Grant Agreement No. 611373

FP7-ICT-2013-10

D2.1. Diver remote sensing technologies in the underwater

Due date of deliverable: 31.05.2014.
Actual submission date: 31.05.2014.

Start date of project: 01 January 2014 Duration: 36 months

Organization name of lead contractor for this deliverable: UNIZG-FER

Revision: version 1

Dissemination level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Contents

1 Introduction

A crucial part in creating a cognitive diving assistant robot is to have a diver sensing mechanism, a way for the robot to perceive the diver, to be aware of its location and movement, communicate with him/her by recognizing symbols or realizing if the diver is experiencing a hazardous state such as nitrogen narcosis, or any other similar problem.
Because of all that, diving assistant (BUDDY) will be equipped with a number of devices used to sense the diver. A stereo camera will be used to provide the means of locating the diver in the surroundings and creating a 3-dimensional model of the diver. It will also be used for recognizing diver's symbolic gestures and behaviour. In addition to stereo camera, a multi-beam sonar will be used as well for the purpose of detecting the diver. Also, a bottom-facing mono camera will be mounted for the purpose of seafloor mapping. While the cameras will be used for close range detection of the diver or parts of the diver, sonar sensing is planned to be used at larger distances or in situations when camera does not provide clear image, such as murky waters.
To detect diver's movement a network of inertial sensors, DiverNet, was developed. This will allow for better tracking of divers behaviour and detecting if something is wrong.
The purpose of this deliverable is to explain the technology that will be used in more detail, how it will be used in the CADDY project and which equipment specifically will be used. Modifications made to allow the underwater use will be shown. Some popular principles in underwater image processing, computer vision and gesture recognition, as well as work in human body tracking with inertial sensors will be mentioned. A general plan for data processing and algorithm implementation will be introduced. Finally, a report on the initial data gathering will be given.

2 Available technologies and principles of operation

In this chapter equipment used for remote diver sensing will be described. Basic principles of operation will be explained for each of the used technologies, as well as some details specific to the equipment used in CADDY project.

2.1 Mono camera

An ordinary (mono) video camera is a device used for electronic motion picture acquisition. It has a single image sensor that converts optical image into electronic signal that, in case of digital cameras, gets converted to sampled (digital) output. Two of the most common sensor types are charge-coupled devices (CCD) and MOS transistor based sensors (most commonly CMOS).
Mono camera's main purpose will be seafloor mapping. However, in case of emergency (e.g. if the forward-facing stereo camera loses the diver) its image might be used for other purposes as well.

2.1.1 Bosch FLEXIDOME IP starlight 7000 VR

Bosch dome mono camera will be used. Some of its basic features include:

- ⅓ inch CMOS HD 1.4MP sensor
- 60 FPS @ 720p
- *motorized focus and zoom adjustment*
- light sensitivity: 0.017 lx in color, 0.0057 lx in monochrome

More details about the camera can be found in [1].
One of the main advantages is its very high light sensitivity, specific for the Bosch "starlight" camera series. This should provide above-average performance in low light conditions. The camera is not intended for underwater use so a modification will be made.

Fig. 1 shows the stock camera photo. Modifications for underwater use will be described in a later chapter.

Fig. 1. Bosch FLEXIDOME IP starlight 7000 VR camera

2.2 Stereo camera

Stereo video camera has multiple (typically two or three) imaging sensors. This enables recording of the same scene from different viewpoints at the same time, which can be utilized to create a 3 dimensional image in the same way as the human brain combines picture from both eyes through a process called stereopsis to create the perception of depth.

Stereo computer vision algorithms designed to do that are based on finding matching objects in multiple images in the same video frame. For this type of algorithm to work, some pre-processing is needed:

- Removing distortion - different kinds of image distortion may occur, causing, for example, straight lines being captured as curves due to distortions from wide-angle lenses. [2]
- Epipolar rectification - each sensor on a stereo camera captures the image from its own point of view. Epipolar rectification is a process of projecting images into a common plane. [3]

Fig. 2 gives a visual explanation of epipolar geometry principles.

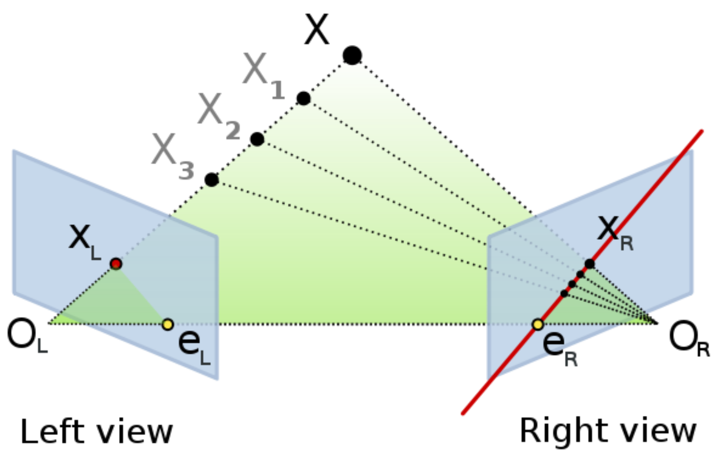


Fig. 2. Epipolar Geometry: Epipoles (eL and eR), Epipolar Line for XL (red). All candidates for correspondence of XL are located on the line in the other image, reducing the search space significantly.

These quality of the pre-processing steps have a great impact on the success of stereo computer vision algorithms. Specifically, the epipolar rectification is a measure of computational optimization - after rectification, all epipolar lines are horizontal, thus allowing easy access of pixels on the epipolar line for matching by iterating through pixels horizontally in the rectified image.
A stereo camera will be used in the CADDY project to create a point cloud model of the observed diver. A point cloud model is a representation of an object (in this case a 3 dimensional object) with a set of points in space.
In the underwater domain, two specific challenges exist that need to be overcome in order to produce high-resolution, high-quality point clouds from stereo images. First and foremost, exact epipolar rectification is not possible due to the refractive effects of the air/water interface of the stereo camera housing. Fig. 3 shows this problem with apparent focal points of camera in water.

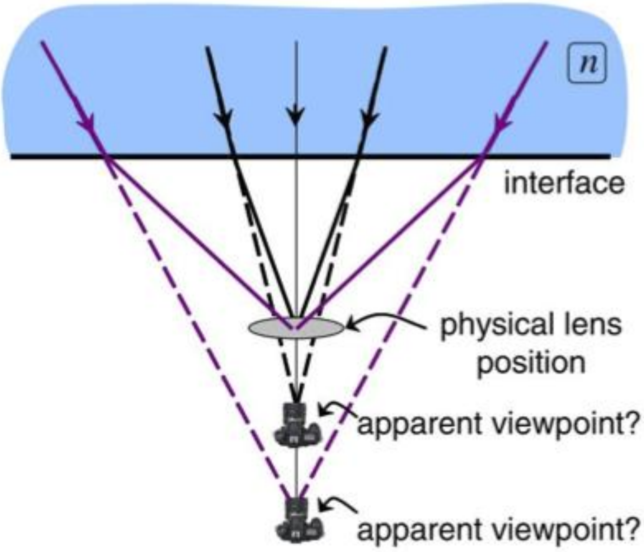


Fig. 3. Illustration of refractive effects on apparent camera focal points in water. [23]

Due to rays in water converging to different apparent focal points inside the housing, the plane epipolar rectification would use does not exist in this setup. Depending on the calibration, though, it is possible to approximate it. However, it is our aim to compute exact look-up tables to follow the epipolar curves on the second (or third) camera's undistorted image instead. Calibration of this phenomenon is currently under investigation at JACOBS University.
The second challenge is color absorption by water. After computing depth, color correction based on the knowledge of the color absorption properties of water is possible. Very recent related work [27] showed that performing simultaneous color correction and depth estimation may be beneficial.
Fig. 4 shows underwater color correction based on depth. Fig. 5 shows an example point clout underwater model.

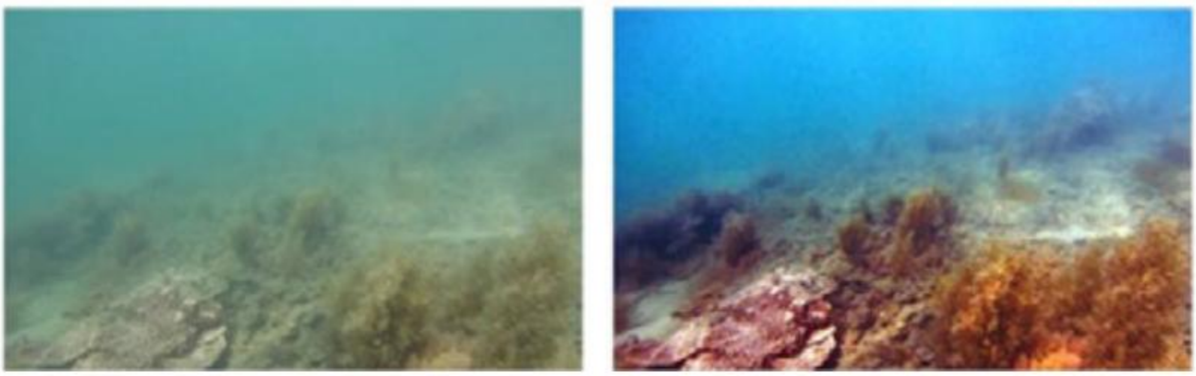


Fig. 4. Before and after color correction from depth cues. [27]

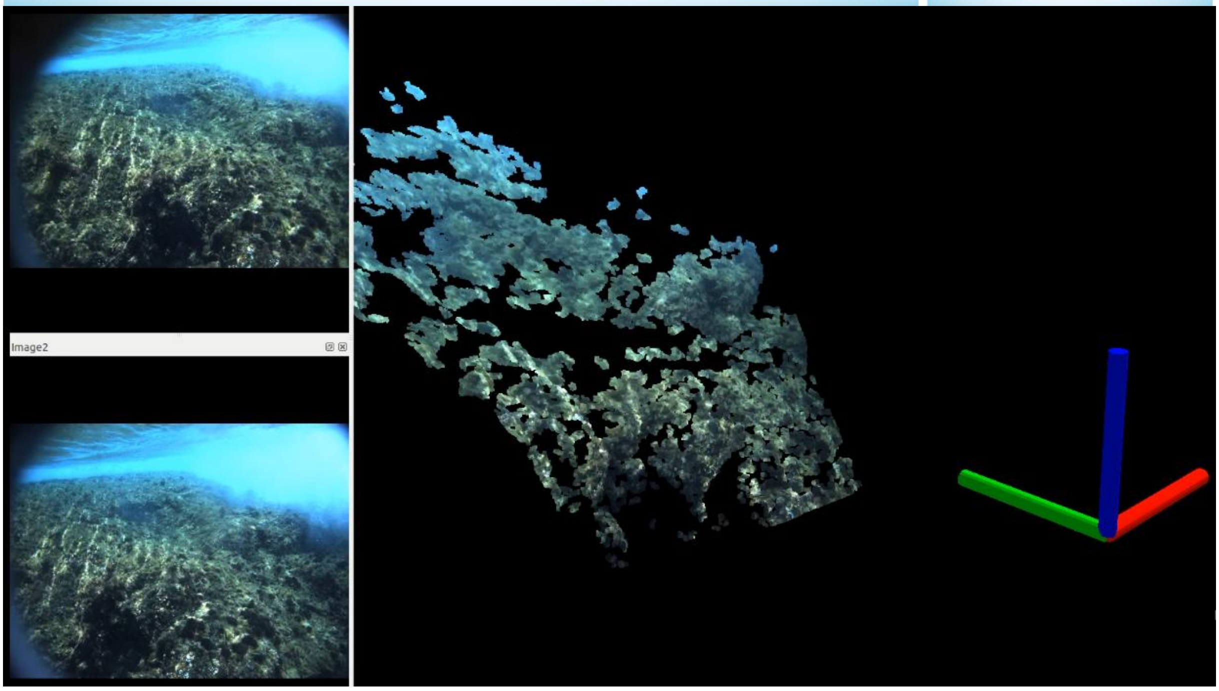


Fig. 5. Example point cloud generated by JACOBS' smart stereo camera. Approximations of refractive effects are used here. While the point cloud seems of good quality visually, it remains to be seen how much the approximation influences metric quality. Color correction is not yet addressed in this preliminary result.

2.2.1 *Point Grey Bumblebee XB3*

Bumblebee XB3 by Point Grey is a 3-sensor stereo camera that will be used on CADDY.

- 3 Sony ICX445 1.3 MP 1/3" CCD sensors, set in a line at 120 mm apart
- Monochrome, fixed focal length 3,8mm
- 16 FPS @ 1280x960
- 2 x 9-pin IEEE-1394b for camera control and video data transmit

More detailed info on camera specification is available in [4].

The IEEE-1394b connection standard has a limit on cable length of maximum 5 meters. This will not pose a problem when the camera will be mounted inside BUDDY as the processing unit will be very close, but limits the options for testing the camera.

Fig. 6 displays the Bumblebee XB3 camera.

Fig. 6. Point Grey Bumblebee XB3 stereo camera; source: ptgrey.com

2.3 Multi-beam sonar

Sonars (originally acronym for sound navigation and ranging) are devices that use sound propagation for navigation and object detection. They are dominantly used underwater where most of the

wireless technologies used in air (radar, GPS…) do not work, but sound propagates much better and faster. Speed of sound underwater is approximately 1500 m/s and propagation losses are much lower, allowing, for example, humpback whales to communicate over distances of hundreds of kilometres.
To understand how multi-beam sonars work some basic knowledge about sonars in general is necessary.

2.3.1 Basic single beam sonar principle of operation

Sonars can be divided into two groups - **passive** and **active**. Passive sonars are used to detect sounds created by other objects in the environment such as ships, submarines and seismic activity. On the other hand, active sonars are creating sounds of their own (of specific frequencies and direction) and are listening to echoes of those sounds reflected from remote objects. From this point on, only active sonars will be described.
To generate a short pulse of sound (often called *ping*) sonars use a device called *projector*. After receiving an echo (obtained with a *hydrophone*, a microphone-analogous device used to convert sound to electrical energy under water) they calculate distance from the object based on speed of sound and time elapsed between transmitting and receiving the ping.
Fig. 7 illustrates the operation of a single beam sonar. During the time t between transmitting the ping and detecting its echo, sound travels a total distance of $2r$ with velocity v. Distance can be calculated from those values.

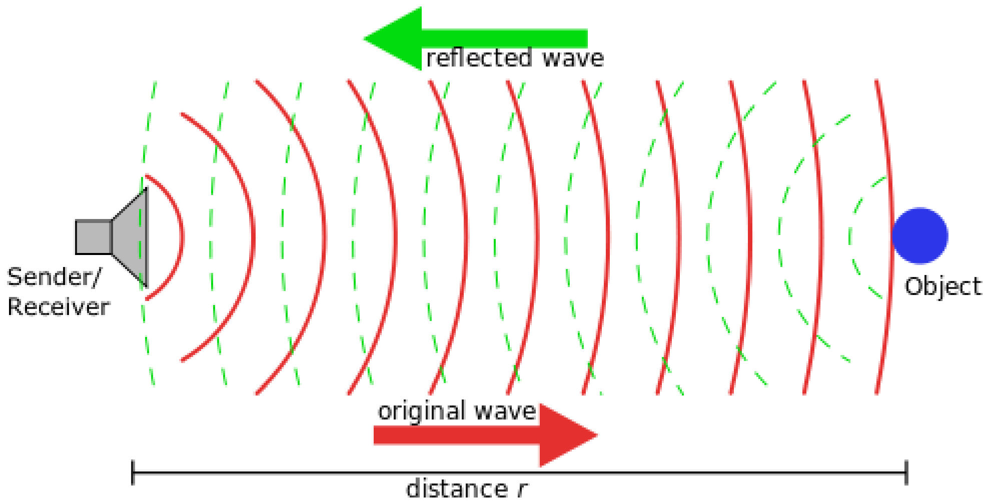


Fig. 7. Single beam sonar operation; source: Wikipedia.org

Single beam sonars have several crucial limitations. If a wide beam is used, it will get reflected of a larger area and many echoes will be received. It is possible to get some useful data out of it, but this process will basically give the information about average distance to the objects or bottom at a large area covered by the beam. If a narrow beam is used (beam forming process will be described in next sub-chapter), information about the distance to a smaller area will be obtained that.
Width of the sonar beam affects the resolution obtained from sonar data. It can be used to obtain, for example, current depth, but for the purpose of detecting more complex objects such as diver with sufficient details (head, limbs, possibly fists and fingers) single beam sonars are not applicable.

2.3.2 Multi-beam sonar operation and advantages

Multi-beam sonars were developed as an answer to main single beam sonar limitations shortly described in previous sub-chapter.

The main idea behind multi-beam sonars is to use a large number (typically around 100) narrow beams of slightly different frequency. The result is effectively the same as firing a single narrow beam sonar 100 times in various directions simultaneously, but is achieved with far less projectors and hydrophones by exploiting constructive and destructive wave interference, as well as different frequency of each beam. Details about narrow multi-beam forming and detection can be found in [5].
Two of the most common uses of multi-beam sonars are *echo sounding* and *sonar imaging*. Beam formed for this purpose is very narrow in the moving direction, but very wide in perpendicular axis.
The second expression, *sonar imaging*, is perhaps not the most specific one, but will be used for purposes of using sonars in a camera-like mode to scan the surroundings.
Echo sounding and sonar imaging, except with different beam forms, also often work on different distances. Echo sounding is used to measure large depths, while sonar imaging is mostly used at shorter distances. This will also be the case for CADDY where the diver will be positioned several meters from the sonar.
Fig. 8 shows two for mentioned sonar uses with different beam forming. The right picture shows a beam typical for echo sounding. The picture on the left displays the imaging sonar how it will be used in CADDY. The beam is facing straight forward from the robot (or is slightly tilted downwards) and is wide in both directions, looking similarly to movie projector beam and covering a large area of interest.

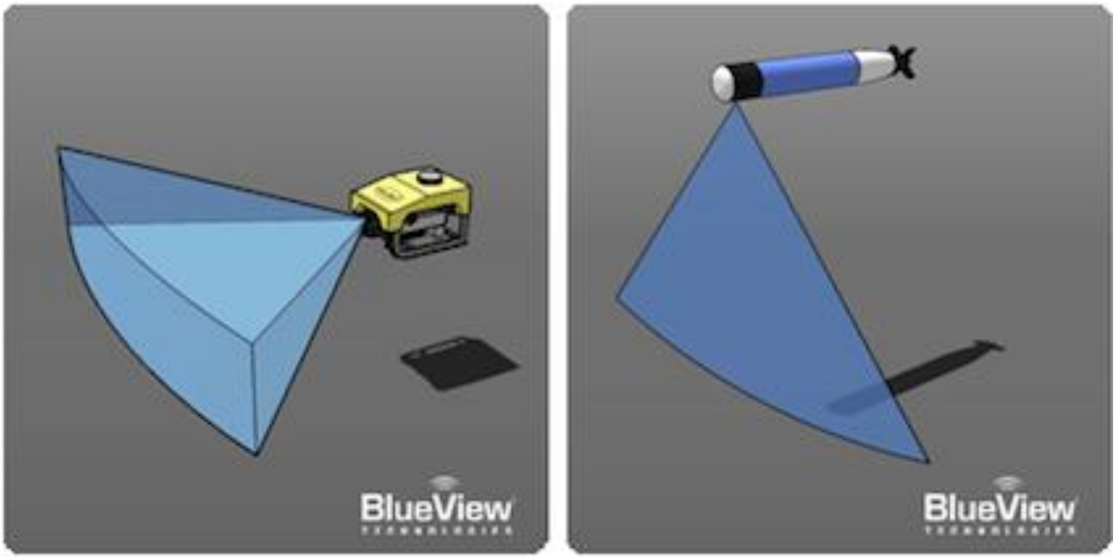


Fig. 8. Sonar imaging and echo sounding beam forming; source: Blueview.com

2.3.3 Sound Metrics ARIS Explorer 3000

Sound Metrics offers three types of ARIS Explorer multi-beam imaging sonars. Model 3000 will be used, which is intended for short-range application.
Key features of the ARIS Explorer 3000 are:

- 128 transducer beams
- Dual frequency
  - 3 MHz, range 5m
  - 1.8MHz, range 15m
- Field of view 30°x14°, beam width 0.25°
- Frame rate up to 15 fps

ARIS Explorer 3000 is displayed on Fig. 9. More info about it is available in [6].

Fig. 9. ARIS Explorer 3000 unit and equipment; source: soundmetrics.com

2.4 DiverNet

DiverNet is an array of inertial measurement units (IMUs) developed at University of Newcastle, with a purpose of capturing the motion of the diver.
Each IMU unit consists of a 3-axis accelerometer, 3-axis gyroscope and 3-axis magnetometer. A total of 17 IMUs is placed at designated positions on the diver's body:

- 3 on each leg (foot, shin and thigh)
- 3 on each arm (hand, forearm, upper arm)
- 1 on each shoulder
- 1 on lower back
- 1 on the chest
- 1 on the head

They are all connected in four main branches, and each branch gets connected into a box mounted on diver's back. Fig. 10 shows the DiverNet system, box with connected IMUs.

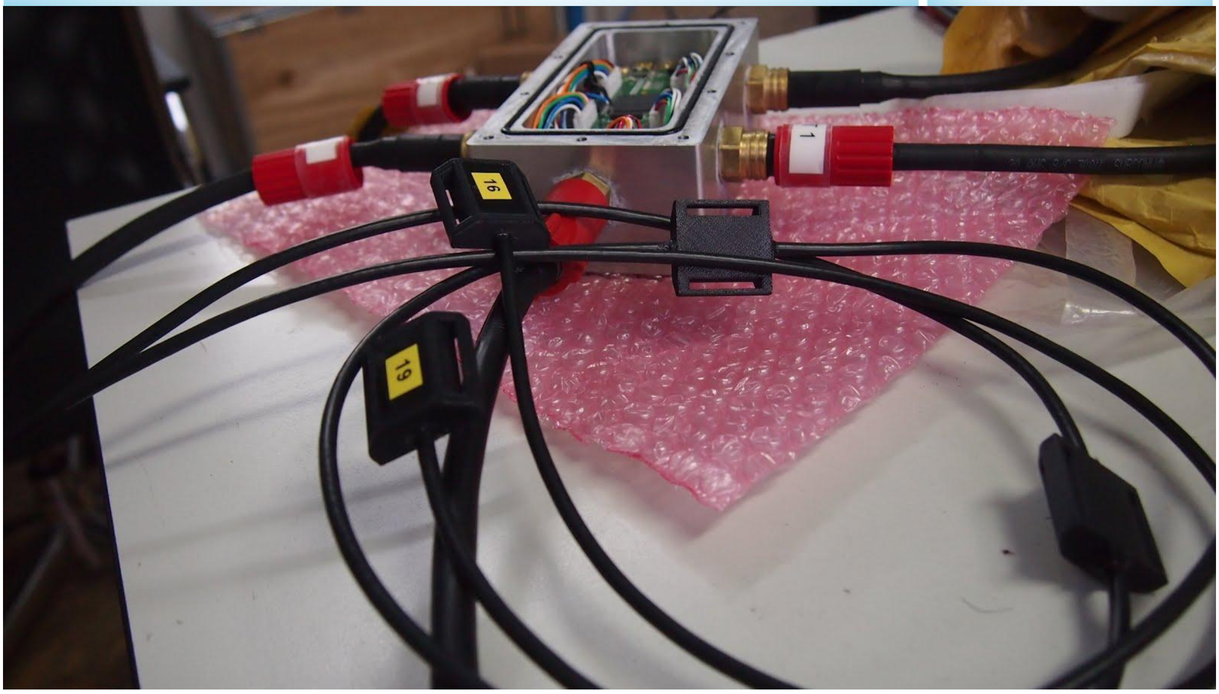


Fig. 10. DiverNet box and IMUs
In addition to IMUs, a breathing belt will also be connected to capture diver's breathing pattern

3 Modification of systems for underwater use

Neither camera is designed for underwater use so it was necessary to make necessary modifications. Waterproof cylinders for housing the cameras were designed and constructed. The cylinder itself is made of acrylic glass, sealed with corks made of polyacetal (POM-C) industrial plastics. They were tested without leaking at 5.5 bar.
Fig. 11 shows a render of the cylinder with stereo camera. Fig. 12 displays both stereo and mono camera mounted in their casings.

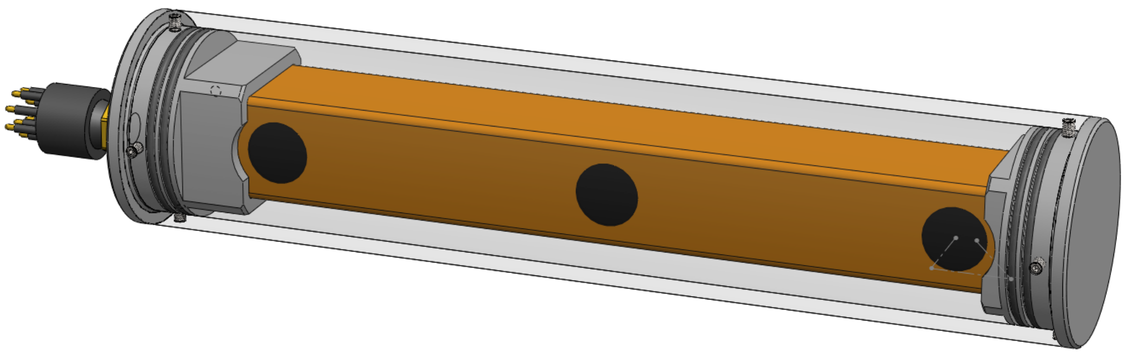
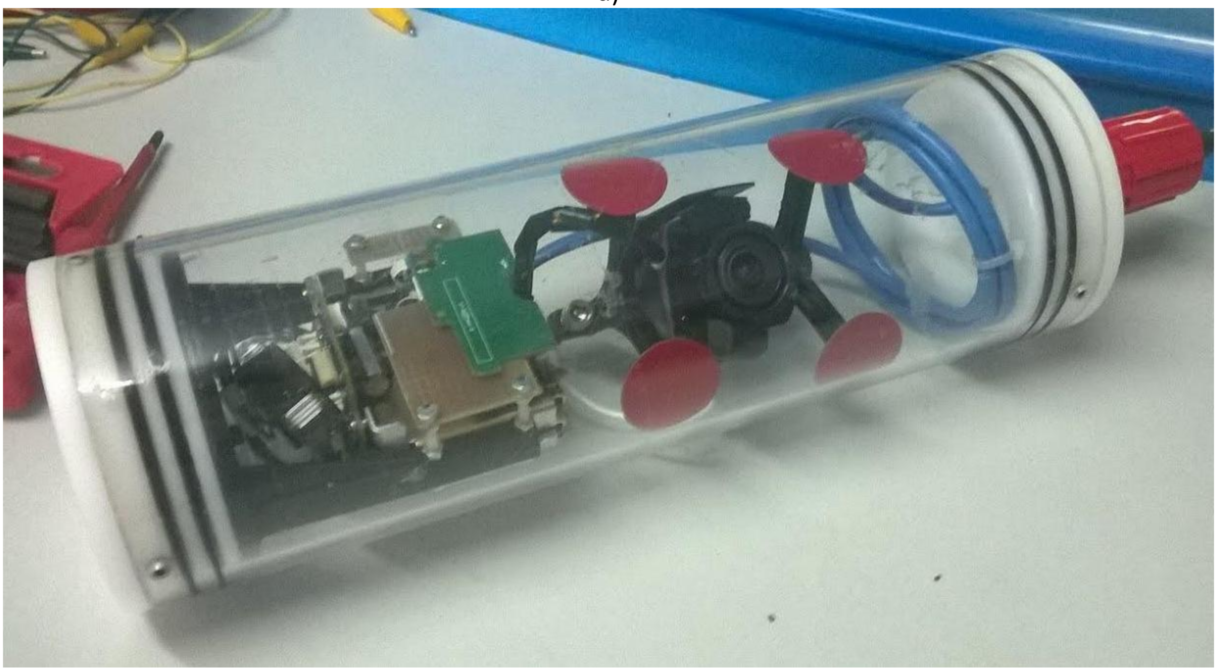


Fig. 11. Rendered case with mounted stereo camera

a)



b)

Fig. 12. a) stereo camera mounted in cylinder casing; b) mono camera mounted in cylinder casing

Subconn ethernet cables and connectors were used for connection. Since the stereo camera operates via IEEE-1394b standard, ethernet cables were just used to transmit IEEE-1394b data and power.

4 State of the art in underwater perception

A number of articles that deals with the specific problem of underwater diver detection is pretty small. In [7] and [8] authors deal with the same problem of detecting and tracking divers with a robot diving buddy. They are using mono camera to detect the diver.
Underwater stereo cameras and point cloud reconstruction are typically used in creating models of stationary objects or seafloor [9], [10], [11].
Sonars in diver detection are mostly used for intruder detection in places such as oil platforms [12]. Using sonar images to locate and track the diver will be a new approach.
A lot of data is available on human detection on land, underwater image processing and underwater computer vision in general. Most of the successful human detection algorithms are based on histograms of oriented gradients, first developed in [14], or the Viola-Jones cascade algorithm [13]. Another very notable work introduces a cascade of histograms [15], combining ideas from the two previously mentioned ones. However, since the underwater images are not nearly as clear as the land ones, some image processing is necessary. Overview of some underwater image enhancement and restoration methods can be found in [16].
As for the DiverNet and estimating orientation, a large number of articles is available on single IMU sensor filtering, using extended Kalman filters [17], [18], a simple IIR-like complementary filter or the class of popularly called Mahony-Madgwick filters [19], [20].
Estimating human body orientation was researched for a specific body part [21], [22] or the orientation of the entire body [23], [24], which is the ultimate goal for DiverNet.
Another interesting option, given the presence of both stereo camera and inertial sensors, can be in combining the data obtained from those two - the stickman figure and point cloud model - to get the diver's orientation and movement more precisely than any of the technologies could do by itself.

5 Processing

In this chapter an approximate collection of ROS nodes that will be used to process the sensor data will be described, together with the planned processing pipeline.

5.1 DiverNet

The first DiverNet ROS node will simply read raw binary data and publish the unfiltered accelerations, angle velocities and magnetic fields. The second node will conduct filtering, both for each of the IMUs individually (using a complementary filter, extended Kalman filter or some other method) and for the entire net (estimating body parts positions from all the IMUs). It will publish data for estimating orientation (Euler angles or quaternions).
Fig. 13 shows a block diagram with planned processing of DiverNet data.

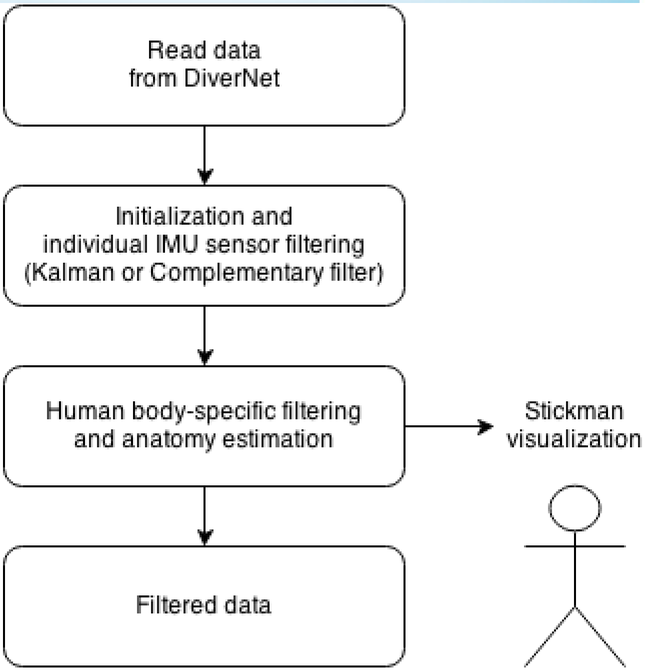


Fig. 13. Block diagram approximating planned DiverNet processing.

5.2 Stereo camera

Several ROS packages that will most likely prove useful are present in *image_pipeline* package.

image_pipeline

- *camera_calibration* - used to calibrate both mono and stereo cameras
- *image_view* - displays images
- *image_rotate* - rotates images
- *image_proc* - mono image processing; distortion removal, color interpolation
- *stereo_image_proc* - stereo image processing; distortion removal, color interpolation, point cloud and disparity image generation
- *depth_image_proc* - processes depth images; most likely will not be used

In addition to existing nodes, a number of new ones will be developed with strong use of OpenCV C++ computer vision library.

Fig. 14 displays a block diagram with planned processing of stereo images. This is just an approximate as it is still unclear which parts of the processing pipeline will be shared by stereo part (point cloud generation) and diver detection part.

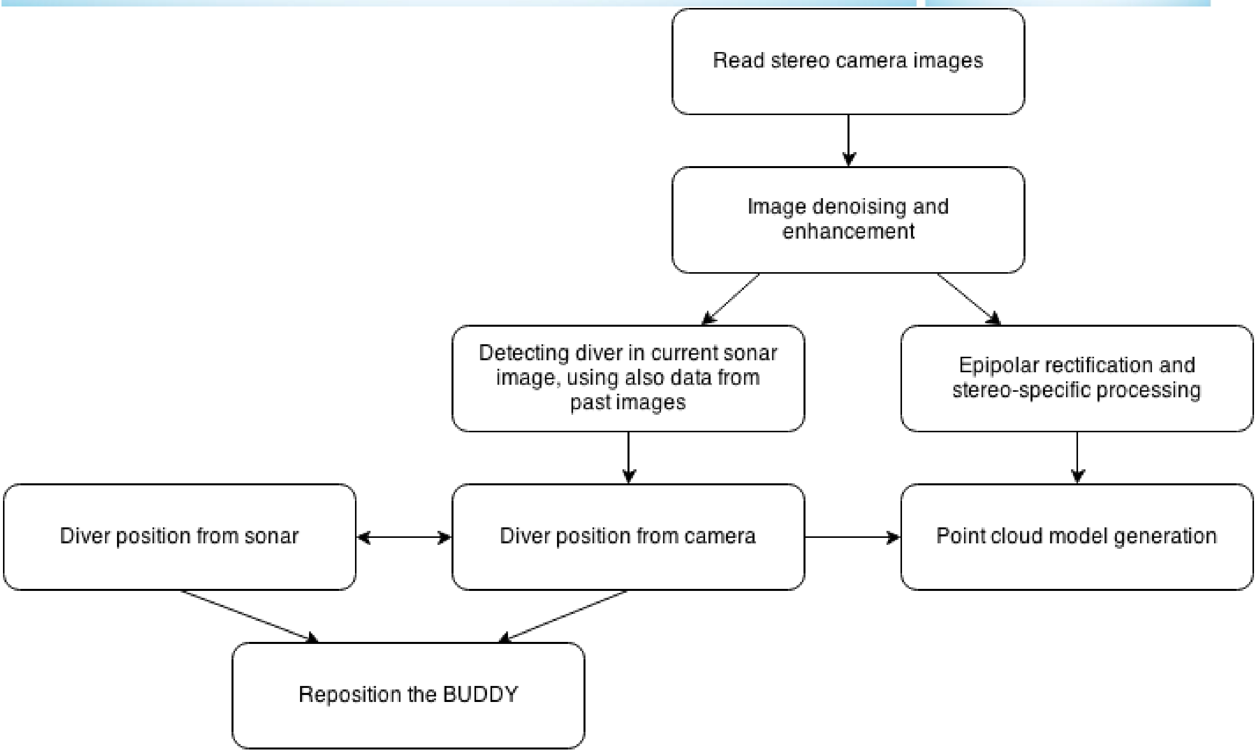


Fig. 14. Block diagram approximating stereo camera processing

5.3 Sonar

After reading the images from the sonar, a ROS node for image denoising and enhancement will attempt to filter the sonar image to preserve only relevant data. After that, a node for diver detection will attempt to locate the diver in the image. This will be done using both the current image and the knowledge from previous ones, most likely combined with a Kalman filter for object tracking and by detecting a region of interest - a part of the image where the diver's position is most likely. Diver position from camera (if available) will also be used.

Based on detected diver's position (from both sonar and camera) BUDDY will be repositioned to follow the diver at a prescribed range and allow an optimal view for both sonar and camera in the future. Depending on sonar's frame rate, some adjustment may have to be done to compensate for the BUDDY's movement, although sonar's frame rate will likely be high enough.

Also, some sonar parameters will be occasionally updated to allow a better view of the diver. This will require some processing to compensate for different ranges, sonar beam width etc.

Fig. 15 shows a block diagram of planned sonar image processing.

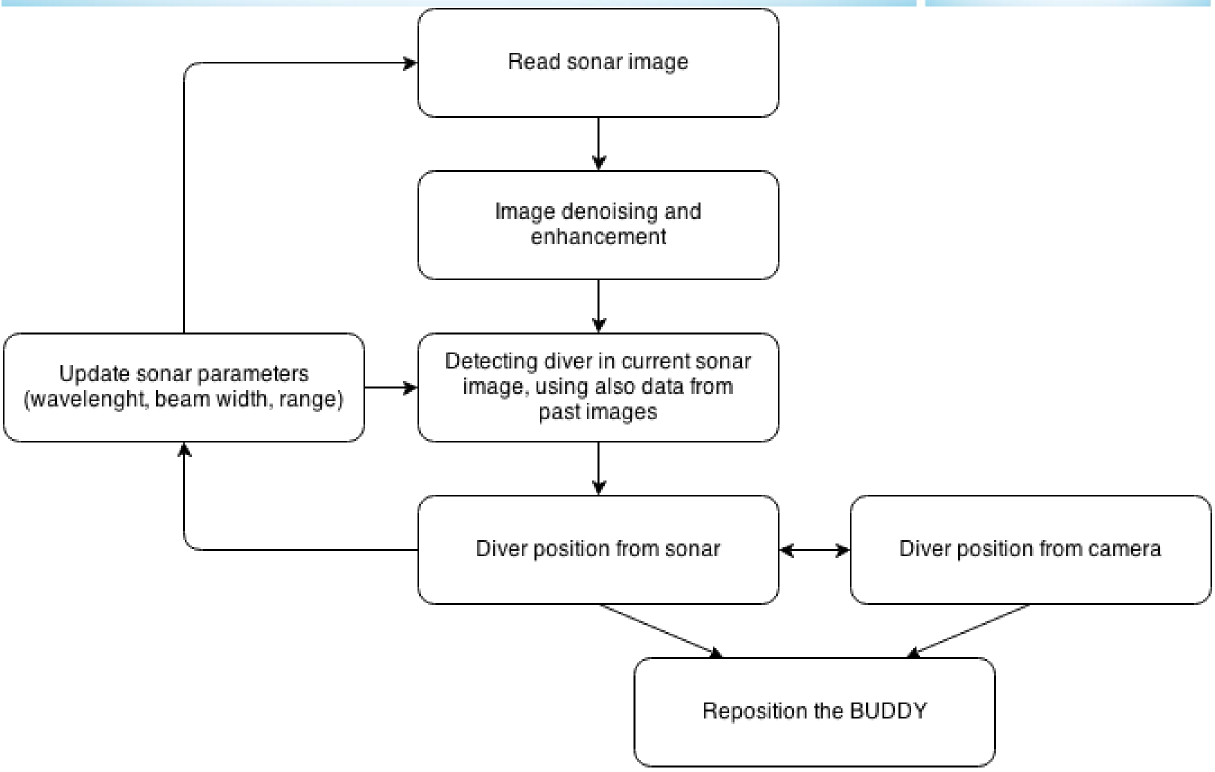


Fig. 15. Block diagram approximating planned sonar image processing.

6 Data acquisition

Initial experiments have been conducted May 13-15 in Caska on Island of Pag, Croatia, where a group of archaeology students from University of Zadar have been conducting underwater excavation and have agreed to participate in recording sensor data.

6.1 Equipment used and initial tests

DiverNet was used to track diver's motion, visualized with a 3D "stickman" model. This was the first time it was tested in real-life conditions, on a diver at sea. A temporary solution with Velcro stripes has been used to mount the sensors on diver. Although it did hold the sensor in places, mounting process was pretty long (approximately 30 minutes) so it was concluded that it is necessary to find alternative solutions.
Fig. 16 shows the DiverNet mounted with Velcro stripes. On the left image a stickman figure is visible in the background on the screen. Right image displays a diver diving with DiverNet.
Fig. 17 shows the virtually reconstructed diver model represented with a stickman figure.

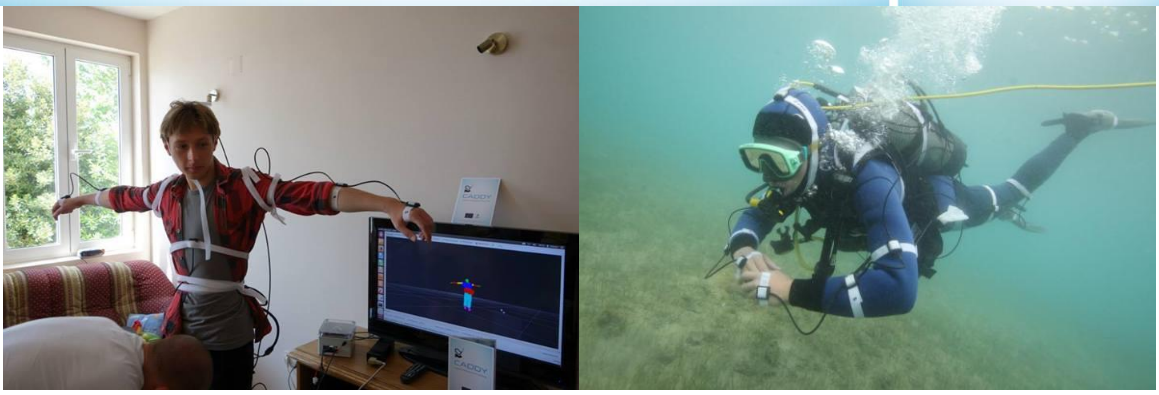


Fig. 16. Left: DiverNet mounted with stickman in the background; Right: diving with DiverNet.

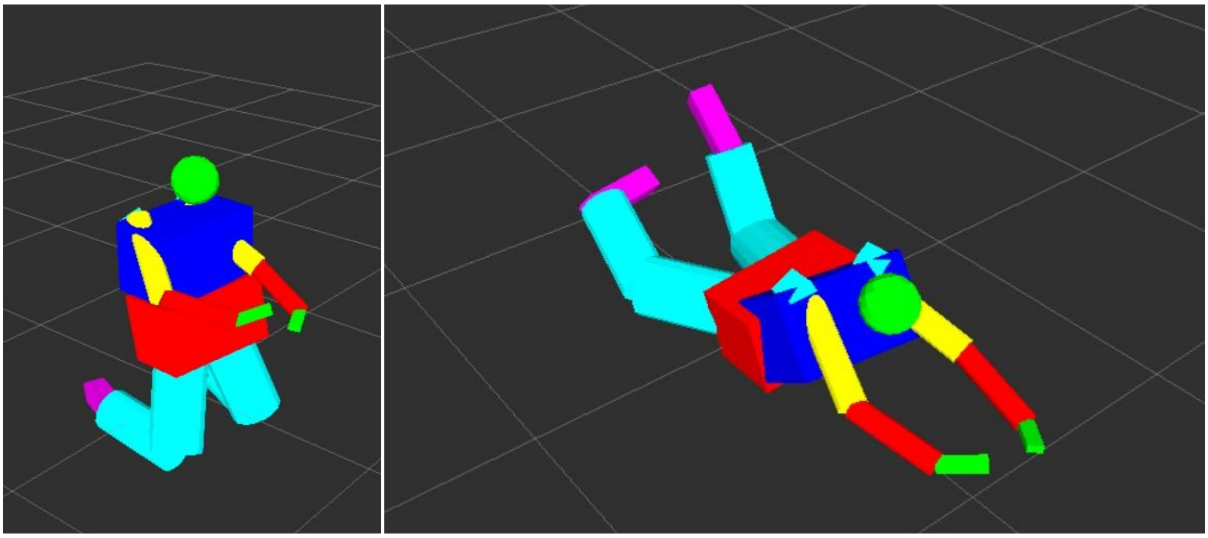


Fig. 17. DiverNet virtually reconstructed stickman

For the purpose of acquiring diver images stereo camera and multi-beam sonar have been used. They have been mounted together on a long pole in order to view the diver from approximately the same lateral angle.

Fig. 18 shows the described setup in the process of filming diver gestures.


Fig. 18. Sonar and stereo camera mounted, filming the diver.

During the first day equipment was tested by recording a single diver with all three devices simultaneously to check if everything was working and to find optimal locations and conditions for conducting experiments.
Sonar image is showing the diver approximately as seen from above when mounted facing the same direction as camera. This is just a consequence of the sonar's principle of operation.

6.2 Experiments conducted

During the second and third day a series of predefined experiments were conducted on a total of six divers.
Due to experiments for DiverNet being different from those for sonar and stereo camera, as well as the time needed for mounting the DiverNet, the experiments were separated and done in parallel. Each of the six divers did both the experiments.

6.2.1 Stereo camera and multi-beam sonar experiments

Divers were asked to perform a number of tasks while being recorded with stereo camera and multi-beam sonar. Some of the tasks were more focused on one of the two technologies, but both were recording during the whole time of the experiment. All the tasks were performed twice.

Diver approach

Data collected in this task will be used to detect the position and orientation of the diver.
Position at around 1 meter away, facing the camera, so that the torso and head are visible. This position will be used in later gesture experiments. Swim away and towards the camera and sonar several times.

Hand gestures

This data will be used to build a gesture recognition system.

1. **Numbers** - First display a 3-digit number (with agreed symbols for hundreds and tens), followed by slowly counting from 0 to 10. Counting was performed in two positions, one favouring the camera point of view and other favouring the sonar. Fig. 19 shows a video

frame of a diver showing number eight.

Fig. 19. Image from stereo camera, diver showing digits.

2. Diving symbols

a. Come closer.
b. Ear problem.
c. Give me light - hand pointing down in the face level and fingers opening like shedding light.
d. Low on air (reserve, 50 bar).
e. Out of air.
f. OK.
g. Something is wrong.
h. Stop.
i. Take me to the boat - two hands above the head forming a “V”.
j. You lead, I follow - right index points to you and goes down below, left index points to me and joins the left index.
k. I don't understand - shoulders going up with hands pointing palms up.
l. I feel dizzy.
m. Go away, you are too close - fingers stretching while palm facing down.
n. Stay there/Don't move - same as stop.
o. Going up.
p. Going down.
q. Which way? - right index points left and then right.
r. Watch me.

s. I'm cold.
t. Danger - pointing away and then forming a shark flipper on top of the head.
u. Slow down - both hands with open palms facing down going up/down.
v. Go faster - both hands, palms facing down, fingers stretching; like chasing away someone.

Fig. 20 shows the diver displaying "OK" symbol on stereo camera and on sonar.

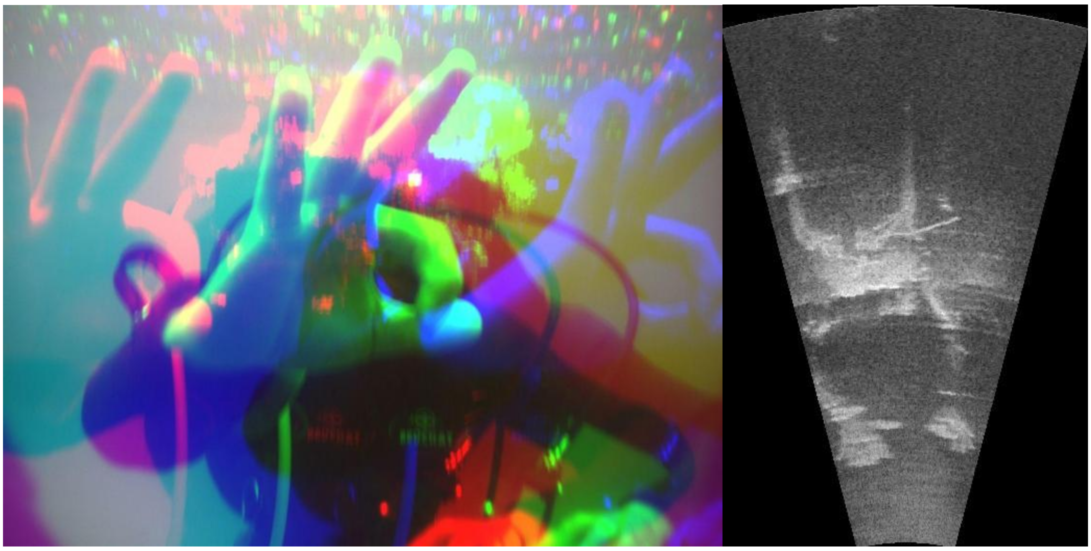


Fig. 20. Left: image from stereo camera, diver displaying "OK" symbol; Right: image from sonar, diver showing palm and "OK" symbol.

Casual diving

This task was done to capture the diver moving freely in directions of his choice to capture the entire body. Data will be used to train algorithms for diver detection and tracking from both sonar and camera image.

6.2.2 DiverNet

The following tasks were conducted with the DiverNet:

1. Take an upright position underwater with the arms stretched out laterally.
2. Swim 10 meters in a straight line RELAXED.
3. Swim 10 meters in a straight line FAST.
4. Swim in a circle RELAXED.
5. Swim in a circle FAST.
6. Swim backwards for 3 meters RELAXED.
7. Remain stationary in standard position for 5 seconds.

The main purpose of these tasks is in psychological studies of diver behaviour, but will also be used together with camera and sonar for tracking the diver.

7 Advantages and disadvantages

Some advantages and disadvantages of the planned diver sensing system were identified from both theoretical point of view and practical testing and data gathering conducted.

A network of inertial sensors, DiverNet, is an excellent tool for estimating human orientation. Similar systems are used and have been tested in different sports, movie and cartoon industry and many other. However, this approach also has many challenges and constraints. Although the filtering problem for a single sensor is well researched and decent solutions are available, it is very difficult to make the entire net represent the human body, estimate diver's anatomy and figure out if one of the sensors is not working as it should.
The latter occurred quite often during tests, mostly because of the difficulties with mounting the DiverNet and with sensor moving or falling off during the dive. It is expected that this can be greatly reduced, but not completely gone with a better mounting solution. Also, the magnetometer was greatly influenced by proximity of ferromagnetic metals in the diver's equipment, limiting options for sensor mounting spots.
Stereo camera should provide the possibility to create an extremely useful point cloud 3 dimensional model of the diver, which could also be combined with the DiverNet model. The biggest issues here is far inferior image quality underwater, making it necessary to develop excellent algorithms for diver detection, as well as for combining videos from different points of view. Also, the cylinder casing does produce a noticeable distortion. Because stereo processing and point cloud generation is a difficult and sensitive task, a new casing with flat surface (rectangular front side) will have to be designed to reduce the distortion.
Using the sonar image to detect diver is not very well researched and will be an interesting approach. The tests conducted are encouraging, with diver silhouette being fairly visible in most of the situations, as well as bubbles created when exhaling. The biggest challenge with sonar image is in beam targeting. If the sonar is not correctly oriented and set up for the situation (diver proximity and position) it could easily completely fail to capture the diver.

8 Conclusion

A description of the technologies that will be used for the purpose of sensing the diver was given. Based on the available parts - a stereo and mono camera, a multi-beam sonar and a network of inertial measurement units - it should be possible to create a robust system that will be able to track the diver and position the BUDDY with respect to the diver. Sonar image is limited, but specific algorithms - bubble detection from breathing, human contour detection in different poses (diving, facing the BUDDY and showing symbolic gestures) should give reasonable results and valuable data from sonar.

Underwater-specific image processing is not as explored as general image processing methods on the surface. Color ranges and contrasts are much different, light conditions unfavourable and images in general are of lower perception quality. Image processing algorithms should account for all that and try to extract as much information as possible to distinguish the diver from the rest of the underwater world.

Creating an accurate 3-dimensional model of the diver based on the stereo image and DiverNet sensors will be a very challenging task. The tests that were conducted, as well as theoretical research examined, have helped identify the biggest initial issues that will have to be addressed to make the model as accurate as possible. A very interesting approach might be in trying to fuse the information from the two systems, stereo camera (point cloud) and DiverNet diver model.

9 References

[1] Bosch FLEXIDOME IP startlight 7000 VR Data Sheet, http://resource.boschsecurity.com/documents/NIN_733_Data_sheet_enUS_10263325835.pdf, retrieved 24.05.2014.

[2] Woods, A.J., Docherty, T., Koch, R.: *"Image distortions in stereoscopic video systems"*, Proc. SPIE 1915.

[3] Oram, D.: *"Rectification for Any Epipolar Geometry"*, proc. BMVC 2008.

[4] Point Grey Bumblebee XB3, http://ww2.ptgrey.com/stereo-vision/bumblebee-xb3, retrieved 24.05.2014.

[5] Multibeam Sonar Theory of Operation, http://www.mbari.org/data/mbsystem/sonarfunction/SeaBeamMultibeamTheoryOperation.pdf, retrieved 21.05.2014.

[6] Sound Metrics ARIS Explorer 3000, http://www.soundmetrics.com/products/aris-sonars/aris-explorer-3000, retrieved 22.05.2014.

[7] Satar, J., Dudek, G.: *"Where is your dive buddy: tracking humans underwater using spatio-temporal features"*, IROS, 2007.

[8] Dudek, G. et all: *"A visually guided swimming robot"*, IROS, 2005.

[9] Beal, C., Lawrence, B.J., Ila, V., Dellaert, F.: *"3D reconstruction of underwater structures"*, IROS, 2010.

[10] Nelson, E.A. et all: *"Surface Reconstruction of Ancient Water Storage Systems: An Approach for Sparse 3D Sonar Scans and Fused Stereo Images"*, GRAPP, 2014.

[11] Schmidt, V.E., Rzhanov, Y.: *"Measurement of Micro-bathymetry with a GOPRO Underwater Stereo Camera Pair"*, IEEE Oceans, 2012.

[12] Wikipedia, *Diver Detection Sonar*, http://en.wikipedia.org/wiki/Diver_Detection_Sonar, retrieved 25.05.2014.

[13] Viola, P., Jones, M.: *"Rapid Object Detection using a Boosted Cascade of Simple Features"*, CVPR, 2001.

[14] Dalal, N., Triggs, B.: *"Histograms of oriented gradients for human detection"*, CVPR, 2005.

[15] Zhu, Q. et all: *"Fast human detection using a cascade of histograms of oriented gradients"*, CVPR, 2006.

[16] Schettini, R., Corchs, S.: *"Underwater image processing: State of the art of restoration and image enhancement methods"*, EURASIP Journal on Advances in Signal Processing, 2010.

[17] Kim, A., Golnaraghi, M.F.: *"A quaternion-based orientation estimation algorithm using an inertial measurement unit"*, Position Location and Navigation Symposium, 2004.

[18] Sabatini, A.M.: *“Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing“,* Biomedical Engineering, IEEE Transactions on (Volume 53, Issue 7), 1346-1356, 2006.

[19] Madgwick, S.O.H., Harrison, A.J.L., Vaidyanathan, R.: *“Estimation of IMU and MARG orientation using a gradient descent algorithm”*, Rehabilitation Robotics, 2011.

[20] Mahony, R., Hamel, T., Pflimlin, JM.: *“Complementary filter design on the special orthogonal group SO(3)”*, 44th IEEE Conference on Decision and Control, and the European Control Conference, 2005.

[21] Luinge, H.J., Veltink, P.H., Baten, C.T.M.: *“Ambulatory measurement of arm orientation“*, Journal of Biomechanics, Volume 40, Issue 1, 78-85, 2007.

[22] Favre, J. et all: *“Ambulatory measurement of 3D knee joint angle”*, Journal of Biomechanics, Volume 41, Issue 5, 1029-1035, 2008.

[23] Roetenberg, D., Luinge, H., Slycke, P.: *“Xsens MVN: Full 6DOF Human Motion Tracking Using Miniature Inertial Sensors“,* http://www.xsens.com/wp-content/uploads/2013/12/MVN_white_paper1.pdf, retreived 19.05.2014.

[24] Brodie, M., Walmsley, A., Page, W.: *“Fusion motion capture: a prototype system using inertial measurement units and GPS for the biomechanical analysis of ski racing“*, Sports Technology Volume 1, Issue 1, 17–28, 2008

[25] Mirzaei, F.M., Roumeliotis, S.I.: “A Kalman Filter-Based Algorithm for IMU-Camera Calibration: Observability Analysis and Performance Evaluation“, Robotics, IEEE Transactions on (Volume 24, Issue 5), 1143-1156, 2008.

[26] Treibitz, T.; Schechner, Y.Y.; Kunz, C.; Singh, H.: “*Flat Refractive Geometry*”. Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.34, no.1, pp.51,65, Jan. 2012.

[27] Martin Roser, Matthew Dunbabin, Andreas Geiger.: “Simultaneous Underwater Visibility Assessment, Enhancement and Improved Stereo”. Robotics and Automation, 2014. Proceedings of the IEEE International Conference on, May 2014.