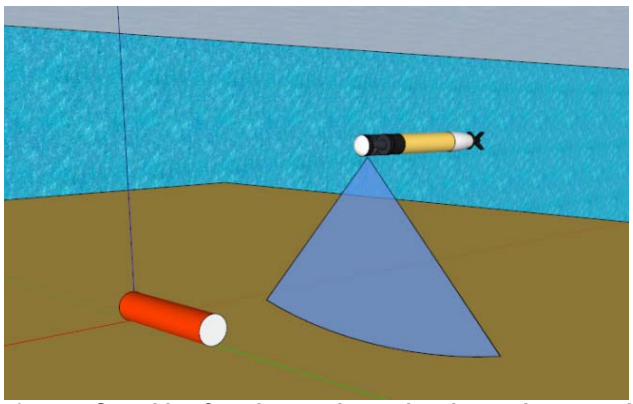

Figure 1 **Searching for mines and mapping the sea bottom using AUV systems**
Slika 1 **Pretraživanje mina i mapiranje morskog dna AUV sustavom**

This paper analyses the self navigation and navigation aiding of the expendable underwater vehicle. The sections that follow describe a collaborative autonomous naval mine countermeasure (ANMCM) system, consisting of a surface vehicle and a small expendable vehicle, offer a typical mission scenario, and define problems occurring in such a system. The navigation system used for remote navigation is derived in the forth section in such a way to

allow sequential data fusion of the remote multibeam sonar measurements and the onboard video target relative localization. The remaining sections will then describe the video processing algorithms derived for target localization and relative navigation.

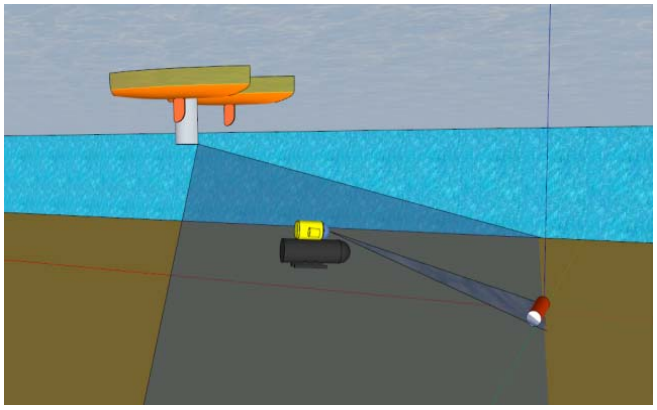

Figure 2 **Mine countermeasures mission - Guiding UUV by ASV sonar imagery**
Slika 2 **Misija u borbi protiv mina – Vođenje UUV-a korištenjem sonarske slike**

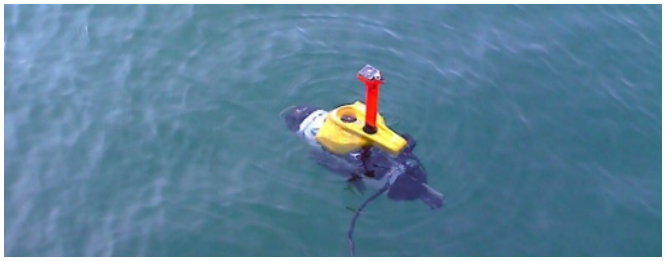

Figure 3 **Unmanned underwater vehicle**
Slika 3 **Bespilotna podvodna ronilica**

## 2 Mission scenario

Let us consider a typical mission scenario where mine detection has to be performed in an underwater region. First, an AUV, equipped with synthetic aperture sonar (SAS) imaging systems is sent to acquire a detailed sea-bottom map of the region. SAS systems can provide imagery of the underwater environment with resolution on the order of a centimeter while achieving an area coverage rate of over a square kilometer per hour [2], which makes it ideal for fast and effective map acquisition. Mine-like object detection is performed on the map and coordinates of the suspicious objects are extracted.

Once coordinates of the objects are available, the autonomous surface vehicle (ASV) is sent for a closer look. The ASV carries a precise navigation system, a forward looking multibeam sonar and one or more small and simple unmanned underwater vehicles (UUVs) [1].

The ASV performs autonomous mine reacquisition of the previously detected target (search AUV with SAS) with the onboard multibeam sonar [3]. Upon detection and verification of the reacquisition the ASV initiates the UUV deploying procedure. First, a new navigation frame $\{T\}$ is created with the reacquired target at coordinates $T_2(0, 0, z)$, see Figure 4a. Then the activated UUV is released into the water (see Figure 3). The activation is performed through a wireless command.

The UUV is commanded to maintain a fixed depth (i.e. 2m) until it is acquired in the multibeam sonar image by the ASV. Once the vehicle is acquired it is switched into the line following mode and follows the line towards the target, see Figure 4**Pogreška! Izvor reference nije pronađen.**b. The navigation is aided by the ASV that sends the sonar detections to the UUV, virtually acting as a remote underwater GPS or USBL. The UUV navigation and control is working in the target centric navigation frame. The line following problem was analyzed in detail in [4].

Once the UUV is close to the mine the target is acquired visually. Depending on the visibility it can be from ten to less than one meter. However, switching to visual updates has the advantage of better refresh rates, and no measurements delays, as opposed to the acoustic navigation.

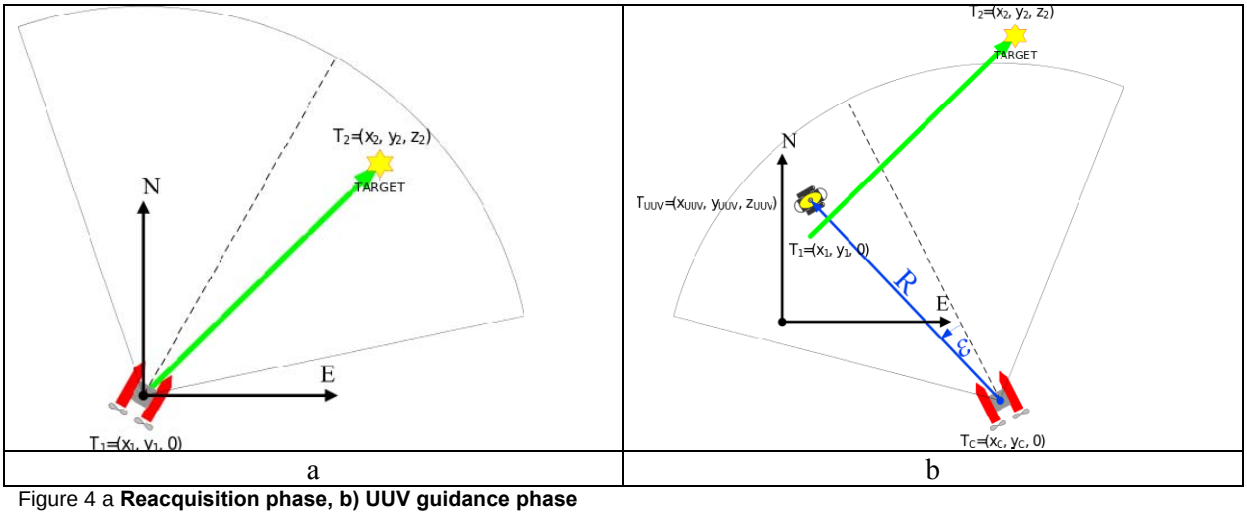


Figure 4 a **Reacquisition phase, b) UUV guidance phase**
Slika 4 a **Faza re-akvizicije, b) UUV faza vođenja**

Current mission scenarios assume the target stationary. Therefore, continuous monitoring of the target is not necessary. The benefit is that the master vehicle can easily maneuver around the intervention UUV depending on surface conditions. In the future an anti-swimmer/diver application can be considered with the same strategy, however, the target position will not be fixed and both tracking of the intervention UUV and the target will be necessary.

## 3 Problem formulation

In the previous section we described the current typical mission scenario. Aspects of deployment, acquisition and sonar tracking have been implemented and tested in sea-trials. More details on this can be found in other papers by the authors, e.g. in [5].

The ASV transmits navigation and mission command data over the acoustic modem. The communication protocol is designed to work one-way, ASV to UUV. This is done to increase the update rate of the communication. However, two-way communication can be used to aid sonar detections and supply additional feedback about the UUV or target, i.e. the target image. Acoustic communication is generally slow and update rates are around 0.5 Hz, without counting packet loss. Also, a delay of around two seconds is introduced due to acoustic message transmission. This makes the navigation updates scarce and non-robust to clutter in near-bottom operation. Therefore, the guidance suffers following weaknesses during near-target operation:

1. *drift away from the path due to clutter detection*

- if false measurements are sent, the vehicle will drift from the path

2. *orientation away from the target*
   - in case of jumpy measurements oscillations around the line can be introduced, which will force the vehicle to oscillate in yaw around the line
   - in case of currents when the UUV is crabbing on the line
   - landing near the target could result in vehicle camera facing to the side of the object making it invisible in the final snapshot that is used to verify successful landing
3. *premature landing*
   - near-target operation can result in false detection of the sonar processing and can force the vehicle to land on the bottom believing it is on the target

We conclude that the used acoustic sensor, placed on a remote location – the surface vessel, is not adequate when the UUV reaches the target area. At the target area, the distance between the UUV and the target is much smaller with respect to the distance between the UUV and the ASV, so the acoustic navigation system has difficulties in differentiating the UUV from the target or surrounding clutter [6]. The value of a critical position in the target area varies depending on the vehicles and target dimensions, sonar resolution and its relative position. Although, the vehicle succeeds to land near the mine, even in presence of clutter and false detections, the final UUV orientation (heading) on landing cannot be directly controlled during line following. Visual loss of the target is important to allow for final verification before neutralization is activated. Usually a visual image is sent over the acoustic link to the operating center with the purpose of verification.

We intend to address these weaknesses by introducing the visual processing algorithm that will use the onboard camera for additional navigation updates when the UUV is in proximity of the target. To improve the navigation and control we need to combine the measurements with estimates of the UUV position. The model based estimation is needed since sonar updates are sparse. It also helps in smoothing and ignoring false measurements and outliers. The estimates and measurements are combined in the Extended Kalman Filter (EKF).

## 4 Extended Kalman Filter navigation

The main control loop of the vehicle is executed at 10 Hz, however, the available sonar measurements depend on the acoustic channel quality and usually arrive every 2-3 seconds. We can identify the dynamic parameters of the yaw, heave and surge degrees of freedom and use them with thruster forces to estimate the movement in between arriving measurements. While relative position measurements are sparsely estimated from sonar contacts, the heading and depth measurements are available continuously as the UUV has the onboard sensors to measure them. Finally, the video estimated position will be available on a higher rate than sonar measurements. The challenge is to perform estimation with measurement that arrives asynchronously. Luckily, the EKF can handle these problems. Usually in standard AUV applications the IMU sensor will drive the estimate at a synchronous rate, while Doppler velocity log (DVL), USBL or surface GPS measurements will arrive at a slower rate in an asynchronous and sometimes out-of-order fusion, i.e. as described in [7].

In order to perform estimation we need to derive the navigation model. We will use the SNAME notation as adopted in [8]. We define the vehicle position, velocities and exciting forces as in Table 1. Where surge, sway and heave are velocities in the body frame.

The frame notation can be defined as follows:

Table 1 **Notation used in control of marine vehicles**
Tablica 1 **Oznake korištene u upravljanju pomorskih vozila**

| DOF | surge | sway | heave | roll | pitch | yaw | defined in |
|---|---|---|---|---|---|---|---|
| **η** | $u$ | $v$ | $w$ | $p$ | $q$ | $r$ | {E}/{T} |
| **ν** | $x$ | $y$ | $z$ | $\varphi$ | $\theta$ | $\psi$ | {B} |
| **τ** | $X$ | $Y$ | $Z$ | $K$ | $M$ | $N$ | {B} |

The reference frames used can be defined as follows:

1. the body fixed frame {B} is attached to the vehicle's center of gravity. The three axis *x, y, z* point along, to the right and under the vehicle, respectively,
2. the NED frame {E} whose axis *x, y, z* point to the north, east and down, respectively,
3. the UUV navigation frame, {T}, which is similar to the {E} frame, except that its origin is above the target.

While the ASV navigates in the {E} frame, the UUV estimates its position in the {T} frame. This simplifies the aided navigation protocol, since the target position is never communicated to the vehicle; also it easily fits with target relative navigation as will be seen later on.

Using this notation we define the full state rigid body dynamic model as in [8]:

$$\underbrace{(\boldsymbol{M}_{RB}+\boldsymbol{M}_{A})}_{\boldsymbol{M}}\dot{\boldsymbol{\nu}}+\underbrace{(\boldsymbol{C}_{RB}(\boldsymbol{\nu})+\boldsymbol{C}_{A})}_{\boldsymbol{C}(\boldsymbol{\nu})}\boldsymbol{\nu}+\boldsymbol{D}(\boldsymbol{\nu})\boldsymbol{\nu}+\boldsymbol{g}(\boldsymbol{\eta})=\boldsymbol{\tau}+\boldsymbol{\tau}_{\boldsymbol{\epsilon}}$$

$$\boldsymbol{J}^{-1}\dot{\boldsymbol{\eta}}=\boldsymbol{\nu} \qquad (1)$$

where $\mathbf{M}_{RB}$ is the mass and inertia matrix, $\mathbf{M}_{A}$ represents the added mass matrix. The vehicle and added Coriolis and centripetal forces are denoted by $\mathbf{C}_{RB}$ and $\mathbf{C}_{A}$ respectively. The linear and quadratic damping term are defined in $\mathbf{D}$, while the $\mathbf{g}$ vector represents the resulting restoring force in presence of gravity and buoyancy.

The rigid body dynamics is modeled in detail by equation (1). Usually, it is hard to identify all the parameters needed to fully utilize the model for controller design. Therefore, the following assumptions are made to simplify the model:

1. vessel dynamics is uncoupled, i.e. coupled added mass terms are negligible, center of gravity *CG* coincides with the origin of the body-fixed coordinate frame {B}, and roll and pitch motion are negligible. As a consequence of these simplifications the total Coriolis and centripetal matrix $\mathbf{C}(\boldsymbol{\nu})$ vanishes, and restoring forces influence only the heave degree of freedom.
2. drag matrix $\mathbf{D}(\boldsymbol{\nu})$ is diagonal and each term can be approximated with a first order speed dependant term.

These simplifications lead to one, generalized, uncoupled, nonlinear dynamic equation (2) that describes surge, sway, heave and yaw degrees of freedom separately and it is given with

$$\alpha_{\nu}\dot{\nu}(t)+\beta(\nu(t))\cdot\nu(t)=\tau_{\nu E}+\tau(t). \qquad (2)$$

where $\nu$ is a single degree of freedom, $\alpha_{\nu}$ and $\beta(\nu)$ are model parameters where $\beta(\nu)=\beta_{\nu}\nu+\beta_{\nu\nu}|\nu|$, $\tau$ is a single degree of freedom (DOF) excitation force, and $\tau_{\nu E}=\tau_{NE}$ represents external disturbances. It should be mentioned that the heave (DOF) includes the difference between the weight and the buoyancy in addition to the control inputs.

The UUV is passive and stable in roll and pitch degrees of freedom and uncontrollable in sway. Therefore we focus on the remaining surge, heave, and yaw DOF. The parameters for these can be identified before the mission or after deployment using the identification via self-oscillation (I-SO) method [17]. The dynamic and kinematic vehicle model used for estimation is then:

$$
\begin{aligned}
\dot{u} &= -\frac{\beta_u(u)u}{\alpha_u} + \frac{X}{\alpha_u} + \xi_u \\
\dot{\omega} &= -\frac{\beta_u(\omega)\omega}{\alpha_\omega} + \frac{(Z+b)}{\alpha_\omega} + \xi_\omega \\
\dot{r} &= -\frac{\beta_r(r)r}{\alpha_r} + \frac{N}{\alpha_r} + \xi_r \\
\dot{b} &= \xi_b \\
\dot{x} &= u\cos\psi + x_c \\
\dot{y} &= u\sin\psi + y_c \\
\dot{z} &= \omega \\
\dot{\psi} &= r \quad \dot{x}_c = \xi_{x_c} \quad \dot{y}_c = \xi_{y_c}
\end{aligned}
\tag{3}
$$

where $[\xi_u, \xi_\omega, \xi_r, \xi_b, \xi_{x_c}, \xi_{y_c}]$ is the noise vector and $x_c$, $y_c$ are estimated currents in x, y, respectively. The input vector, τ, is defined as [X, Z, N], since these are the only three controllable degrees of freedom (DOF). Additionally, the estimate *b* is added to the heave DOF to estimate the difference between the weight and buoyancy forces. We can now use the discrete version of this model with the standard EKF equations, (9).

### 4.1 Sonar measurement updates

The multibeam sonar that is mounted on the ASV vehicle acquires the UUV highlight in the sonar image. Since the multibeam sonar is a 2D sensor it will detect the slant range and bearing of the echo. The sonar processing algorithm with choose a 2D position estimate between clutter echoes and the UUV echo. To calculate the 3D position the UUV depth has to be known, however, due to the one sided communication the depth is unknown. Therefore, it is necessary to send the measured sonar slant-range and absolute bearing in the {E}/{T} frame with the ASV position in the {T} frame. This way we can calculate the UUV position in the {T} frame as:

$$
x_{uuv} = x_{asv} + \cos(\psi_{asv} + \varepsilon)\sqrt{R^2 - \hat{z}^2}
\tag{3}
$$

$$y_{uuv} = y_{asv} + \sin(\psi_{asv} + \varepsilon)\sqrt{R^2 - \hat{z}^2}$$

Range (R) and bearing (ε) of the UUV with respect to the sonar head are determined from the sonar image while the depth of the UUV ($\hat{z}$) is measured onboard. The measurement covariance should be estimated based on the ASV state covariance and the sonar tracker but that would require sending additional data through the communication link. In cases where it is not possible we can assume the covariance matrix to be constant or diagonal and increasing with the slant range.

### 4.2 Video processing updates

The video processing is performed on board and data are available depending on the computing power of the mine intervention UUV. Usually the computing power will be smaller to reduce the price of the vehicle. However, with increased popularity of general purpose graphical processing units (GP-GPU) and their lower cost, improved video processing power is available and we can assume a decent refresh rate of 3-5 Hz for the video processing. This is a great improvement compared to the acoustic updates. Still, the challenge of detecting the target in the image with often poor visibility underwater makes it applicable only in the final approaching stage.

The video processing determines the relative position of the vehicle directly in the mine-centric coordinate frame {T}. Therefore, it is straight forward to use $x_{uuv}$, $y_{uuv}$ measurement during the correction step. The video image processing approach is described more in detail in the following sections. The measurement covariance is estimated by the image processing software and it is used during the sequential measurement fusion.

### 4.3 Measurement delay

Distributed systems often exhibit transmission delay or out-of-sequence packet arrival. This can present a problem in cases where delay becomes noticeable. The problems might arise when incorporating the delayed measurements into the Kalman filter. Measurement delays that are smaller than the sampling time can be incorporated optimally by adjusting the Kalman filter output equations [10]. However, delays larger than the sampling time require backward and forward propagation in time.

One way to optimally incorporate delayed measurements is applying the correction in the past and propagating the filter into the present time step. This requires us to save the initial state, initial covariance, all inputs and measurements that occurred until the delayed measurement arrived. Additional computation complexity arises with the EKF, since matrix derivatives, which are dependent on the state, need to be recalculated. The computational load in the sampling time when the delayed measurement arrives is increased by $N_s$ prediction/correction steps, where $N_s$ represents the approximate delay time in samples. However, in our case where we assume a constant delay of 20 samples the recalculation can be performed within the 100ms sampling time.

Alternatives to the standard Kalman filter recalculation can be found in [11], [12], [13]. The fact that is exploited in [12] is that for linear systems the covariance propagation is independent of the state and measurements. It only depends on the variances of these two values. Nonlinear systems generally have dependence between the estimate covariance and state. However, the Jacobian of the transition matrix will depend only on the heading and surge state. The heading is propagated by the compass and the effect of the *x*, *y* measurements are negligible. The surge speed is affected by the x, y measurements but in the line following

mode we assume a constant forward surge and therefore the surge speed could be treated as constant. Using these assumptions we can apply the Alexander's method described in [12].

Consider that we update the covariance at time ***s*** as if the position measurement was available. Propagating such a covariance yields suboptimal estimates between ***s*** and ***k***. However, at time step ***k***, when the measurement is available, a correction term is added to the estimate, which assures its optimality. The correction term is derived as follows.

The systems which have the measurement covariance **R** diagonal or constant can integrate arrived measurements sequentially, as described in [9]. Therefore, we can separate the filter correction into two sequential steps. We first perform filter correction with the measurements that are not delayed and then with the delayed measurements. Using the EKF equations we can write the correction equations for time step ***s***:

$$K_{i,s} = P^{-}_{i-1,s} H^{T}_{i,s} \left(H_{i,s} P^{-}_{i-1,s} H^{T}_{i,s} + R_{i,s}\right)^{-1}$$

$$\hat{x}^{-}_{i,s} = \hat{x}^{-}_{i-1,s} + K_{i,s}\left(y_{i,s} - H_{i,s}\hat{x}^{-}_{i-1,s}\right) \qquad (4)$$

$$P^{-}_{i,s} = \left(I - K_{i,s} H_{i,s}\right) P^{-}_{i-1,s}$$

where the index $i \in \{1, 2\}$ denotes the sequence number. We can now perform the first sequence to incorporate the available non-delayed measurements into the filter. The second sequence cannot be calculated because the measurements are not available. However, $\mathbf{H}_s$ and $\mathbf{R}_k$ are known at time ***s*** so we can update the estimate covariance as if the measurement is available. Note that the correction term $\delta_s = K_{2,s}\left(y_{2,s} - H_{2,s}\hat{x}^{-}_{1,s}\right)$ can be fully evaluated at time ***k*** when $y_{2,s}$ is known. Tracking the propagation of this correction from time ***s*** to ***k*** gives us:

$$\delta_k = \prod_{i=0}^{N-1} \left(I - K'_{k-i} H_{k-i}\right) A_{k-i-1}\, \delta_s \qquad (5)$$

where K' represents the Kalman gains calculated after the estimate covariance matrix was updated. This gain is different from the optimal gain because the measurement variance was incorporated in the past. As mentioned earlier this yields suboptimal estimates between ***s*** and ***k***. However, there is no need to recalculate the filter and the only computational load is to update the correction term at each time step. In [12] an additional modification is suggested where at time ***s*** a new Kalman Filter is started that will run in parallel to the optimal estimator. Then, at time ***k***, the delayed measurement is fused and the parallel filter will have the optimal estimate. Note that this imposes almost the same computation burden as recalculating the filter. However, now the burden is distributed over $N_s$ time steps.

# 5 Fusing vision detection and acoustic detection

Target detection is designed as an information fusion that uses several different methods. Each of these methods is independently sufficient for target detection, but in combination with others, more reliable detection can be achieved.

a) b)

Figure 5 **The results of two different analysis of the same frame: a) detection of characteristic key-points; b) image segmentation using threshold**

Slika 5 **Rezultati obrade istog kadra različitim metodama: a) detekcija karakterističnih točaka, b) segmentacija slike na temelju specificirane vrijednost.**

The first method detects the large differences in light intensity (see Figure 5a). This method is very variable with depth, light intensity, and small patterns on the bottom, but is very handy in detecting information about the environment, i.e. to differ the sea bottom from the sea surface, to detect large shadows, etc. By using this method, it is possible to eliminate areas where the target is less likely to be.

The second method uses the assumption that the target usually has a higher pixel intensity gradient on its surface than the background. Therefore, the area with more key-point density than other parts of the image has a larger probability of being the target (see Figure 5b).

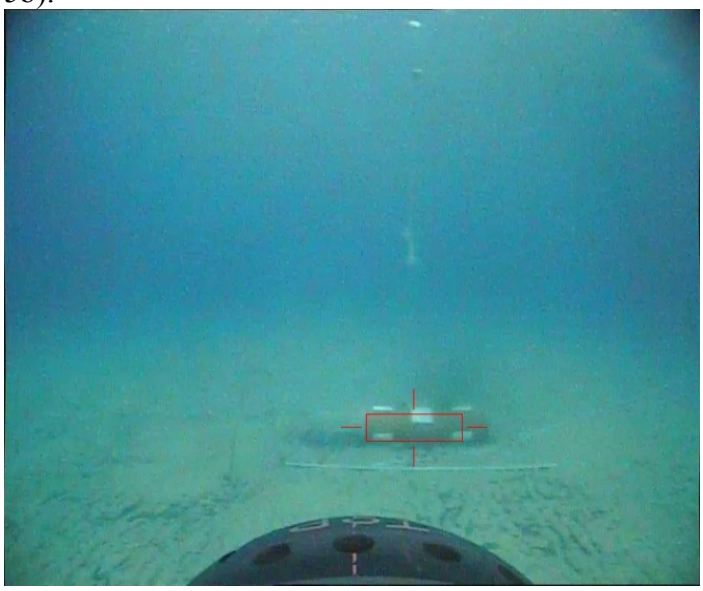

Figure 6 **Target recognition based on fusion**
Slika 6 **Prepoznavanje mete temeljem fuzije**

At the end of the frame analysis, the result is a possible detection with a combined probability calculated from the probabilities of different methods (see Figure 6).

Decision-making level: there is a need to choose between two measurements: a) position data from the sonar and b) position data from the video processing. The first one is better when the distance between the target and the vehicle is bigger. This border position value can be easily given by empiric approximation. Environmental conditions in marine and underwater operations are variable and unobservable. Therefore, the mentioned approximation values could be a reason of bad measurement results. An example: in good sonar conditions, the sea is quiet (without waves) and still (no currents), sea surface is flat and sonar visibility of the target and the vehicle from the surface is ideal. This leads to good measurement data and if detection is correct, we get the position with smaller deviations and variations. (Correct detection is not subject of this, but is very important for good results. It depends on the target location, existence of target-like obstacles and bottom configuration). Good visibility is a good underwater visual condition; however it does not need to be in correlation with sea surface conditions. It is fairly easy to decide to use the sonar up to 1m from the target and then the camera for final homing.

The method should calculate which measurement position has better precision and then decide. Alternatively, both measurements can be used to update the navigation solution. However, the measurement quality will vary based on the environmental conditions. Mathematically this will be reflected in the measurement covariance **R.**

### 5.1 Past target detection correlation

The same target is detected frequently when the UUV is close to the target. False recognitions can be eliminated with a target correlation algorithm. By matching more than three key-points in the target area it is possible to find the homograph matrix. Homography is a perspective transformation between two projective planes. After a series of plane transformations the relative route is determined. Knowing the exact time, the speed of the vehicle can be calculated easily. The information is now sufficient to predict the target position in the next frame.

## 6 Fusing vision and position information from remote distance acoustic sensors

Distance calculation is based on stereo vision with the difference that the stereo base is not fixed but variable. The value of the stereo base must be calculated because there is only one camera which moves in time. There is no need to know these two absolute positions but the relative movement must be known [14]. Underwater vehicles have 6 degrees of freedom. In this scope we will ignore 2 degrees (roll and pitch) because on the used UUV these can be assumed passively stable. So, there is an assumption that the roll and pitch are zero in each iteration. Yaw (heading), heave, surge and sway movements are left, the 6 axes system is simplified to a four-dimensional coordinate system. The stereo base calculation of the rotational movement is independent and can be separated from other movements.

### 6.1 Basic stereo vision theory

Measuring relative distance from the target using visual techniques means calculating depth $d$ of an object in image by using two different images of the same object. The concept of stereoscopic vision is to fuse these two images of the object taken by the same camera at different positions under the assumption of no change in camera orientation [15]. Clearly,

camera frames share an area in which the object can be seen from both positions (see Figure 7 (a) and Figure 7 (b)). The object projects towards the first and the second frame image measuring different width projection (L & R). The horizontal position difference of the same object point in different image is denoted as *a* while *f* presents the distance between the focal point and the image plane on the camera. The distance between two camera positions before and after the movement, called the stereo base, is denoted as *D*. If the camera is moved just lateral the distance value is a one dimensional vector. In order to calculate the depth, both object projections are considered together. In Figure 7 (*d*) the projection schemas are merged in one image, so now it is easy to notice the similarity between two triangles. Thus, equation (*7*) applies.

$$d = \frac{Df}{a} \tag{6}$$

If *a* is replaced as a product of the pixel number *n* and pixel length *p*, the equation (*6*) can be written as follows:

$$d = \frac{Df}{np} \tag{7}$$

### 6.2 Stereo vision at rotational movement

After a simple vehicle rotation around the focal point of the camera, the assumption of equal camera orientation is not fulfilled. The value of the stereo base is zero. Doing simple rotational and lateral transformations like in Figure 7 the projection schema can be transformed into the basic stereoscopic shape where the rotation becomes zero and the stereo base is described with the equation**:**

$$VirtualStereoBase[pixels] = \overline{RS} = \overline{RQ} - \overline{R'T} - \overline{TQ} = \overline{RQ} - 2 * \overline{R'T} \tag{8}$$

where:

*RQ - image width [pixels]*

*∢RTR' = ∢ROR'*

*∢ROR' - delta heading*

*R'T=R'O sin ∢R'OT ∢R'TO*

*∢R'OT= ∢ROQ- ∢ROR'2*

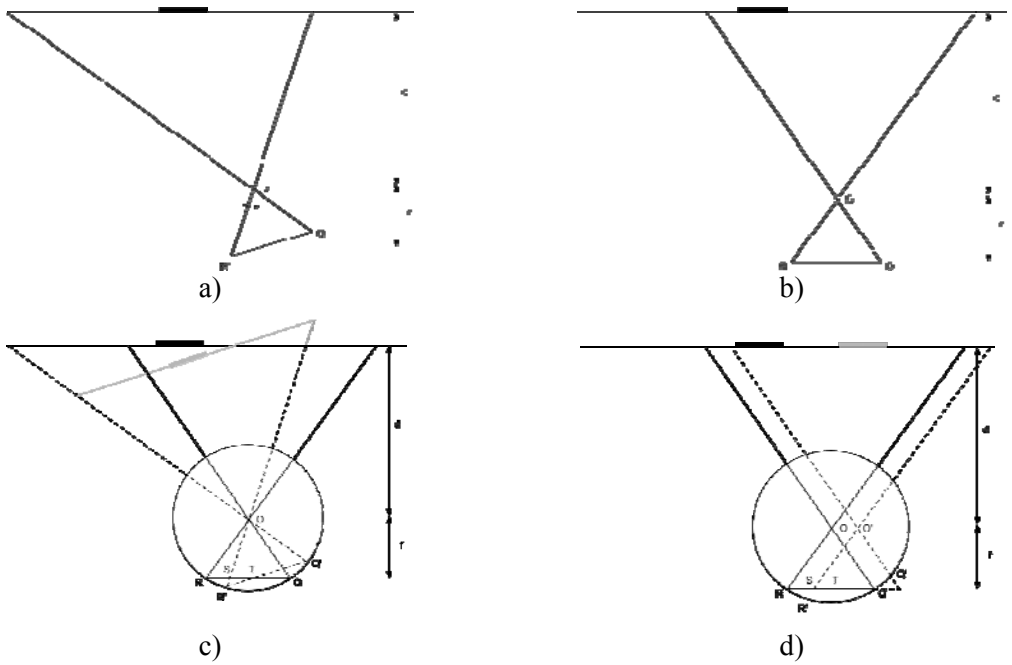


Figure 7 **Top view of camera, target, and photo projection due to camera rotation. Photos of target have been shot from the same location but with different orientation (a and b). Projection schemes are first shown in the same figure (c) and then the first scheme is rotated to match projection plane of the second scheme (d).**

Slika 7 **Tlocrt foto projekcije kamere i mete uslijed rotacije. Kadrovi s metom su snimljeni s iste lokacije ali pod različitim kutem (a i b). Projekcijske sheme su prikazane u istoj slici (c), a potom je prva shema zarotirana na način da se poklapaju projekcijske ravnine**

## 6.3 Stereo vision at longitudinal and transversal movement

Transversal movement is a basic stereoscopic case when the position of the camera is changed in the lateral direction. When the position is changed in both longitudinal and lateral directions at the same time, then the stereo base becomes a two dimensional vector. Figure 8 shows how to move the projection schema in order to accent similarities between triangles in the merged projection schema.

$$d = \frac{\frac{f_p}{2} * \sqrt{dx^2 + dy^2}}{\cos\left(\arctan\frac{\Delta y}{\Delta x}\right)} \qquad (9)$$

where:

dx - longitudinal shift value
dy - transversal shift value
fp - focal distance in pixels
a - horizontal position difference of the same target point in different frame

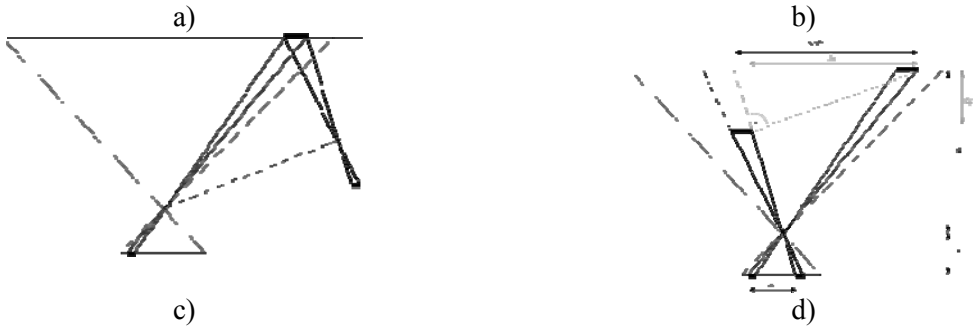


Figure 8 **Top view of camera, target, and photo projection due to camera translation. Photos of target have been shot from different locations but with equal orientation angle. Projection schemes are first shown in the same figure (c) and then the scheme of position A is translated to match the focal point of position B.**

Slika 8 **Tlocrt foto projekcije kamere i mete uslijed rotacije. Kadrovi s metom su snimljeni s različite lokacije ali pod istim kutem (a i b). Projekcijske sheme su prikazane na istoj slici (c), a potom je prva shema pomaknuta na način da se poklapaju projekcijske ravnine**

## 6.4 Stereo vision at vertical movement

In practice, there is vertical movement. The influence of this effect is just added to the relative depth due to the equation:

$$d = \sqrt{dx^2 + dy^2 + dz^2} \tag{10}$$

Data from the pressure sensor are used to get the value of the depth shifts that are usually very small. Due to that small value, the vertical shift has very small effects on the total value of the relative position.

## 6.5 Experimental research

The first task that was performed was developing an offline application for simulating conditions in UUV system. The simulation is necessary to test the algorithm efficiency in real conditions by using real data from the past field experiments.

Video and data logging, needed as simulator inputs are separated files without exact time of sampling. The method used for synchronization is based on manual detection of the moment that could be recognized in video and data log. The landing is usually the synchronization moment. Figure 8 shows the results of stereoscopic ranging with the monocular vision system. Rotation movement is excluded from the calculation on purpose to avoid errors caused by very small values and imprecise synchronization [16]. Rotation impact is minimized by taking only short steps of the UUV absolute position while progressing forward to the target.

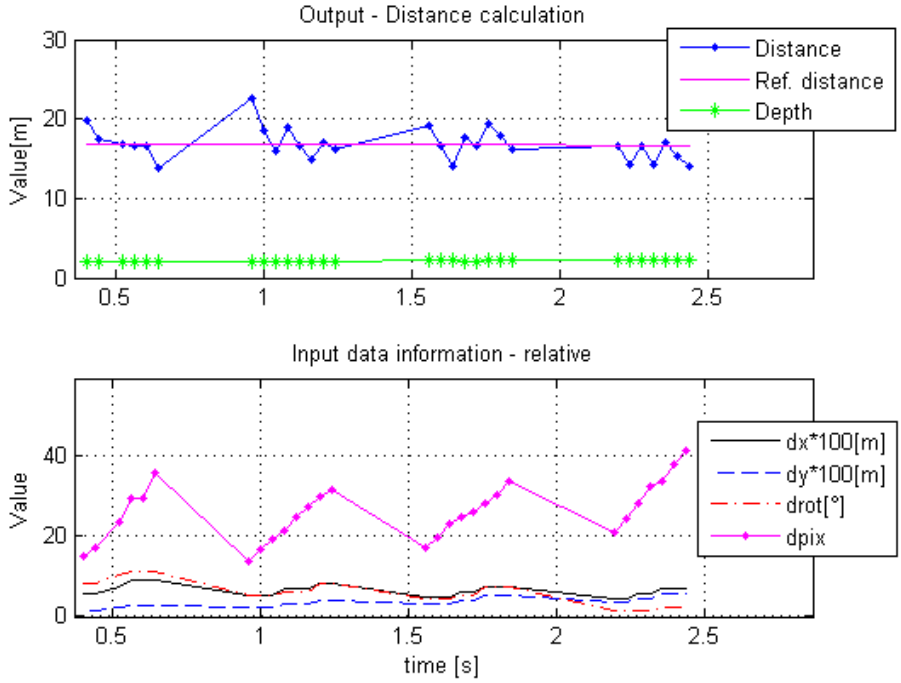


Figure 9 **Stereoscopic ranging results with monocular vision system during mission**
Slika 9 **Rezultati stereoskopskog mjerenje udaljenosti mono-vizualnim sustavom za vrijeme misije**

In the first graph, the calculated absolute distance value is compared with the estimated value. Also rotation is shown but not added to the result value. Observing behavior of all input parameters from the second graph, the effect of the deviations on the result can be easily seen.

UUV visual navigation system uses data information that comes from different sensors and even different vehicles. Different sensor values are gathered to calculate relative distance using equations 6 - 10. Small deviations in sensor values are multiplied in these equations. Deviation impact on small values in stereoscopic ranging causes error divergence, which becomes the main method disadvantage. This is solved by limiting the minimum values of the input signals.

Additional error generator is time delay and time drifts between different types of sensor values. These values are extracted from video frames and sonar images. This is also a weak link in algorithm because an extra time is needed to process the video frames. The process time depends on mission requirements for the UUV. Generally, it is assumed that the time delay of this signals is constant during missions. The position information comes from ASV assuming the both vehicles (ASV and UUV) are changing position during mission so the time delay is not constant. Using the mathematical model of ASV and all known variables about the system, the time delay estimation is done.

## 7 Conclusion

The paper presented a part of a collaborative autonomous system. The navigation of the mine intervention UUV subsystem was analyzed. The first part of the paper presented the current state of the aided navigation system that is based on remote sonar sensing from a more advanced autonomous platform. Potential problems with the system in cluttered environments and during the final approach sequence were identified. A solution to use the onboard visual sensing for final homing to the target is proposed. A situation where underwater visibility is poor limits the use of the onboard visual sensor. However, in cases where we have limited, but usable visibility underwater, visual sensing can be employed. Therefore, in the second

part of the paper we focused on employing the onboard camera of the mine intervention UUV to aid in the final approach stage to the target. The first testing by using past video data and navigation logs has promising applicability of the method in real-life situations. We have shown that fusing monocular vision measurements of a single feature with history position data generates a feasible sensing strategy for determining reliable position information.

The main contribution to this research is the design of stereo photogrammatery algorithm with a monochrome camera. This article represents a first step in the exploration of a new sensing strategy. The sensor fusion problem has been researched only in aspect of position value determination. Detection effectiveness, reliability and value precision are just a few areas where fusion should be used to improve result.

Future work will include testing of visual aided guidance in real-time on the intervention UUV. Extending the system to two or more vehicles guided in the same manner from a single master vehicle will increase cost effectiveness, area coverage and extend on possible applications of the autonomous system. An active issue, as mentioned earlier, is the anti-swimmer/diver application for harbor protection. Without extensive adaptations the autonomous system described can be utilized in such applications. Therefore, future work will include extension to the navigation system to be handled in relative terms, meaning, the distance between the movable target and the intervention UUV will be estimated. Estimating the target movement from received acoustic and possibly visual measurements will also be used in the multi-vehicle scenario to estimate relative distance and movement of the other vehicle in order to allow collision-free and coordinated operation in the area.

## References


[1] DJAPIC, V., NAĐ, Đ.: **"**Collaborative Autonomous Vehicle Use in Mine Countermeasures"**,** November 2010, Sea Technology.

[2] GROEN, J., COIRAS, E., WILLIAMS, D.P.**:** "False-alarm reduction in mine classification using multiple looks from a synthetic aperture sonar", Proceedings of the IEEE Oceans. 2010.

[3] GALCERAN, E., et al.: "A Real-time Underwater Object Detection Algorithm for Multi-beam Forward Looking Sonar", 2012. Navigation, Guidance and Control of Underwater Vehicles. 2012, Vol. 3.

[4] MIŠKOVIĆ, N., NAĐ, Đ., VUKIĆ Z.: "3D line following for unmanned underwater vehicles", Brodogradnja / Shipbuilding: Croatian Journal of Naval Architecture and Shipbuilding Industry. February 2010, Vol. 61.

[5] NAĐ, Đ., MIŠKOVIĆ, N., DJAPIĆ, V., VUKIĆ, Z.: "Sonar aided navigation and control of small UUVs", Proceedings of the 19th Mediterranean Conference on Control and Automation (MED). 2011. .

[6] MIŠKOVIĆ, N., et al.: "Multibeam sonar-based navigation of small UUVs for MCM purposes", Proceedings of the 18th IFAC World Congress. 2011.

[7] MILLER, P.A. et al.: **"**Autonomous Underwater Vehicle Navigation"**,** 3, July 2010, IEEE Journal of Oceanic Engineering, Vol. 35.

[8] FOSSEN, T.: "Guidance and control of ocean vehicles", 1994. Chichester , John Wiley & Sons, 1994.

[9] SIMON, D.: “Optima state estimation: Kalman, H-infinity and nonlinear approaches”, s.l. , Wiley-Interscience, 2006.

[10] ANDERSEN, G.I., CHRISTENSEN, A.C., RAVN, O: “Augmented models for improving vision control of a mobile robot”, 1994. Control Applications.

[11] GOPALAKRISHNAN, A., KAISARE, N. S., NARASIMHAN, S.: **“**Incorporating delayed and infrequent measurements in extended kalman filter based nonlinear state estimation”, 2011, Journal of Process Control, pp. 119-129.

[12] LARSEN, T.D., et al.: **“**Incorporation of time delayed measurements in a discrete-time kalman filter”, 1998. Proceedings of the 37th IEEE Conference on Decision and Control.

[13] BAR-SHALOM, Y.: “Update with out-of-sequence measurements in tracking: exact solution”, 38, 2002, IEEE Transactions on Aerospace and Electronic Systems, Vol. 3, pp. 769-777.

[14] HUSTER, A.: “Relative position sensing by fusing monocular vision and inertial rate sensor”, Ph. D. Dissertation. s.l. , Department of Electrical Engineering and the Comitee on Graduate Studies of Stanford University, July. 2003.

[15] COPERTARI, L.: **“**Stereoscopic vision for depth perception”, No. 2, Zacatecas, Mexico : Investigacion cientifica, 2007, Vol. Vol. 3. ISSN 1870-8196.

[16] MATTHEIS,L., SHAFER, S.: “Error modelling in stereo navigation”, Computer Science Department, [Online] 1986. http://repository.cmu.edu/compsci/1593. 1593.

[17] MIŠKOVIĆ, N., VUKIĆ, Z., BIBULI, M., BRUZZONE, G., CACCIA, M.: “Fast in-field identification of unmanned marine vehicles”, Journal of Field Robotics, February 2011, Vol. 28.