EUROPEAN COMMISSION

Community Research

Deep-Learning for Multimodal Sensor Fusion

D2.2 Sensor Concept

Version 1.0

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 101016958

Deliverable number:	D2.2
Due date:	30.4.2021
Nature:	Document
Dissemination Level:	Public
Work Package	WP2
Lead Beneficiary:	KRA
Contributing Beneficiaries	DFKI, UDG, UH

Document History

Version	Date	Author	Contributors	Reaon for Change / RID #
V0.1	01/02/2021	Tom Runge (DFKI)		Initial template
V1.0	26/04/2021	Jakob Schwendner (KRA)	UDG, UH, DFKI, TA	Input to V1.0

Document Approval Sheet

Version	Reviewed and approved by	Organization	Date
V1.0	Thomas Vögele (PC)	DFKI	30/04/21

Project Coordinator

Organization	DFKI Robotics Innovation Center
Responsible Person	Dr. Thomas Vögele
Address	Robert-Hooke Str. 1
Phone	+49 17845 4130
e-mail	thomas.voegele@dfki.de

Consortium

Participant name	Short name	Country
Deutsches Forschungszentrum für Künstliche Intelligenz GmbH	DFKI	Germany
Universitat de Girona	UdG	Spain
University of Haifa	UH	Israel
Kraken Robotik GmbH	KRA	Germany
Bundesministerium des Inneren	THW	Germany
Israel Nature and National Parks Protection Authority	INPA	Israel
Tecno Ambiente SL	TA	Spain

Copyright

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

1 EXECUTIVE SUMMARY

This document collects information on relevant sensors and sensor pairing options in the context of the DeeperSense project. The goal is to identify and specify sensor candidates for each of the inter sensoric use-cases within the project.

The list of potential sensor candidates is restricted to sensors that are both suitable and available for the respective application within the project. It is not meant to be a generic list for sensor pairing possibilities.

Based on the available sensors and their specifications, a set of different pairing options is proposed and evaluated according to the requirements defined in "D2.1 Use-Case Requirements". The pairings include subsets as well as configuration options. For example, the same set of sensors could be used in separate pairings when the have a different geometric configuration.

This document also intends to be used as a reference in the integration phase of the project, where consortium members can quickly find key features of each sensor, such as mechanical and electrical characteristics as well as the nature and the volume of output data.

The proposed sensors and sensor pairings are part of a conceptual design. Therefore, sensors and sensors combinations can still be added or removed during the project development depending on their level of contribution to the results.

2 USE CASE I: HYBRID AUV FOR DIVER SAFETY MONITORING

2.1 SENSORS

2.1.1 Seavision

2.1.1.1 Description

SeaVision is the world's first RGB underwater laser imaging system that offers the resolution, range, and scan rate to deliver dense full color 3D point cloud images of subsea infrastructure with millimeter accuracy. SeaVision uses a full color laser scanning process repeated thousands of times per second to generate coordinate values of millions of points on a reflected surface. The coordinates and intensity associated with each reflected laser pulse are processed in real time to generate a high-resolution point cloud. Unlike other underwater laser scanning systems. Figure 2.1 shows on the left a 3D point cloud with millimetric accuracy which can be used for metrology, in the right a RGB colored 3D reconstruction of a mooring chain.

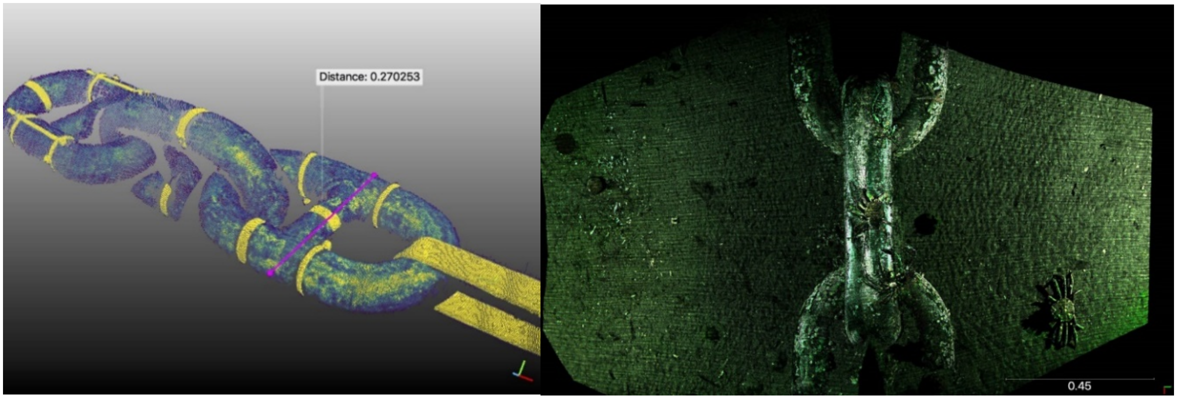
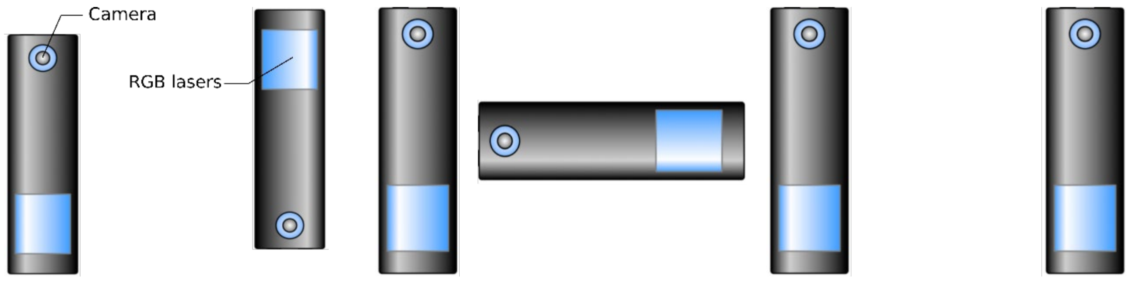


Figure 2.1: SeaVision Data Output

SeaVision does not have any externally moving parts. It is integrated in a compact twin pod configuration (Figure 2.2). Each SeaVision pod is composed of a red, green, and blue laser on a steerable unit with a color camera and LED light. Internally, each pod has embedded electronics for on-board processing, embedded micro inertial measurement unit to assist motion estimation algorithms and solid-state hard drive for on-board data storage. SeaVision requires a twin pod configuration to operate.

Figure 2.2: SeaVision Twin System

The SeaVision sensor can be used as a 3D laser scanner but is also suitable as a color camera (both mono and stereo). Typical Seavision mounting is achieved by rotating one unit by 180 degrees (Figure 2.3.a). This configuration allows to use the Seavision PODs as two laser scanners and increases the scan volume of the system. The SeaVision mounting flexibility can be exploited to create different mounting schemes to favor particular aspects, such as an increased overlap between both cameras to create a stereo system. Figure 2.3. shows examples of possible mountings:



(a) (b) (c)

Figure 2.3: Possible SeaVision Twin mounting configuration
(a) classical mounting with best field of view. (b) cross mounted for horizontal and vertical scanning. (c) mounting optimized for stereo-vision.

Seavision lasers are Class 3R (IEC/EN 60825-1:2014) with an optical output power of 40, 50 and 75 mW spread over the respective lines. For the lasers to be used in the vicinity of divers, protective measures have to be taken.

2.1.1.2 Sensor Characteristics

Table 2.1 Physical Characteristics

Specification	Description
Dimension	0.11 m x 0.11 m x 0.68 m (each pod)
Power	20W per tube
Input Voltage	24 VDC
Depth Rating	1000 m
Weight in air	13.6 kg (6.8 Kg each pod)
Weight in water	4.72 kg (2.35 Kg each pod)
Max Water Temperature	40°C
Minimum start-up temperature	-5°C

Table 2.2 Performance Characteristics

Specifications	Description
Measuring Principle	Active triangulation
Light Emitter	2x Red, Green, Blue Laser Diodes
Light Pattern	Line
Pixel Sensor Resolution	2x 1280 x 960

Depth Resolution (mean observed)	0.2 mm @ 1 m 0.8 mm @ 2m 3.0 mm @ 4m
Field of View in water	79° x 56° @ 0.5 m-6 m 107° x 56° @ 1 m-6 m
Scan Angle	0-130°
Working range	1.0-8 m
Frame Rate	2x125 FPS
Maximal Profiles / sec	250 profiles/s
Maximal Images / sec	125 images/s
Interface	GigE
Time Sync.	NTP, Hardware Sync Line
Protocol	gRPC
Scan Speed	5 to 11 degrees/sec
Latency Point Cloud Processing	< 50 ms
Video Streaming	Yes, full resolution for both cameras
Internal HDD capacity	2 x 1TB
Point Cloud Data format	LAS, LAZ and PLY
Point Cloud Visualization	Open source programs such as CloudCompare or MeshLab
Image Data Format	JPEG, BMP
SeaVision processing	Embedded or Post-processed

Figure 2.4 shows Seavision working area, based on the opening angles of laser and cameras.

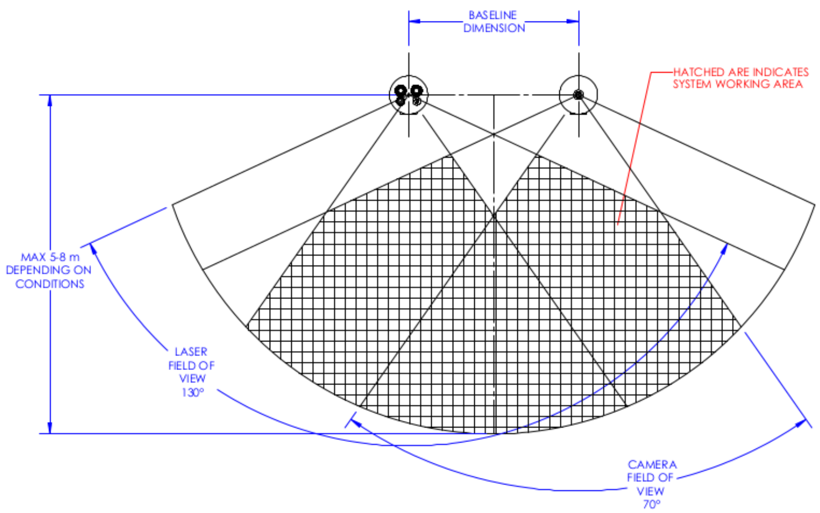


Figure 2.4: SeaVision working area.

2.1.2 Multibeam Imaging Sonar

2.1.2.1 Description

A multibeam imaging sonar is a sensor that produces video-like acoustic imagery by emitting sound pulses and measuring the intensity of the reflected signals. By using an acoustic lens, multiple beams can be formed and emitted simultaneously and thus covering a larger view field as shown in Figure 2.5. All points lying vertically on the same radius from a single beam are measured as one reflection. To generate a 2D image, the sonar maps the intensity of the reflected waves to the azimuth angle and distance between the sensor and the reflecting object. A simple example of a generated image using one beam is shown in Figure 2.6.

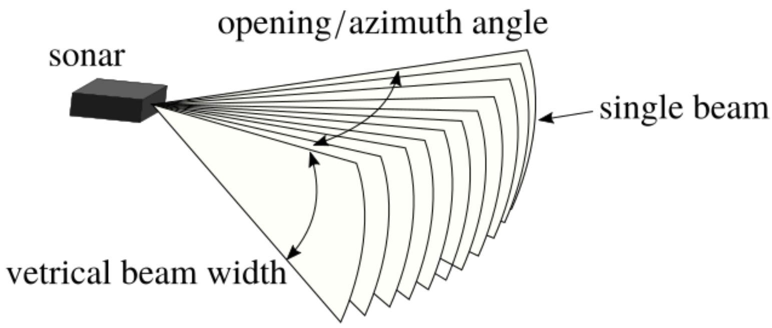


Figure 2.5: Schematic showing the acoustic beams emitted by a multibeam sonar

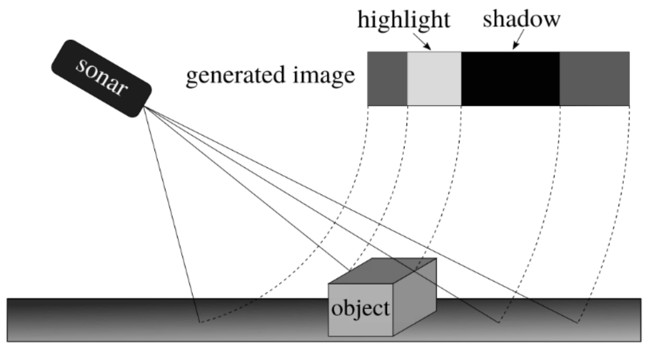


Figure 2.6: Example of a generated image using one beam

Unlike visual sensors like cameras, imaging sonars among other types of acoustic sensors can operate in very low visibility conditions such as turbid or cloudy waters or even in complete darkness. Additionally, traditional cameras lack the depth information that is inherent to sonars. The performance of imaging sonars and the clarity of the produced images relies on several factors such as the operating frequency, the width of the acoustic beam, the scanning field of view (opening angle), and the angular and range resolution. In general, a lower operating frequency increases the range of the sensor, however sacrificing the resolution of the captured image. Higher operating frequencies and a narrower beam width provide a clearer and more detailed images.

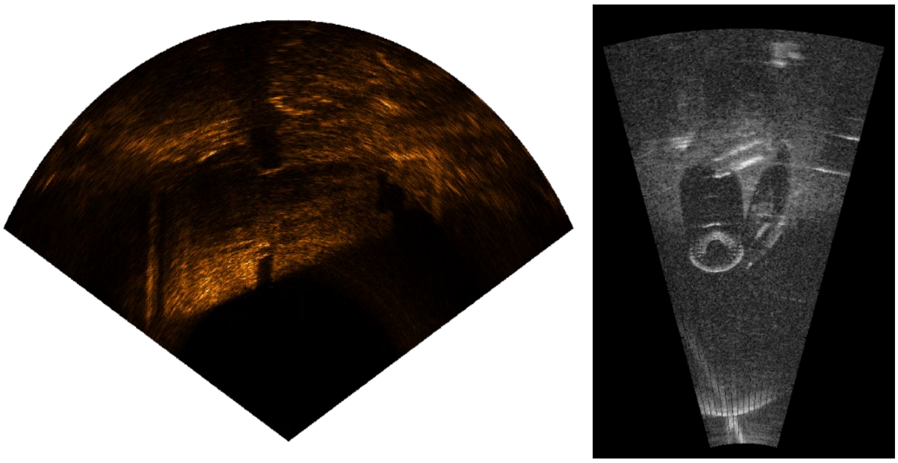


Figure 2.7: Examples of images produced by a multibeam imaging sonar

Imaging sonars can be operated from a boat, mounted to a fixed tripod, hand-held by diver or mounted onto underwater vehicles such as autonomous (AUVs) or remotely operated underwater vehicles (ROVs).

In the context of use case 1, DFKI will use two imaging sonars, the Gemini 720i by Tritec, and the BlueView P900 by Teledyne.

Figure 2.8: Multibeam imaging sonars. Left: BlueView P900, right: Gemini 720i

2.1.2.2 Sensor Characteristics

Table 2.3: Gemini 720i – Physical characteristics

Specification	Description
Dimensions	0.11 m x 0.135 m x 0.228 m
Power	35W maximum (depending on range)
Input Voltage	20 -75 VDC
Depth Rating	300 m
Weight in air	3.9 kg
Weight in water	1.2 kg
Max Water Temperature	35°C
Minimum start-up temperature	-10°C

Table 2.4: Gemini 720i – Performance characteristics

Specifications	Description

Operating Frequency	720 KHz
Angular Resolution	1.0° acoustic, 0.5° effective
Transducer Angle	10° downward tilt
Opening angle (Field-of-view)	120°
Number of Beams	256
Vertical Beamwidth	20°
Range	0.2m to 120m
Scan Rate	5-30 Hz (depending on range)
Range Resolution	8mm (depending on range)

Table 2.5: BlueView P900-45 – Physical characteristics

Specification	Description
Dimensions	0.127 m x 0.127 m x 0.287 m
Power	9.5W maximum (depending on range)
Input Voltage	12 -48 VDC
Depth Rating	1000 m
Weight in air	2.4 kg
Weight in water	0.6 kg

Table 2.6: BlueView P900-45 – Performance characteristics

Specifications	Description
Operating Frequency	900 KHz
Angular Resolution	1.0°
Opening angle (Field-of-view)	45°
Number of Beams	256
Vertical Beamwidth	20°
Range	2m to 60m
Scan Rate	Up to 15Hz (depending on range)
Range Resolution	25mm (depending on range)

2.2 PAIRING POSSIBILITIES

Table 2.7: Sensor Pairing Use case 1

Specifications	Pairing Options		
	1	2	3
3D Laser Scanner			x
Mono Camera	x		
Stereo Camera		x	
Imaging Sonar	x	x	x

2.2.1 Pairing 1: Optical Mono Camera – Imaging Sonar

The first sensor pairing possibility for Usecase I is that of an optical monocular camera and an imaging sonar. A mono-camera produces a 2-dimensional image where 3D points are mapped into 2D image plane, thus losing the depth information. On the other hand, the imaging sonar provides a different field of view where the depth and swath coordinates are observed while the elevation gets projected on the image plane. By placing both sensors in the right configuration, the field of view of both sensors could be aligned to observe simultaneously the same scene. In the DeepSenser context, we will investigate two different camera-sonar positioning by creating a single rig and placing two cameras on it. Figure 2.9 shows two different setups.

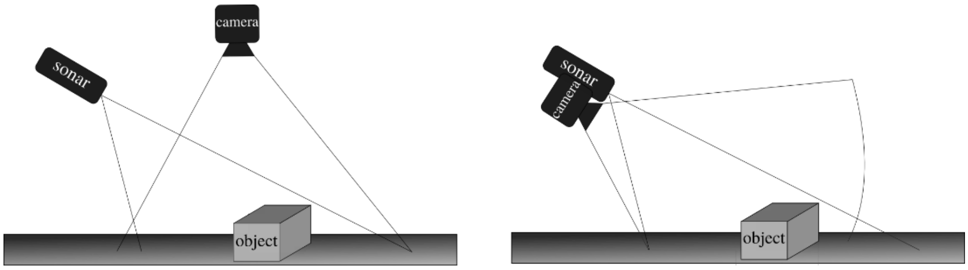


Figure 2.9: Schematic depicting possible positions of the imaging sonar and camera

By having the view fields of both sensors aligned, image transformation techniques similar to [Wang and Gupta, 2016], [Isola et al., 2017] and [Frans, 2017] could be applied to retain visual-like images from acoustic imaging provided by the sonar. The figure below shows example of a car tire where the view fields of sonar and camera are aligned.

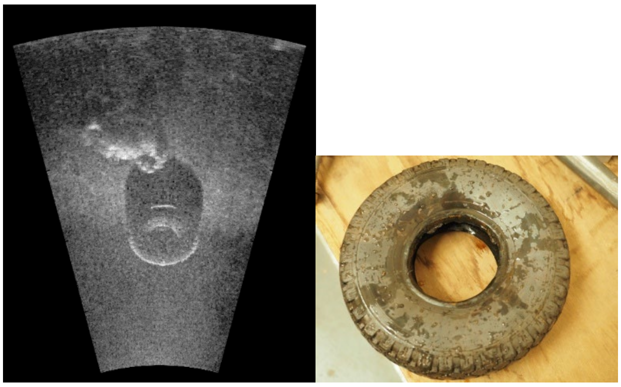


Figure 2.10: An example of a camera and sonar images with an aligned view field source: [Valdenegro-Toro, 2019], DFKI

2.2.2 Pairing 2 and 3: Stereo camera/ 3D Laser Scanner – Imaging Sonar

The second and third sensor pairing aims to leverage 3D information from optical sensors and combine it with the imaging sonar. The modeling and output format of 3D laser scanners (structured light) and stereo camera systems are quite similar since the laser projector is commonly considered as an inverse camera and depth information is retrieved by triangulation. The differences between them reside on accuracy, computational cost, field of view and color sensitivity. For training the network both systems will contribute with a depth image of the scene in addition to the color image of the camera.

This depth knowledge is not directly captured with a mono camera, therefore we expect that we can improve our training with that. Additionally, due to the nature of sonar data acquisition, explained in section 2.1.2.1, multiple objects lying vertically are seen by the sensor as a single reflection. We expect pairing 2 and 3 to be complementary since the optical sensors can provide the information that would be otherwise lost. Here the position between the sensors is as important as it is on the pairing option 1, therefore we will also investigate both positions depicted in Figure 2.9.

As a characteristic of the 3D laser scanner, when collecting the data set the scene needs to be static, otherwise data have to be motion compensated. While this is not suitable for operations, it can be a good input data for training the network and provide the information required to retrieve image having only sonar.

3 USE CASE II: SURVEYING AND MONITORING COMPLEX BENTHIC ENVIRONMENTS

3.1 SENSORS

3.1.1 Forward Looking Sonar (Multibeam Imaging Sonar)

3.1.1.1 Description

Same as described in section 2.1.1.1.

3.1.1.2 Sensor Characteristics

We are going to use two different Sonars that mainly differ in their operating frequency (900Khz vs. 2100 Khz). As explained in Section 2.1.1.1 this results in different resolutions and ranges. Lower frequency enables a longer range (60m vs. 30m) but provides lower resolution (examples in Figure 3.1).

Table 3.1: BlueView M900- 130: mounted on the AUV (lower resolution) – Physical characteristics

Specification	Description
Dimensions	0.1016 m x 0.1016 m x 0.192 m
Power	20W maximum (depending on range)
Input Voltage	12 -48 VDC
Depth Rating	1000 m
Weight in air	1.84 kg
Weight in water	0.39 kg

Table 3.2: BlueView M900- 130: (lower resolution) – Performance characteristics

Specifications	Description
Operating Frequency	900 KHz
Angular Resolution	0.18°
Opening angle (Field-of-view)	130° V
Number of Beams	768
Vertical Beamwidth	20°
Range	2m to 60m
Scan Rate	Up to 25Hz (depending on range)
Range Resolution	25mm (depending on range)

Table 3.3: Blueprint M1200d: on the blueROV (higher frequency) – Physical characteristics

Specification	Description
Dimensions	0.125 m x 0.122 m x 0.062 m
Power	35W maximum (depending on range)
Input Voltage	18 - 32 VDC
Depth Rating	300 m
Weight in air	0.98 kg
Weight in water	0.36 kg

Table 3.4: Blueprint M1200d: on the blueROV (higher frequency) – Performance characteristics

Specifications	Description
Operating Frequency	1200 KHz (LF) / 2100 KHz (HF)
Angular Resolution v	0.6° (LF) / 0.4° (HF)
Opening angle (Field-of-view)	130° V (LF) / 60° (HF)
Number of Beams	512
Vertical Beamwidth	20° (LF) / 12° (HF)
Range	0.1m to 30m (LF) / 0.1m to 10m (HF)
Scan Rate	Up to 40Hz (depending on range)
Range Resolution	2.5mm (depending on range)

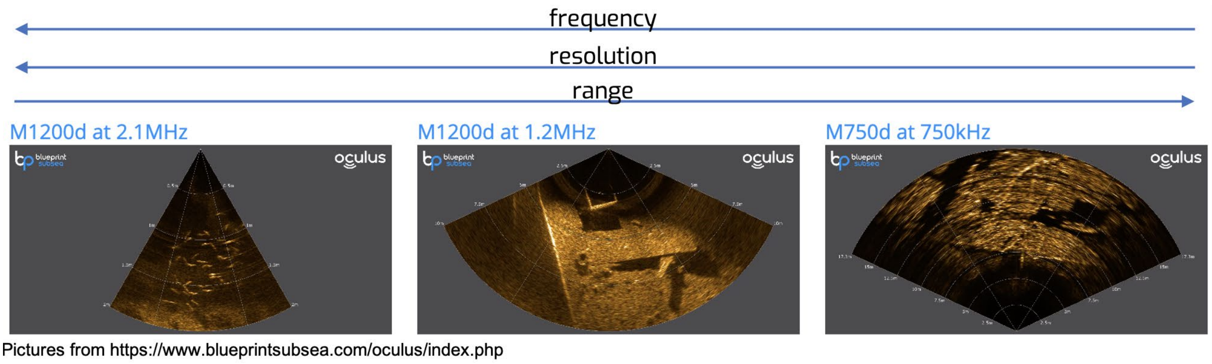


Figure 3.1: Figure 3.1: An example of images acquired by sonars with different frequencies. Source: https://www.blueprintsubsea.com/oculus/

3.1.2 Camera

The optical camera provides high-resolution on the expense of the range dimension. In addition, it has limited visibility underwater.

3.1.2.1 Description

An optical camera acquires a 2D image of the scene, while operates with limited range. Underwater, visibility is limited because of the optical properties of the water. The water attenuates light as it propagates through it and also scatters it. Both attenuation and scattering depend on the wavelength of the light and change both temporally and between water bodies (see Figure 3.2). Therefore, we do not expect to have more than 10m visibility in coastal waters.

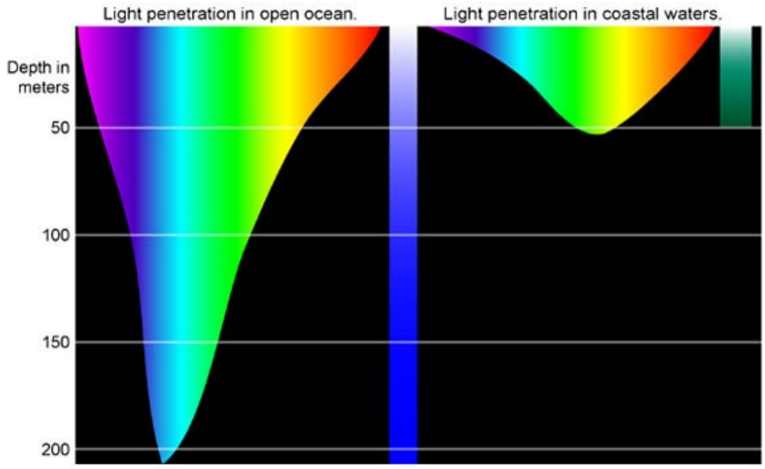


Figure 3.2: Light attenuation underwater.
Source: http://oceanexplorer.noaa.gov/explorations/04deepscope/background/deeplight/media/diagram3.html

3.1.2.2 Sensor Characteristics

Table 3.5: Prosilica GT 6600: on the AUV – Physical characteristics

Specification	Description
Dimensions	0.096 m x 0.066 m x0.053 m
Power	8.1W maximum
Input Voltage	7 - 25 VDC
Weight in air	0.372 kg

Table 3.6: Prosilica GT 6600: on the AUV – Performance characteristics

Specifications	Description
Resolution	6576 (H) x 4384 (V)
Sensor	ON Semi KAI-29050
Sensor type	CCD Progressive
Sensor size	Type 35 mm
Pixel size	5.5 µm × 5.5 µm
Max. frame rate at full resolution	4 fps
ADC	14 bit

Table 3.7: Prosilica GT 6600 Attached lens: AF Nikkor 24mm f/2.8D

Specification	Description
Dimensions	0.0645 m X 0.046 m
Weight in air	0.27 kg
Focal Length	24mm
Aperture	f/ 2.8 - f/ 22
Format	FX/35 mm
Max reproduction ratio	0.11x
Min focus distance	0.304 m

Table 3.8: IDS UI-3260CP Rev.2: on the blueROV – Physical characteristics

Specification	Description
Dimensions	0.029 m x 0.029 m x0.029 m
Power	3.2 W maximum
Input Voltage	12-24 VDC
Weight in air	0.052 kg

Table 3.9: IDS UI-3260CP Rev.2: on the blueROV – Performance characteristics

Specifications	Description
Resolution	1936 (H) x 1216 (V)
Sensor	Sony IMX249LQJ-C
Sensor type	CMOS
Sensor size	1/1.2''
Pixel size	5.86 µm
Max. frame rate at full resolution	47 fps
ADC	12 bit

Table 3.10: IDS UI-3260CP Rev.2 Attached lens: Tamron M112FM12

Specification	Description
Dimensions	0.0406 m X 0.0385m
Weight in air	0.064 kg
Focal Length	12mm
Aperture	f/ 2 - f/ 16
Min focus distance	0.1 m

3.2 Pairing Possibilities

Table 3.11: Sensor pairing Use case II

Specifications	Pairing Options	
	1	2
Mono Camera	x	x
Low res Imaging Sonar		x
High res Imaging Sonar	x	

Forward-Looking Sonars (FLS) and Forward-Looking Cameras (FLC) are used in conjunction in many under water platforms because of their complementary abilities. The FLS has a long range and its performance does not depend on water conditions. However, it produces low resolution images that lack vertical information and its input is unstable in short ranges due to reverberations. On the other hand, the FLC has excellent spatial resolution with color information but only works in short ranges. In our use case the two sensors will be co-located as they mounted on a vehicle (see Figure 3.3). Because of the FLS mode of operation, this yields two very different viewpoints (illustrated in Figure 3.3). The FLS displays a top view 2D acoustic image of the scene, whereas the FLC provides a 2D side-view optical image, despite being located in proximity. The sensors can be calibrated before hand, aiding the coordinate transfer between the images. The two options consist of the same configuration, but with different models of the FLS, with different frequencies.

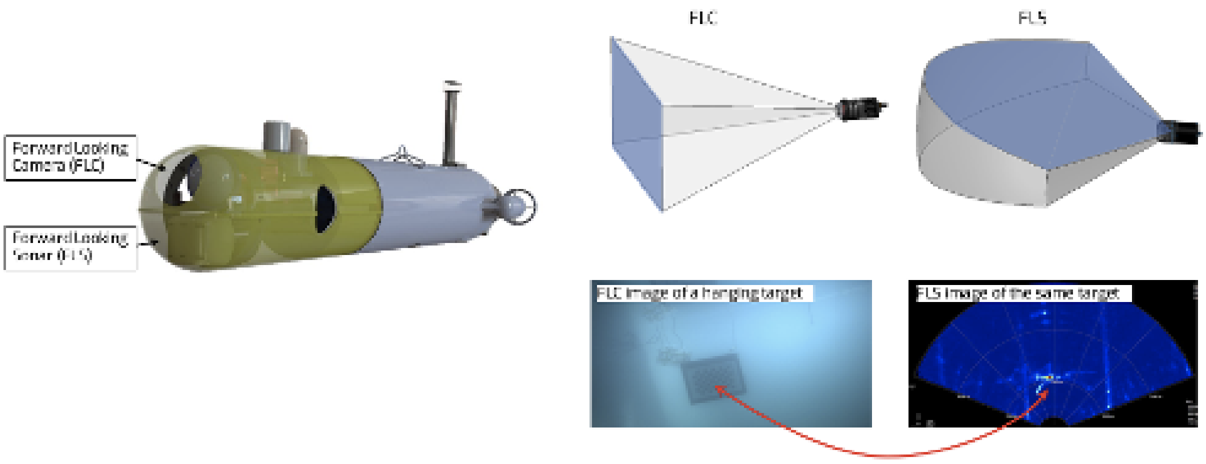


Figure 3.3: Sensor pairing configuration.

4 USE CASE III: SUB-SEA MULTISENSOR BOTTOM MAPPING AND INTERPRETATION FOR GEOPHYSICS

4.1 SENSORS

4.1.1 Side Scan Sonar

4.1.1.1 Description

Towed side scan sonar surveys use a cylindrical vehicle with hydrodynamic design provided with fins, towed behind the stern of the boat. Two transducers send acoustic signals across the water in a determinate frequency, each one. The side scan sonar can work at different frequency ranges: systems working in high frequencies, between 500 kHz and 900kHz, offer high resolution but low ranges, and systems working in low frequencies i.e. 100 kHz, offer low resolution but higher ranges.

The following figures explain how beam patterns in an SSS is configured.

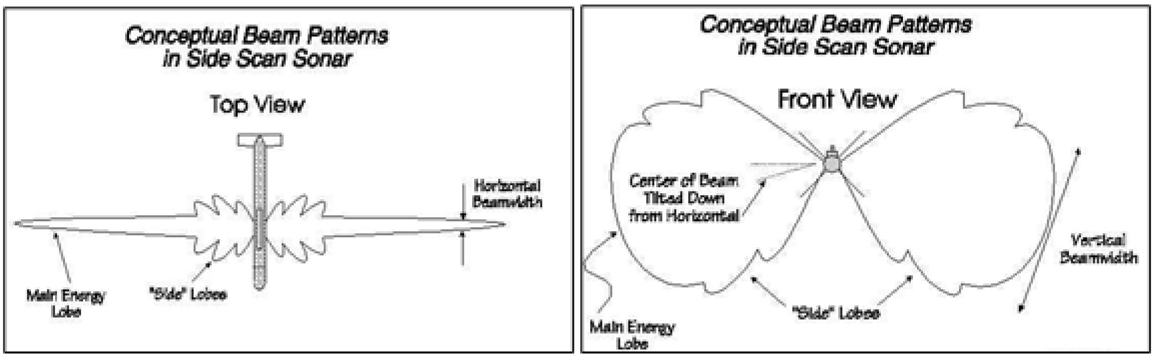


Figure 4.1: Beam patterns scheme of an SSS.

The reflection of this signal, coming from the bottom, is acquired by the same transducers. With side scan sonar it is possible to identify sea bed morphology and configuration, as well as its nature, identifying different substrates of hard bottoms (rocky or consolidated), soft or sedimentary bottoms and sea weeds.

In order to have a full coverage of the seabottom, the side scan sonar has to be towed following a project lines. This project line will depend on the frequency, range and overlap chosen.

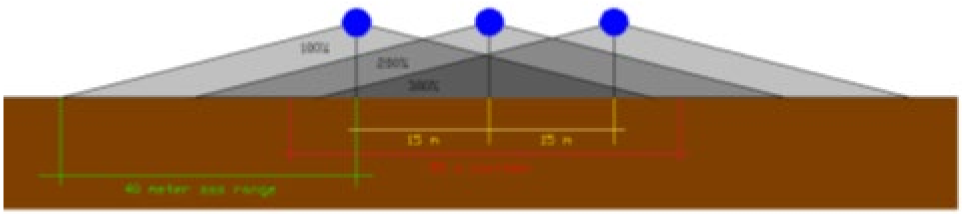


Figure 4.2: Side scan sonar survey overlap example.

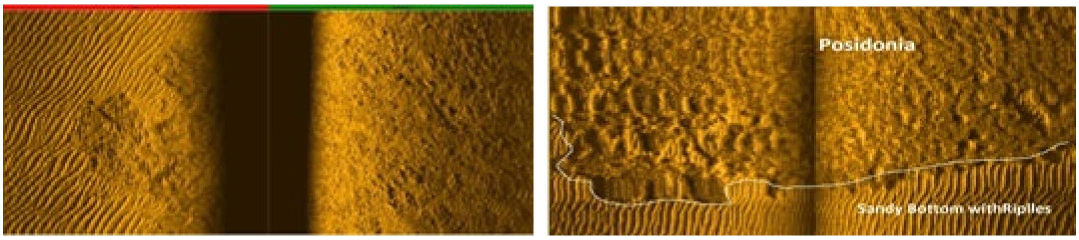


Figure 4.3: Left, side scan sonar record of cemented sands. Right, Posidonia field and sandy bottom with ripples.

Figure 4.4: Left, side scan sonar record of the Spicara Smaris nest. Right, the camera photo on the same area.

4.1.2 Klein 3000 towfish

4.1.2.1 Sensor Characteristics

The Klein System 3000 presents the latest technology in digital side scan sonar imaging. The simultaneous dual-frequency operation is based on new transducer designs, as well as the high-resolution circuitry recently developed for the Klein multi-beam focused sonar.

Figure 4.5: Klein 3000 side scan sonar.

Table 4.1 Towfish technical specifications.

System 3000 towfish	
Frequencies	100 kHz (132 kHz, ± 1% actual) 500 kHz (445 kHz, ± 1% actual)
Transmission Pulse	Tone burst, operator-selectable from 25 to 400 µsecs; independent pulse controls for each frequency
Beams	Horizontal: 0.7° @ 100 kHz 0.21° @ 500 kHz Vertical: 40°
Beam Tilt	5°, 10°, 15°, 20°, 25° down, adjustable
Range scales	15 setings – 25 to 1000 metres

Maximum range	600 m @ 100 kHz 150 m @ 500 kHz
Depth Rating	1500 m standard; other option available
Constrution	Stainless steel
Body lenght	122 cm
Body diameter	8.9 cm
Weight	29 kg in air
Standard Sensors	Roll, pitch and heading

Table 4.2: TPU technical specifications.

Transceiver Processor Unit (TPU)	
Operating system	VxWorks® with custom application
Basic hardware	Splash-Proof 2 (SP2) TPU
Outputs	100 Base-Tx, Ethernet LAN
Navigation Input	NMEA 0183
Power	120 watts @ 120/240 VAC, 50/60 Hz (includes towfish)
Interfacing	Interfaces to all major sonar data processors
Options	19-in rack mount TPU

4.1.3 Imagenex Ethernet kit

4.1.3.1 Sensor Characteristics

A smaller side scan sonar is also available in the form of a kit that can be AUV deployed. However due to the small transducer size, the performance characteristics are also considerably lower than the Klein 3000 towfish described above. The full specifications of the Imagenex SSS kit can be found in the following link: https://imagenex.com/products/ethernet-sidescan-kit

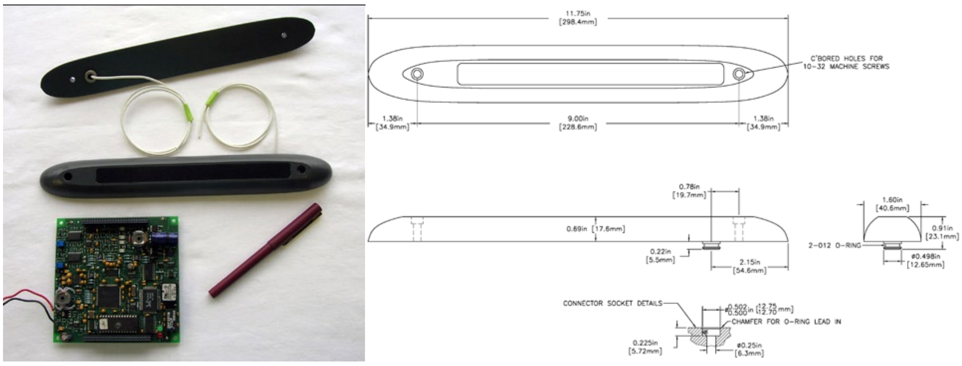


Figure 4.6: *Imagenex side scan sonar.*

Table 4.3: Imagenex side scan sonar

Imagenex Ethernet Side Scan Sonar	
Frequencies	260 kHz / 330 kHz / 800 kHz nominal
Transducer Beam Width	260 kHz: 2.2° x 75° 330 kHz: 1.8° x 60° 800 kHz: 0.7° x 30°
Range scales	Range scale ÷ 1000
Power Supply	Nominal 24 VDC (22 VDC – 33 VDC) at less than 2.5 Watts
Weight	506gr
Interface	Ethernet

4.1.4 MULTIBEAM ECHOSOUNDER (MBES).

4.1.4.1 Description

Hydrographic surveying has experienced a conceptual change in depth measurement technology and methodology from the first generation singlebeam echosounders, where major development and use in hydrographic surveying was in the mid-1900s, to the current generation of multibeam echosounders (MBES).

Bathymetric echosounder works by estimating the distances through the knowledge that sound propagates a given distance at a given time. The echosounder sends out a pulse of sound, a ping, and observes the echoes. Angles and distances are estimated for each received echo through beamforming algorithms, to determine the polar coordinates of each sea floor point. In practice, sound velocity varies through the water column (as it is affected by conductivity, pressure and temperature), so a layered sea model is used and the sound velocity in each layer is measured, and the ray path is determined.

By following the above procedure, large areas on the sea floor can be mapped. This is done by putting the information from many such succeeding pings together. If a vessel sends several pings while going forth (x direction), these pings will when put together make out a stripe of depth-measured area (z values) where the vessel is running. Several such stripes lying side by side will then constitute an entire bathymetric map.

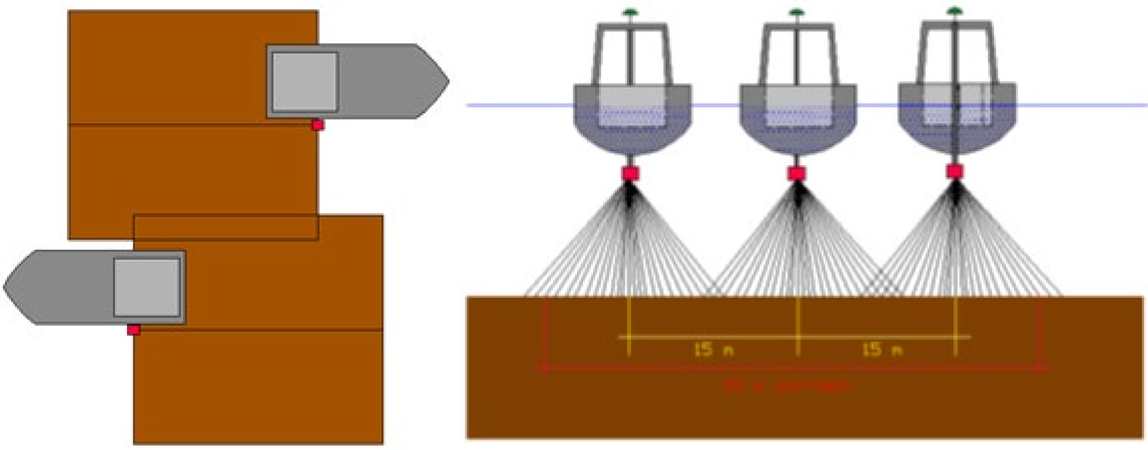


Figure 4.7:. Multibeam survey overlap example.

In order to have a corrected bathymetric data, the system has to be complemented with some other equipment: typically this will be a GPS (global positioning system) and a movement reference unit as a minimum.

4.1.4.2 Sensor Characteristics

The bathymetric echosounder selected for potential use in DeeperSense is the R2Sonic 2024 multibeam, This model stands out from other models in the market by features such as selectable operating frequencies and extended depth range.

Figure 4.8: R2Sonic 2024 multibeam echosounder.

Table 4-3 Multibeam echosounder technical specifications

Technical specifications	
Selectable frequencies	170-450 kHz. Optional 700 kHz
Minimun frequency increase	1 Hz
Beamwith, across track and along track	0.3° x 0.6° at 700kHz (optional) 0.45° x 0.9° at 450kHz 1° x 2° at 200kHz
Number of soundins	Up to 1024 soundings per ping
Max speed (vessel)	11.1 knots for full coverage
Near-field focusing	Yes
Roll stabilized beams	Yes
ROBO™ Automated Operation	Auto Power, pulse width, rangeTrac™, GateTrac™, SlopeTrac™
Saturation monitor	Yes
Selectable swath operation (also referred as max coverage)	10° to 160° User selectable in real-time
Sounding patterns	Equiangular Equidistant Single / double / quad modes Ultra High Density (UHD)
Sounding depth	Up to 400 m+
Pulse length	15µs - 1ms
Pulse Type	Shaped CW
Ping rate	Up to 60 kHz
Bandwidth	Up to 60kHz
Inmersion depth	100 m Optional 4000 and 6000 m
Bottom detect resolution	3 mm

Operating temperature	-10ºC to 50ºC
Storage temperature	-30ºC to 55ºC

Table 4-4. Electrical Interface specifications.

Electrical Interface	
Mains	90-260 VAC, 45-65 Hz
Power consumption	50W avg
Uplink/downlink	10/100/1000 Base-T-Ethernet
Sync in, Sync out	TTL
Deck cable length	15 m, optional 25 and 50 m

Table 4-5. Mechanical specifications.

Mechanical	
Receiver DIM (LWD)	480x109x190 mm
Receiver Mass	12.9 kg
Projector DIM (LWD)	273x108x86 mm
Projector Mass	3.3 kg
Sonar Interface Module	280x170x60 mm
Sonar Interface Module Mass	2.4 kg

4.1.5 Towed Camera

4.1.5.1 Description

Performing transects with georeferenced underwater video recording is a common technique used for the mapping of marine benthic communities or confirmation of the substrates detected by the side scan sonar. The quality of the recorded imagery is strongly dependent on the experience of the technicians conducting the survey and on the characteristics of the equipment.

The methodology of towed video transects consists of using a camera attached to a towable sled. This camera allows you to have real-time viewing on a monitor onboard the boat. The camera for better background viewing has lighting rings with adjustable intensity. This camera, lights and sled system will be towed by a cable that will make, along with the ship, as a tow agent and send the signal to the monitor on the boat.

The navigation height of the sled will depend on the speed of the boat, as well as the visibility in the waters under study, the most suitable being the one that provides a good image along with a wide field of view.

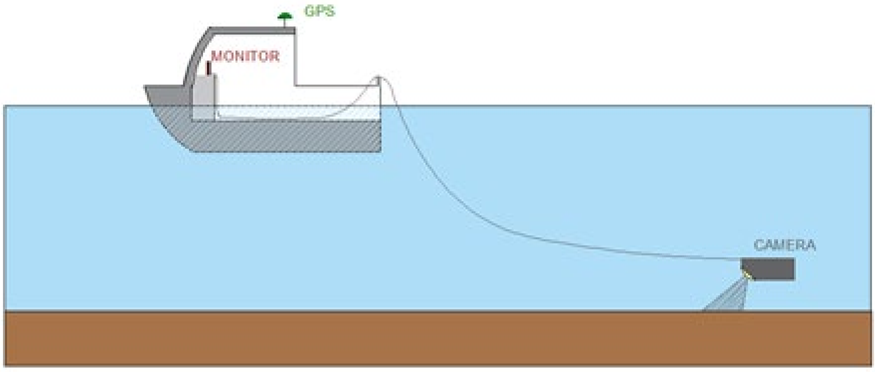


Figure 4.9: Towed camera deployment.

Figure 4.8 shows examples for images from the sea-bottom created by a towed camera.

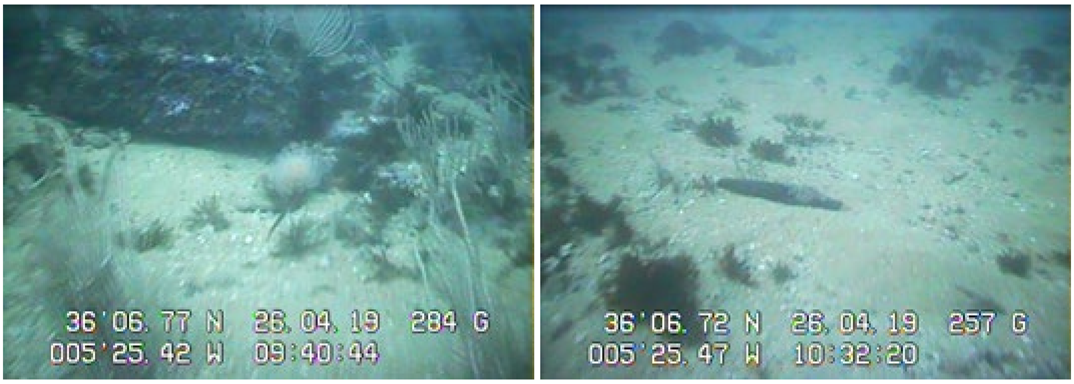


Figure 4.10: Seabottom images examples.

4.1.5.2 Sensor characteristics

The GNOM GTC-01 is a towed camera system with live video that can be deployed and towed by a small vessel. The towed camera system GTC-01 is ideal for small boat and platform operations, where a single operator or quick launch is required.

Figure 4.11: GNOM GTC-01 towed camera

Table 4.4: Technical specifications - Vehicle

Technical specifications - Vehicle	
Maximun operating depth	Up to 150 m
Dimensions	Length: 330 mm Width: 205 mm Height: 190 mm
Weight in air	2.5 kg
Operating temperature	-5ºC to 45ºC
Max cruising speed	Up to 1-2 knots
Vehicule protection	Crash frame assembly constructed of light weight and durable marine compatible polyethylene material

Table 4-7. Technical specifications – Towing Cable.

Technical specifications – Towing Cable	
Max tether length	400 m
Cable diameter	4 mm
Breaking strain	80 kg
Weight in air	3 kg/100m
Weight in fresh water	Neutrally bouyant
Type	Ultra-thin flexible coax cable, polyurethane jacketed internal Kevlar reinforcing (strain) member.

Table 4.5: Technical specifications - Camera System.

Technical specifications – Camera System	
Camera model	Sony Super HAD 2 CCD
Camera resolution	700 TV Lines
Effective Pixels	752(H) x 582(V), 437664 pixels
Image Sensor	1/3" Interline Transfer CCD
Mini Illumination	0.1 lux (0.01 - b/w camera)
Lens	3.6 mm / F 2.0
Iris control	Auto
Focus	Auto
Field of View (FOV)	66º
Camera Tilt	+/- 50º

Table 4.6: Technical specifications – Lighting System.

Technical specifications – Lighting System	
Light source	White ultra-bright LEDs
Qty of lights	2x

Power	2 x 3 Watt (6 W)
Luminous Flux	2 x 300 lumen (600 lm)
Beam Angle	105º each
Colour temperature	5600-6000º Kelvin
Control	Variable intensity

Table 4.7: Technical specifications – Navigation System.

Technical specifications – Navigation System	
Sensor	Compass and depth
On-screen overlay	Yes
Heading Accuracy	+/- 3º
Compass Resolution	0.5º
Depth Sensor Accuracy	1% F.S.

Table 4-11. Technical specifications – Surface Control Unit

Technical specifications – Surface Control Unit	
Power supply	Single Phase 100-240 VAC, 60-50 Hz.
Recommended Input Voltage	220 VAC
Maximun power output	100 Watt
Monitor	15“ LCD
Control panel system	Yes
ON-Screen Display	Yes
Navigation system	Wireless joystick
Integrated single board computer	Optional
System protection	Full system packed in two high-performance waterproof cases.

4.1.6 AUV Monocular Camera

4.1.6.1 Description

A monocular camera is available for use on the Girona 500 AUV. The Blackfly S GigE camera (BFS-PGE-27S5C-C) is a 12-bit camera with a 2.8Mp resolution (1936x1464) able to capture images at a speed of 43fps with its 2/3” Sony IMX429 CMOS sensor in global shutter. In addition to the base hardware, the camera capabilities can be expanded with specific C Lenses, polarization filters, and the Spinnaker SDK.

Figure 4.12: FLIR Blackfly S GigE camera (BFS-PGE-27S5C-C)

The lens chosen to be fitted with the camera is the GMTHR36014MCN of Goyo optics, a 6mm single focal length lens with a f number between F1.4-16 and a dimension of Ø30 x 32.8(mm)

Table 4.8: Technical specifications – Camera BFS-PGE-27S5C-C

Specifications	Description
Pixel Sensor Resolution / Size	2.8Mp (1936x1464) / 4.5 µm
Sensor	Sony IMX429, CMOS, 2/3"
Lens Mount	C-mount
ADC (Analog to Digital Converter)	12 bit
Gain Range	0 db to 47 dB
Exposure Range	15 µs to 30 seconds
Frame Rate	1-43 FPS
Image Buffer	240Mb
Flash Memory	6Mb non-Volatile
Interface	1G PoE
Time Sync.	IEEE 1588 PTP

4.1.8 Stereo Camera

4.1.8.1 Description

The camera stereo system is built in a cylindrical housing made from hard-anodized aluminum alloy and with two rectangular viewports made of highly transparent polymethyl methacrylate (PMMA). The cylinder contains two Canon EOS 5D Mark II still cameras, with 21MPixel sensors and Canon 24mm lenses. The cameras are connected to a PC-104 computer stack which can store and post-process the images if necessary. The computer is also in charge of controlling and logging data from an echo sounder whose transducer is mounted between the two viewports (Figure 4.11, right)

Figure 4.13: Camera Stereo schematic

The echosounder ranging assists the focusing of the cameras. This mechanism is helpful under mild image turbidity, as the cameras are not required to do optical based focusing. The housing has several connectors which make it possible to interface with the Girona 500 AUV, to connect with auxiliary systems such as an external multibeam echosounder and to control the lighting system.

Table 4.9: Physical Characteristics – Camera Stereo System (DLRS)

Specification	Description
Dimension / Mass	52 cm x 23cm x 19cm / 15.5kg
Connectors	Echosounder 2 USB for camera control Serial Communication Multibeam echosounder connection
Operating mode	Programmed / Remotely controlled
Depth	500 m

Table 4.10: Technical specifications – Camera Canon EOS 5D Mark II

Specifications	Description
Resolution	22MP (5616x3744) - Video FullHD (1936x1464)
Sensor	CMOS, 3:2
Lens	Canon 24mm
ADC (Analog to Digital Converter)	8 bit
Shutter range	1/8000 to 30 seconds
Frame Rate	25FPS @ FullHD, 50 FPS @ 1280x720
Video max length	30min (4Gb)

4.2 PAIRING POSSIBILITIES

Table 4.11: Sensor Pairing UC 3

Specifications	Pairing Options		
	1	2	3
SSS Klein towfish	X		
SSS Imagenex kit		X	X
MBES R2Sonic 2024	X		
Towed Camera	X		
Mono Camera		X	
Stereo Camera			X

In contrast to the first two use cases, where it is important that the different sensors be used simultaneously and/or deployed from the same platform, for use case 3 this condition is not a requirement. In fact, the acoustic side scan sonar and the optical cameras have optimal survey altitudes that are quite distinct. Whereas the Klein 3000 SSS is towed at approximately 12-17m above the seafloor, this distance is far too large for the optical cameras to collect discernible data from the bottom. Conversely, optical cameras present a good balance between bottom discernibility and image footprint when they are deployed at 5 or less meters above the seafloor. The lack of intersection of the operating altitude ranges of the acoustic and optical sensors is the main reason why the surveys are typically done with two distinct platforms and frequently at distinct times. Regarding this aspect and the bottom classification goals of UC3, what is most important is that the data is acquired with navigation sensing information that allows proper georeferencing of the two sensor modalities.

For the above reason we consider two sensor pairing possibilities. The first possibility concerns the use of the two towed platforms (the SSS Klein towfish and the towed camera), and the ship-based MBES. The relevance of this possibility comes mainly from the act that the existing legacy data provided by TA has been collected with this sensor pairing. Furthermore, additional data to be collected during the project will be primarily done with this pairing, given that it is the *de facto* sensor set used by TA in their activities. The MBES is envisioned to provide acoustic backscatter information of the seafloor as a complement to the information provided by the SSS. However, preliminary assessment of the relevance of the MBES-based acoustic backscatter indicates that it does not provide significantly more discriminative information than the SSS. For this reason, the need (and the future use) of the MBES will be re-evaluated during the initial year of the project, based on analysis of legacy data.

The second pairing possibility considers the use of the AUV for both SSS data and optical data collection. Under this possibility the Girona500 AUV would be equipped with the SSS Imagenex kit and with the monocular camera. This possibility has one important drawback which has to do with the lower performance of the Imagenex kit when compared to the Klein towfish. As such, the primary focus will be on the use of the Klein towfish, for two main reasons: (1) there is already a significant amount of available training data with the Klein SSS whereas there is currently none with the Imagenex, (2) data from the Imagenex is expected to be of considerably lower quality, therefore posing an unnecessary failure risk to the performance of the algorithms to be developed.

The third pairing possibility is a variant of the second, where the monocular camera is replaced by the stereo camera. The stereo capabilities of this camera are not required in the context of UC3, given that the objective of the optical imagery is to provide texture information of the seafloor appearance, rather than 3D structure information. The advantage of the stereo camera with respect to the monocular camera lies on the higher quality of the sensor which can potentially allow for better image preprocessing results and therefore better classification performance. However, it has the drawback of being considerably bulkier and not allowing direct control of the acquisition and storage of the images by the AUV.

5 CONCLUSIONS

The exploitation of acoustic and optical sensors is intrinsically related to the goals of DeeperSense project. Not surprisingly all use cases propose a combination of those two sensor types to some extent. Different types of acoustic and optical sensors will be applied aiming to maximize their performance for specific environment and application.

For use case one, the main goal is to generate realistic visual images from sonar images. Therefore, the straightforward sensors usage consists in an imaging sonar and cameras. One of the key points is to investigate the optimal placement of both sensors that would provide the results. We will also investigate the benefits of using stereo camera and 3D laser scanners during the training process, because, contrary to the mono-camera approach, they provide depth information besides RGB images. We expect that by having depth as an additional modality, the generation of visual images could be further improved.

On the use case II we will explore two forward-looking sonar options that have different resolutions: low-resolution and high-resolution. We expect that the high resolution FLS will result in better matching with the optical image. However, this is a more expensive sensor and because of its range limitation it is less commonly used. Therefore, we will also aim to achieve good results on the lower frequency FLS.

Finally, when it comes to the use case III, the distinct operating altitudes of the SSS and the optical cameras preclude their simultaneously use on a single platform. When comparing boat towed sensors with AUV deploy sensors, there is a clear advantage of former in terms of spatial coverage and therefore of quantity of available data for training. As such, the choice of the best sensor pairing will be guided primarily by the ability of the sensors to provide data with adequate quality, the availability of legacy data, and the ability to collect larger amounts of new data during the project, rather than by platform integration issues.

6 BIBLIOGRAPHY

Frans, K., 2017. Outline colorization through tandem adversarial networks. *arXiv preprint arXiv:1704.08834*.

Isola, P., Zhu, J.Y., Zhou, T. and Efros, A.A., 2017. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1125-1134).

Valdenegro-Toro, M., 2019. Deep neural networks for marine debris detection in SONAR images. *arXiv preprint arXiv:1905.05241*.

Wang, X. and Gupta, A., 2016, October. Generative image modeling using style and structure adversarial networks. In *European conference on computer vision* (pp. 318-335). Springer, Cham