# D1.2 - Roadmap and final requirements specification



| | |
|---|---|
| **Project number:** | 248623 |
| **Project acronym:** | TACO |
| **Project title:** | **T**hree-dimensional **A**daptive **C**amera with **O**bject Detection and Foveation |
| **Start date of the project:** | February 1st, 2010 |
| **Duration:** | 42 months |

| | |
|---|---|
| **Deliverable type:** | Report |
| **Deliverable reference number:** | 248623 / D1.2 |
| **Deliverable title:** | Roadmap and final requirements specification |
| **WP contributing to the deliverable:** | WP1 |
| **Due date:** | April 2012 – M27 |
| **Actual submission date:** | June 4th 2012 |



| | |
|---|---|
| **Responsible organisation:** | TEC |
| **Authors:** | TEC, TUW, SHADOW, OTL |
| **Abstract:** | This technology roadmap has three major uses. It helps to reach a consensus about a set of specifications and the technologies required to satisfy the user's need; it provides a mechanism to help forecast technology developments; and it provides a framework to help plan and coordinate future technology development. |
| **Keywords:** | Roadmap, Technology, Use cases, Benchmarks, Risks |



| | |
|---|---|
| **Dissemination level:** | Public |
| **Revision:** | V1.0 |

| | |
|---|---|
| **Instrument:** | STREP |
| **Thematic Priority:** | ICT |

# Table of Contents

# List of Figures



# List of Tables

# 1 Executive Summary

TACO develops a 3D sensing system with 3D foveation properties endowing service robots with a higher level of motion and affordance perception and interaction capabilities with respect to everyday objects and environments.

This document has three major uses. It helps to reach a consensus about a set of specifications and the technologies required to satisfy the user’s need; it provides a mechanism to help forecast technology developments; and finally it provides a framework to help plan and coordinate future technology development.

A preview of possible strategic directions of the technology with a focus on foveation aspects is given. The different capabilities of the TACO system are being described by introducing applications that are viewed primarily important and best demonstrate the advantages that the new technology will bring.

The selection process of use cases is presented. A subset of use cases are considered throughout. System requirements are charted up, as well as an estimate of potential markets where the TACO system could be introduced commercially. Requirements are being compared with actual data received from work package 2, 3 and 4.

One section deals with benchmarking. Existing benchmarks are being analysed and new possibilities for benchmarking with regards to a few selected use cases are being discussed. As far as we know, there are no established benchmarks for foveating sensors yet. Currently established benchmarks only exist for normal non-foveating sensors.

Three different benchmarking approaches have been chosen: the use of existing benchmarks for similar sensors to provide the characteristics of low-level hardware performance, the establishment of new benchmarks to characterize the new foveation functionality and as already said, a number of use cases will also be selected to benchmark the sensor towards real-world applications. These approaches are described in chapter 6.

In the Technology Roadmap and Future Implementation section the requirements to optimize the components of the TACO sensor with a focus on weight, size and robustness are being described. An outlook for future options and improvements is given. A section dedicated to the implementation possibilities beyond the project lifetime concludes the document. To complete the picture, a list of risks and corresponding mitigation strategies is presented to ensure that likely risks or problems are detected early and adequately attended and responded to.

# 2 Document Scope

The technology roadmap of Three-dimensional Adaptive Camera with Object Detection and Foveation (TACO) has been developed to reach a consensus about a set of needs and technologies required to satisfy those needs. It provides mechanisms to forecast technology development and a framework to help, plan and coordinate future technology developments.

The roadmap is significant to understand the status quo of the costs and performance and also to define the cost and performance goals. In case of market positioning the focus is the high-end market because the TACO sensor provides higher quality, better field of view (FOV) and higher resolution than cheaper 3D sensors. It outlines multiple pathways to achieve costs, size, weight, performance, capability and robustness goals.

This roadmap document exclusively deals with the achievement of the TACO project's short- to long-term objectives.

The specific scientific and technological objectives of the TACO project are:

- Develop a flexible, compact, robust and low cost 3D imaging device providing high resolution 3D data of high quality
- Achieve 3D measurements of increased spatial and temporal resolution in detected regions-of-interest by developing adaptive and intelligent software for sensor control and 3D foveation
- Benchmark the 3D sensing system on robots in an everyday environment test bed interacting with everyday object

More precisely, these objectives can be classified into short term and long term objectives which define the scope of the roadmap.

**Short term:**

Short term objectives reflect those targets the consortium wants to achieve during the project itself. The primary short-term targets are the following:

- The main short term target is to demonstrate how 3D foveation can be applied to fill the gap in 3D robotic sensing. This gap exists because at present it is necessary to choose between fast coarse resolution 3D scene imaging and slow fine resolution 3D imaging of details of interest. To achieve the short-term target we must develop some critical system components and integrate these in a 3D foveation camera.
- A bidirectional MEMS laser scanning mirror with large aperture
- A synchronized laser distance measurement unit including steerable focusing optics and time-of-flight electronics compliant with the MEMS laser scanning mirror
- A control software for enabling 3D foveation principles

Further short term targets are:

- Define design goals for both the hardware and software parts of the system, enabling both to become competitive towards other sensors and sensing principles. This is a necessity to ensure later exploitation of project results. These are addressed in section 2.
- Ensure that there is an easy to use and adapted interface to the sensor enabling later widespread use. Due to the technical nature of this material, this objective is addressed in D2.1 chapter 3.
- Ensure that links are made to other robotics projects for which the sensor would be of interest, to further generate interest with regards to the TACO project itself and towards foveated imaging in general. This sub objective is addressed in D6.2.

**Long term:**

Long term objectives are objectives not to be achieved during the project lifetime, but rather after the project has finished.

The primary long-term target is to enable future robots to be equipped with a fast adaptive–resolution camera sensor to enable interaction with the environment in a more natural manner.

The consortium believes this objective is primarily enabled through achieving the project's market and outreach (M&O) objectives during the project lifetime:

- Offer new technology to the European robotics industry, and make the new 3D sensing system attractive to both large and small scale actors within robotics
- Make TACO knowledge visible within industry and the scientific community
- Carry out proof-of-concept validation of the concept

The implementation of these M&O objectives is detailed in D6.2 – Project Dissemination Plan.

# 3 The TACO Vision

Service robots are currently mainly found within "4Ds" segment, performing tasks which are dull, dangerous, dirty or distant. As service robots become more sophisticated, new markets will open for robot technology within the fields of cleaning, construction, maintenance, security, health care, entertainment and personal assistance robots. According to EUROP's Strategic Research Agenda (2006 and 2009), the key challenges faced by the robotics sector in Europe are:

- to have greatly improved 3D vision sensors with higher frame rates
- to reach advanced task-dependent sensor fusion and a step change in visual serving
- to achieve a decrease in costs to further the opening up of mass markets

The TACO project will develop a flexible, compact, robust and 3D image acquisition device providing high resolution, high quality data for robot real-time operations. The development of such a device will entail the introduction of novel MEMS mirror devices, a novel 3D foveation concept and the provision of a novel camera that enables real 3D vision for robotics. The availability of such a device will be a key element for meeting all the challenges listed above.

TACO addresses the challenges in 3D sensing by an innovative concept for hardware-based 3D attention management based on a principle of 3D foveation. 3D foveation is an intrinsic property of the 3D sensing system based on the process of acquiring 3D images with coarse level of details, applying fast object recognition techniques to select areas of interest in the coarse 3D image and then concentrate image acquisition on regions or details of interest. This foveation process takes advantage of the finding that animal visual systems have solved the problem of limited resources by allocating more processing power to central than peripheral vision.

TACO introduces 3D foveation as an important concept for service robot interaction with their natural environment. By 3D foveation properties we mean properties based on the process of acquiring 3D images with coarse level of details, applying fast object recognition techniques to identify areas of interest in the coarse 3D image and then concentrate the image acquisition on details of interest allowing for higher resolution 3D sampling of these details. The robot and image acquisition system will autonomously be able to increase the level of detail whenever needed for interaction between the robot and everyday objects and humans. TACO will produce a novel 3D sensing system that includes three important parts:

- A novel concept for fast attention level management based on the 3D foveation principle enabled by dedicated sensor hardware.
- A 3D laser scanner sensor based on a miniaturised MEMS micro-mirror device combined with time-of-flight measurement technology, which will enable operation in different modes ranging from coarse broad field-of-view 3D image acquisition to higher resolution narrow field-of-view 3D image acquisition.
- A software framework for fast object recognition in everyday scenes based on saliency and visual cues. These cues allow efficient selection of details of interest, controlling the foveation process of the 3D sensing device.

The 3D sensing system will be fast, small, lightweight and relatively energy-efficient to facilitate use in real-time operations and efficient and practical mounting on a service robot or a robot arm.

# 4 Strategic Direction of TACO

TACO develops a 3D sensing system with 3D foveation properties by utilizing the power of micro-mirror MEMS technology combined with state-of-the-art time-of-flight methods. The proposed 3D sensing device will have the ability to concentrate 3D image acquisition and processing on the real details of interest for a robot allowing the robot to inspect and interact with these objects autonomously. The long-term goal of TACO is to equip robots with an affordable and fast 3D sensing system. TACO reaches this goal in 3 essential steps:

1. Development of a 3D sensing system based on micro-mirror MOEMS technology that enables large scale low cost production of a robust moving mirror needed for scanning coupled with fast 3D ranging measurements. Fraunhofer IPMS, Fraunhofer IPM and CTR are front-runners within these scientific fields and capable of building the 3D sensing system.
2. Creating the scientific foundation for fast and flexible processing and steering of the 3D sensing device in a manner that allows high resolution 3D image acquisition and processing only when needed to solve the robots tasks. SINTEF has the expertise within 3D computer vision and image analysis to implement a software toolbox for exploring and exploiting the novel 3D sensing system and its capabilities.
3. There is a goal in the project to create a test bed for proof-of-concept verification and testing in a real robotic environment. Oxford Technologies and Shadow Robot Company will perform the actual proof-of-concept study together with the Technical University of Vienna. The consortium will work closely with an industrial advisory board to promote the novel approach to 3D sensing and its uses in the robotics and manufacturing industry. The consortium already has links to organizations with possible future application of robotics where the TACO sensor could be of great value.

The aim is that TACO will form the basis for how robots in the future sense their environment and interact efficiently with objects. The ability to handle diverse situations and interaction modes with the environment will be important for robot manufacturers in Europe, for European research, for flexible industrial manufacturing and the advancement of service robots and personal robots in Europe. While Europe has a strong tradition of industrial robotics and biomedical applications, it is lagging behind compared to Korea and Japan concerning service robotics and personal robotics products. TACO aims at bridging this gap by utilising one of Europe’s strengths – we are excellent at producing robust industry standard sensors and sensing systems. Of particular importance to affordance perception is the possibility to integrate task-specific feature extraction and feature combination into the software layer of the sensor itself which is addressed in the TACO project.

In sum, the TACO advantages over existing devices can mainly be described as follows:

- High resolution - Simultaneous high spatial and temporal resolution (in regions of interest)
- Adaptability - possibility to vary between high frame rate/low spatial resolution and low frame rate/high spatial resolution
  - change in operation modus – from navigating (large FOV; key point tracking, obstacle avoidance) over searching (large-medium FOV; medium-high resolution; object detection) to handling (small FOV; high resolution)
  - typically, when in need for several 3D sensors, those can be put into one TACO sensor to reduce overall system cost and complexity

- Large focus depth - equivalent beam divergence at all distances (~0.2 -15 meters)
- Multi-targeted acquisition – the ability to focus attention and resolution on multiple targets simultaneously within the scene

The above listed advantages brought about by the TACO project are further highlighted by certain aspects of the selected use cases in 5.2.

# 5 Capability Areas

The TACO system aims to develop a sensor targeted at improving the operation of robotics and manipulators in real life operating environments. The system will endow the robot platform with the following capabilities:

- Sensor input for accurate interaction and manipulation of objects
- Obstacle detection and avoidance
- Real time localization and knowledge of changing environments and structures
- Human-robot physical interaction

Here, possible areas of application for the TACO sensor are areas inaccessible (or restricted for humans) for reasons of distance (space applications), confined environment (urban search and rescue robots), or dangerous areas (field robotics such as decommissioning or nuclear facility maintenance fields). Other applications also include fields where the tasks are routine and tedious, where robot systems can perform with an excellent level of repetitive accuracy over extended periods of time. Examples of such applications would be in home or factory environments. These applications all require a reliable and rugged sensor that can be applied to robot systems with a various levels of autonomy from fully automatic to operator assistance of teleoperated manipulator systems. However, there are capabilities beyond the robotics segment, such as:

- Faster 3D measurements of industrial objects (e.g. parcels and manufactured goods)
- Inspection of structures (e.g. subsea or land-based)
- Semi-autonomous vehicles (e.g. wheel chairs, cars and trains)
- Security surveillance

Moreover there are notable capabilities displayed by the individual components of the TACO system. The bidirectional scanning mirror has a number of potential application areas outside the TACO camera. The time-of-flight electronics could be reused in other settings. The foveation software can be reused with other camera and laser scanning systems such as 3D machine vision applications or automatic robotic maintenance applications.

This section of the report describes the process used to establish the prime capabilities of the TACO sensor, through applications envisioned for the sensor (section 5.2), analysis of competitive sensors (section 5.4) and final establishment of the system requirements (section 5.3). A basic market and pricing analysis (section 5.5) is also described.

## *5.1 Use Case Gathering and Requirements Establishing Process*

The TACO consortium contains both RTD and industrial end user partners. The end users are particularly able to bring to the project knowledge of sensor requirements use for real life applications. This also applied to some of the RTD partners where the state of the art of existing sensor technology has direct bearing on their research.

As a starting point the partners were requested to supply a number of use cases where partners envisioned that the TACO system would be of real benefit. This involved introducing the field of application, the current problem or limitation where TACO could assist. This then lead to how TACO would be applied and would be able to show a definite benefit over the current state of the art of sensor technology. The templates which the partners used for those sections that required them to specify the requirements for the sensor in terms of data format and quality, sensor physical properties and acceptable cost range that would be necessary for the sensor to be of use in the discussed applications. Finally the partners were asked to discuss the market that this application would have reach over both in terms of potential sales in units for a commercially available TACO system and different sectors of industry that the application covers. These applications are summarized in section 5.2, with market analysis into section 5.4.

Furthermore, each partner was asked to provide input on how they would prioritize different aspects of the sensor. For instance, due to hardware constraints, it is difficult to obtain simultaneously high frame rate and high resolution. For each use case, the users were asked to submit a prioritized list of the importance of each aspect of the sensor. These are summarized in section 5.2.2.

SINTEF, in its role as the technical leader, then reviewed these use cases to verify that they indeed were good use cases for a foveating sensor. The major finding was that the foveation aspect was immaturely established amongst partners, meaning that many of the use cases would be easily solved with existing sensors. One typical example is a use case involving static scenes requiring very high spatial accuracy – these are very well suited for e.g. structured light systems. We believe this is natural due to the new sensing principle employed, which needs establishment amongst partners. Furthermore, this concern will need to be addressed when the sensor later will be marketed towards the robotics industry.

Based on a dialogue with partners, some use cases were selected as good cases thought to exhibit the foveation aspect of the sensor well. This means that they should highlight well at least one of the advantages given by the sensor, as presented in section 5.2.2. Use cases were adapted accordingly, and the revised version of the selected use is presented in section 5.2.

To establish final hardware parameters, we chose to look towards existing sensors to ensure that the hardware had a competitive and sustainable advantage. This means that while use cases were used as basic input for establishing numbers, the prime source for actual hardware specifications were competitive sensors. This analysis is further detailed in section 5.4.

The RTD partners used this input to establish first of all a list of actual priorities of the sensor, plus a design goal for the sensor. As there are numerous uncertainties remaining with regard to sensor development, we decided that it was most important to decide on priorities amongst sensor capabilities than actual target values (as they may be unreachable). This is presented in section 5.3.

The TACO requirements were established so they will primarily reflect the TACO advantages. When trading features to be prioritized this must reflect the benefits of the TACO sensor compared to other already existing sensors. I.e. traditional laser scanners will be superior compared to the TACO sensor with regard to spatial resolution, but the temporal resolution and the field of view will be lower.

## *5.2 Selection of Use Cases*

This section introduces the use cases that were proposed by the partners. They represent real life applications where each partner envisions that the TACO system would demonstrate a clear and significant advantage over existing technologies. The use cases are arranged to introduce the individual use case, i.e. its area of application, why the TACO system would be of use to the application and how it will be implemented to demonstrate this advantage.

### 5.2.1 Overview of the Use Cases

#### 5.2.1.1 Augmented Reality and Task Safety

JET (Joint European Torus) is a large magnetically confined experimental fusion reactor near Abingdon, UK. Experiments carried out here advance the field towards the creation of fusion power stations and increase understanding of tokamak design and control. ITER is the international follow on project currently being built in the south of France. In both cases

remote handling is used for a variety of reasons, some areas can be highly radioactive, others have hazards such as beryllium dust preventing unsuited access and many components are simply too large for humans to handle.
Currently all operations are performed with a man in the loop due to the high uncertainty in the status, location and condition of components with information of the environment being supplied through the use of 2-d cameras and monitors. Undesirable shadows and contrasts when combined with the standard limitations of camera viewing have a significant impact on human image processing and make depth perception impossible in some circumstances. Moreover, full manual operation has a large impact on operator fatigue rates especially with repetitive tasks.
At the simplest level, data from the TACO system could be used to augment the camera views of the operators giving additional depth cues or could be used to draw the operator's attention to certain features. At a higher level, data from the TACO sensor could be employed to allow partial automation of the manipulators in some tasks if the necessary level of environment awareness could be provided. In areas which allow a mixture of manual and remote work, this could be accomplished with human or human controlled remote manipulation being performed within its workspace.

#### 5.2.1.2 3D Inspection

Following a campaign of operation within a fusion vessel one of the first tasks for the remote handling operators during a shutdown is to inspect the vessel for damage to the first wall. This involves essentially manually looking for any differences to the structure from prior to the fusion campaign. This may be, for example, cracks to blankets, something being bent, broken or missing.
Inspection operations are performed with a man in the loop controlling an articulated boom mounted manipulator to move the position of the cameras around the vessel whilst inspecting the images on the monitor for visible signs of faults. Initially a video survey is performed which is implemented using the manipulator following an automated trajectory. The video must be manually inspected in the control room to locate areas of further interest, e.g. potential damage sites. Following this either high resolution video or still images (which may be overlaid to create a surface map) may be obtained. The camera will be positioned with a mix of man in the loop and automated control of the manipulator system. The most accurate system uses a photogrammetric camera to obtain accurate 3D position data. This requires targets to be placed onto the surface of components. It would be desirable if an intelligent system such as TACO could be used to obtain accurate position data using just visual cues from features such as edge etc.
Currently used sensor technology requires a significant amount of manual assistance which is tedious, time consuming and prone to errors and missed visual cues. During a shutdown there is a significant cost per hour for any delay, therefore, any new system that can shorten the necessary length of the shutdown is very attractive. TACO could potentially automate this procedure by controlling the remote handling manipulator system to perform a gradual systematic sweep of the vessel wall structure looking for obvious signs of damage. Any detected differences could override the sweep to perform more detailed scans of the area to gain a better indication of the size of the damage.
3D inspection is also performed prior to installing new components. When a new fixing point is required it is welded into position using remote handling means. Currently only approximate positioning is possible (<10mm) and subsequent 3D inspection with a photogrammetry camera obtains accurate values of its actual position before make-up pieces are manufactured to assure that the new component when installed is in its desired location.

TACO is thought to be able to significantly improve this process by aiding accurate positioning of the fixing point in the first instance and hence avoiding the requirement of make-up pieces.
The vessel prior to a campaign is completely structured. Detailed models of the vessel exist due to the design process used to create the vessel, which are also used by the VR system to help with the remote handling operations. Variations (i.e. an unstructured condition) from this known state when entering the vessel following a campaign of operation potentially indicate damage to the vessel and areas that the TACO sensor should foveate upon to gain further detailed information, which should be flagged and indicated to the operator.

#### 5.2.1.3 Public Safety

In the public safety sector, robots are an invaluable tool to save lives and prevent people from having to manually attend to suspect devices. After detecting a suitcase, a swab will be passed to the robot (and tracked by the TACO sensor). The suitcase will be swabbed to simulate the search for explosives. Then the suitcase will be emptied until a suspicious object is found. Here, the TACO system will need to identify an object the size of an AAA battery in a cluttered space, then to plan a route to extract it without contact. This use case will test the TACO sensor against a cluttered and unstructured environment with varying light condition.
During the extraction of the object, the foveated region will ensure both a high spatial (for finding the AAA battery for example) and temporal (to track the swab) resolution in the ROI. At the same time data coming from the non foveated region will simultaneously ensure the extraction path avoids any collisions.

A UML timeline describing this use case step by step can be found in the appendix. It also distinguishes the work done by the robot, the foveation software and the sensor itself.

#### 5.2.1.4 Home Grasping

The purpose of this use case is to test the sensor within a cluttered domestic environment. A platform with the TACO sensor, a robotic arm and a robotic hand would be used to locate and manipulate some day to day objects. The tasks would be: identify and pick up a specific pen from a pot of pens, pick out a specific page of a document from a pile of documents, pick up a full mug from a desk and put it somewhere else on the desk.
This uses will highlight the adaptability of the sensor. While moving the full mug, the foveated region of the sensor will be used to ensure a high spatial and temporal resolution of the sensor in the foveated region.
The foveated region of the sensor will pick up the interesting region and allow for a good grasping of the different objects. The information gathered by the sensor out of this foveated region will be used for obstacle avoidance.

The UML diagram in section A.1 shows the typical use case scenario for home grasping: The robot navigates in the kitchen (unfoveated TACO sensor), closed doors need to be opened (TACO sensor foveates on door handles). In the kitchen the robot starts the search for the object (TACO sensor foveates on object candidates). The UML diagram presents the communication between the robot software, foveation software and sensor hardware.

### 5.2.2 Prioritization of sensor capabilities

For each of the use cases, the partners were asked to prioritize the following aspects of the sensor:

- Field-of-view unfoveated area
- Frame rate unfoveated area
- Resolution unfoveated area
- Field-of-view foveated area
- Frame rate foveated area
- Resolution foveated area

These capabilities were chosen as they are competing constraints in sensor design (i.e. increasing resolution decreases frame rate).
Based on this, we created a comparison graph that graphed these abilities against each other (Figure 1).

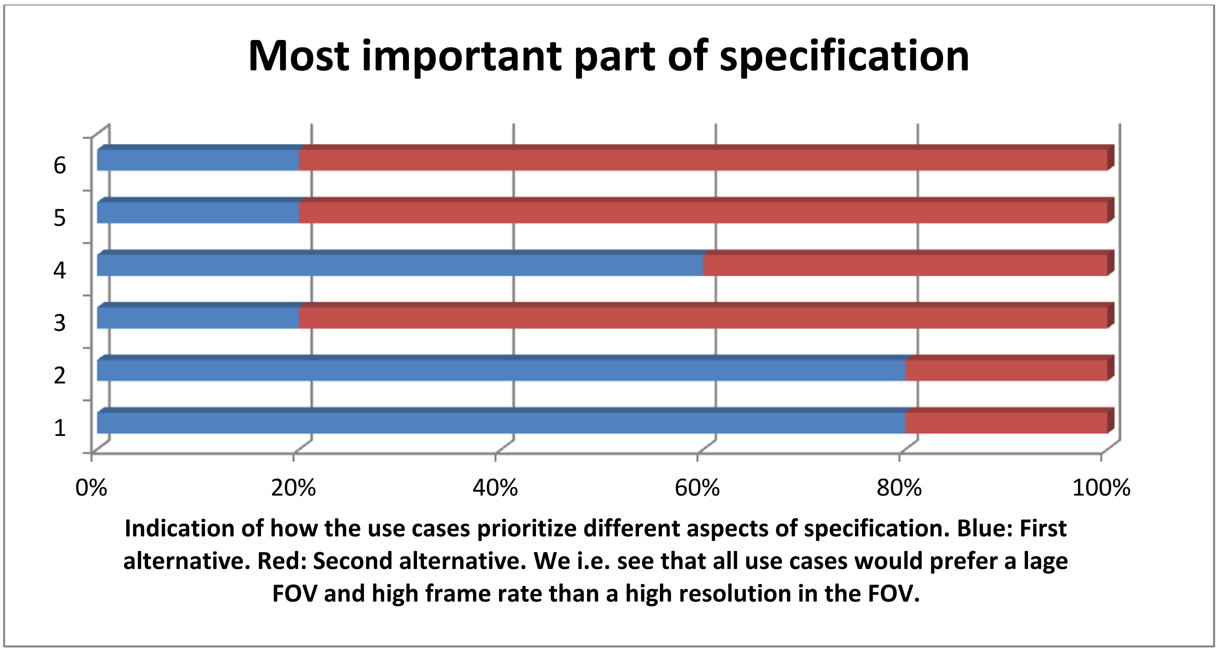



Figure 1 Prioritization of sensor aspects. Blue/red bars indicate how many use cases desire one sensor aspect above another. For instance in line 1, we see that ~60% (100%-40%) of the use cases would prefer resolution over frame rate (Hz) in the foveated region.

From this we see that most use cases would like high resolution in fovea and large FOV outside fovea. This could provide direct input for sensor design, but needs to be compared with providing a sustainable competitive advantage, which will be discussed later.

### 5.2.3 Comments on specifications given in use cases

The specifications of the use cases selected to showcase TACO are shown in Table 1. When reviewing the use cases and their specifications, SINTEF in its role as a technical leader, observed the following issues:

- It would be fair to say that not all numbers were given with full and qualified judgment. There are some contradictions in the specifications – i.e. specifying 20 points per degree in fovea and angular resolution requirement of ±5mm (at 1m) which corresponds to around 3-4 points per degree.
- We believe some parameters where influenced by "limiting beliefs" about the sensor specification possibilities (based on numbers from kick-off meeting etc), meaning that both must-have and nice-to-have numbers cannot be interpreted directly.

Some of the use cases were not particularly good for promoting the main advantages of TACO and may have been better off with another type of specialized sensor (also after pruning)
In total, this meant that we decided that exact numbers for hardware specification could not be derived directly from the use cases, not even for the selected ones as shown in Table 1. Instead, we decided on looking at competing sensors whilst simultaneously taking into account the use cases envisioned.

| | TUW | Shadow | OTL_AR | OTL_3D |
|---|---|---|---|---|
| Distance min | 0.1m | 0.7m | 0.02m | 0.02m |
| Distance max | 7.5m | 2m | 10m | 10m |
| FOV must have (hor) | 60° | 40° | 30° | 30° |
| FOV must have (ver) | 60° | 40° | 20° | 20° |
| FOV nice (hor) | 180° | 60° | 90° | 90° |
| FOV nice (ver) | 120° | 60° | 70° | 70° |
| Fovea must have | 5° | 1.2° | 1.2° | 1.2° |
| Fovea nice | 12° | 9° | 4° | 4° |
| Resolution FOV (points per degree) | 2 | 3 | 4 (2) | 4 (2) |
| Resolution Fovea (points per degree) | 10 | 20 | 90 (35) | 90 (35) |
| Accuracy FOV depth and angular | ±1mm | ±5mm | ~±3mm | ~±3mm |
| Accuracy Fovea depth and angular | ±1mm | ±5mm | ~±3mm | ~±3mm |
| Frame rate | 30Hz | 50Hz | 30Hz (10Hz) | 30Hz (10Hz) |
| Size | 10x10x10cm | 10x10x10cm | 5x5x5cm | 5x5x5cm |
| Weight | <0.5kg | <1kg | <1kg | <1kg |
| Price | 1 000 € | 2 000 € | N/A | N/A |

Table 1 Summary of selected use cases from TUW, Shadow and OTL

## *5.3 System Requirements*

Mostly based on competitive analysis, while simultaneously looking at the use cases, the following design goals were set.
**Primary goal: As high as possible resonant scan frequency (=> frame rate) but with reasonable beam divergence (=> resolution)**

Other goals:

- Large FOV (both in resonant and quasistatic view). FOV of coarse view more important than FOV of foveating region.

- We should have a measurement range of ~7.5 meters; preferably 10 meters.

| Sensor aspect | Design goal |
|---|---|
| **Distance min (cm)** | 20 (10) |
| **Distance max (m) [Fovea max 3 m long vs. short range system]** | 7.5 (15) |
| **Mirror resonance fast axis freq (Hz)** | 2000 |
| **Mirror resonance slow axis freq (Hz)** | 120 |
| **Mirror diameter (mm)** | 2..3 x 3..5 |
| **FOV (hor/ver) – degrees (sym, quadr. nice to have)** | 90 / 60 |
| **Fovea (hor/ver) – degrees** | 60 / 40 |
| **Beam divergence (mrad) better than TOF in 1 m** | 1.1 .. 2.2 |
| **Points per degree (beam divergence)** | 8.0 .. 16.0 |
| **Accuracy of angular MEMS position (points in FOV)** | 2000 |
| **Points per degree (mirror positioning)** | 22 |
| **Angular accuracy at 1m (mm)** | 2 .. 3 |
| **Depth accuracy (mm) (single shot 5, 3 nice )** | 5 (3) |
| **Frame rate (HQVGA 240x160 (Hz) [FOV 90 x 5..50])** | 25 |
| **Frame rate (HQVGA, interlaced 1:3 [FOV 30..90 x 5..50])** | approx. 8 |
| **Size** | 10 - 30cm³ |
| **Weight** | <1 - 4 kg |

Table 2 Indicative sensor specification. This is a design goal based on early simulation, not a final specification

## 5.4 Competitive analysis

It is important for the sensor that the hardware in itself is competitive, and sustainably so. While foveation might be nice, and the enclosed TACO software will provide a competitive advantage in itself, the hardware must be competitive on its own. If not, we might well end up with the TACO foveation software being used with another competing sensor. This is not a desirable outcome of the project.
We have further decided to exclude Stereo vision systems due to the incompleteness of the 3D data that they provide.

***Comparably sensors for 3D-imaging have been divided into different technologies:***
**2D laser (with tilting unit):**
A perception system consists of a camera and a three-dimensional laser range finder, which is based on a two-dimensional laser scanner and a tilt unit as a moving platform.
Constructing an elevation map of the environment for example is possible when utilizing a 2D laser range finder which is mounted on a tilting unit. The point cloud, which is generated by the laser range finder, is converted to an elevation map. Various filters are investigated to reduce the noise effects to get a consistent map.
**2D laser triangulation:**
With help of the laser triangulation principle, 3D cameras – like the so-called "Ranger" – can measure 3D data. Hence, to be able to measure 3D shape, an external line-generating laser

source is required. The laser module is mounted to project its laser line on to the object. By measuring the camera's laser line, the height of the object can be computed. This unique camera technology is capable of finding the position of the laser line by itself and reducing to whole image information into compact laser coordinates.

**3D laser:**

The RIEGL LMS-Z420i for example is a rugged and fully portable sensor especially designed for the rapid acquisition of high-quality three dimensional images even under highly demanding environmental conditions, providing a unique and unrivalled combination of a wide field-of-view, high maximum range, and fast data acquisition. The range finder electronics of the 3D laser scanner are optimized in order to meet the requirements of high speed scanning (high laser repetition rate, fast signal processing, and high speed data interface).

**3D structured light:**

A single camera and a LED projector using the DMD technology is used to perform 3D measurement procedures with structured light and a phase shift algorithm. The 3D scanner system from Vialux offers short measurement periods, as well as high precision and the operational mode is very flexible in its adaptation, which are very diversified and vary from the medical to the industrial field.

**Light Coding:**

Coded structured light is a computer vision technique which aims to reconstruct objects. This kind of technique belongs to the group of active triangulation methods. The setup usually consists of a Infrared-LED projector and a camera. The LED device projects a pattern or a set of patterns so that finding correspondences between the projected and the viewed images becomes easier. The coding strategy used to generate the patterns determines the performance of the system in terms of resolution and accuracy.

**Time-Of-Flight cameras:**

A time-of-flight camera is a range imaging camera system that resolves distance based on the known speed of light, measuring the time-of-flight of a light signal between the camera and the subject for each point of the image. The time-of-flight camera is a class of scannerless LIDAR, in which the entire scene is captured with each light pulse, as opposed to point-by-point with a laser beam such as in scanning LIDAR systems.

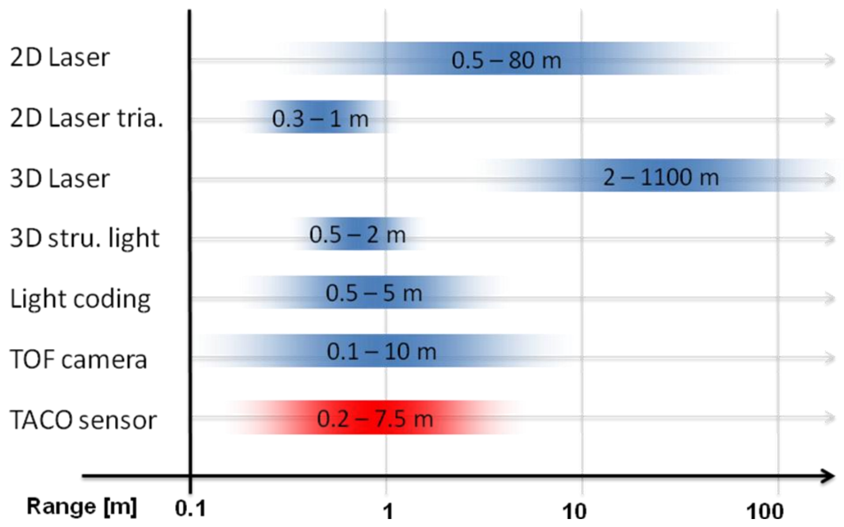



Figure 2 Operating distances of different sensors

The diagram above displays the different operating distances of sensors. There exist quite some different application fields. 2D laser triangulation is only adaptive for the close-up range. 3D laser (eg. For modeling whole buildings) have a very high range compared to other sensors. The TACO sensor itself is similar to the TOF and light-coding and has a range from 0,2 to 7,5 m.

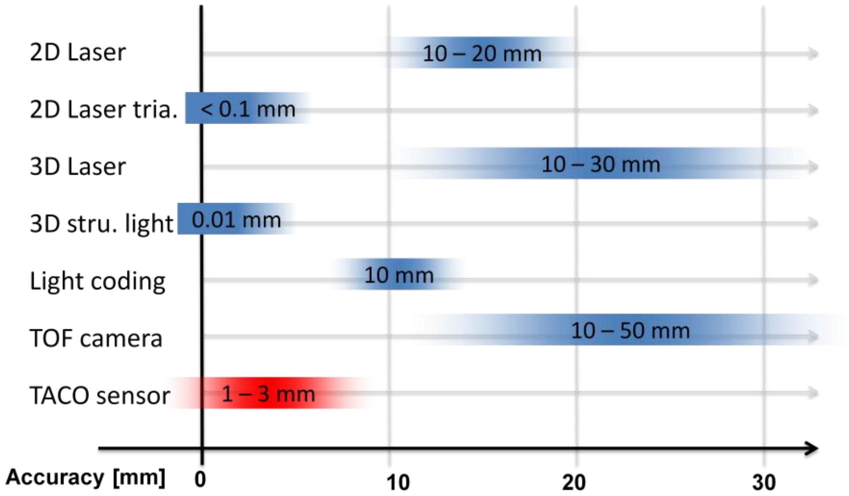



Figure 3 Depth accuracy of different sensors

With regards to the accuracy 2D/3D laser have a typical sigma from 10-20 mm, which is shown in the figure above. In contrast 2D laser triangulations are very detailed (applicable for measurement for drugs packaging). 3D structured light sensors features the highest dimension of accuracy and therefore these sensors are especially qualified for fancy and fine configurations. In comparison to the TACO sensor, the TOF camera is significantly less accurate.

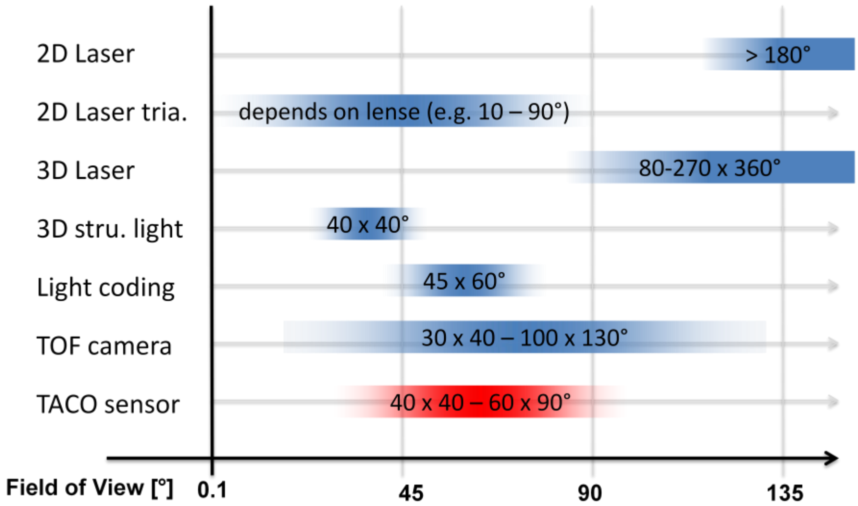



Figure 4 Field of view of the sensors

Concerning the field of view, all mentioned laser cameras have different properties. 2D lasers are generally bigger than 180° (the vertical disbanding is controllable with tilting-unit). The 2D laser triangulation is addicted to the lens which is used. 2D laser triangulation is fixed to the certain setup and not mutable changeable. The TACO sensor has a similar field of view like the FOV of light coding and TOF cameras.

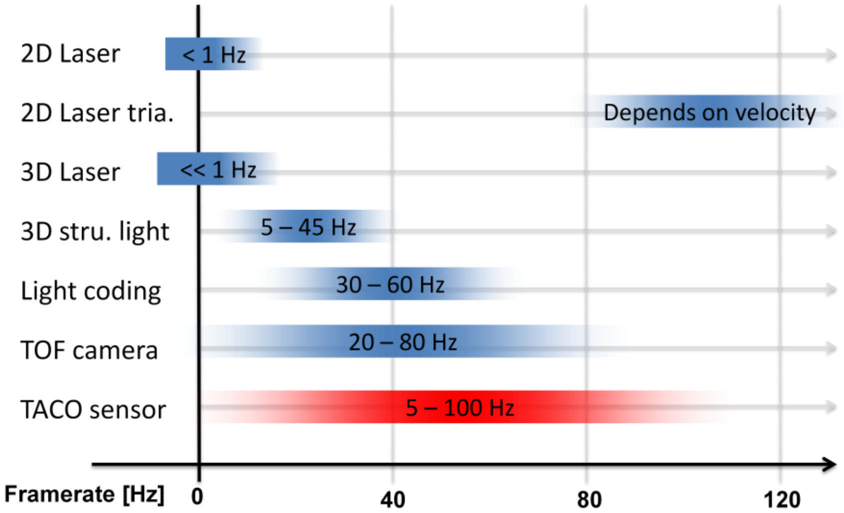



Figure 5 Framerate of different sensors

In terms of the framerate of the sensors, 2D laser are very slow, because of the slew round of the tilting-unit. 2D laser triangulation depends on the speed of the measured objectives, which pass through. The TACO sensor offers a very flexible framerate compared to light coding and TOF cameras. Typically the light coding has a framerate of 30 or 60 ps.

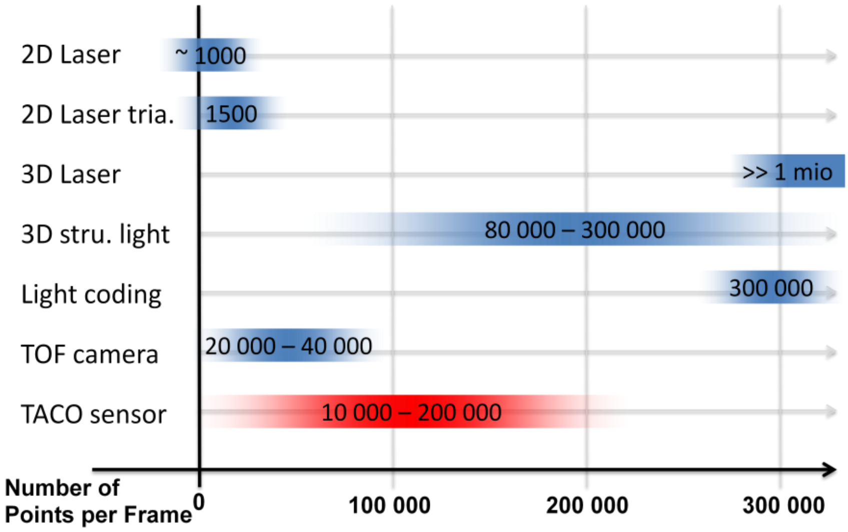



Figure 6 Number of points per frame

The figure above shows the number of points per frame. Concerning this parameter, 2D laser provides around 1000 points per scan. 2D laser triangulation features almost the same dimension. 3D laser in contrast have more than 1 million points per frame. The TACO sensor itself is variable from 10000 to 200000 points, depending on the framerate. A higher framerate means fewer points at the same time and vice versa. Light coding typically has a resolution of 640x480 (that is approximately 300000 points).

**Conclusion:**
To sum up, only TOF, light coding and 2D Laser (with tilting unit) offers comparable fields of applications to the TACO sensor.

- 2D laser on the one hand provides a higher operating distance, but on the other hand they are less exact and can´t keep up with the framerate.
- TOF cameras indeed have relating to FOV, framerate, operating distance and disbanding similar specifications as the TACO sensor, but TOF cameras also offers a lesser accuracy. In exchange TOF cameras are not that expensive than the TACO sensor. One of the disadvantages of the TOF is termed by: the integration time can limit frame rate and can cause motion blur, and data points can influence others (it is hard to get sharp corners of concave surfaces)
- Since the Kinect has been introduced, the light coding is widely-used. It has convenient acquisition costs (approximately EUR 100), but offers a lesser accuracy as the TACO sensor. As further disadvantages of the Light Coding could be named:
    - Dependency on background lighting (not useable for outdoor-areas)
    - Contingent on materials´properties (eg. Mirror finishes provides holes in the range image)
    - Contingent on the geometry of the objects (curves are not offered accordant to the reality)
    - Accuracy depends on the distance (depth resolution depends on the range, far away=> worse depth resolution)

## 5.5 Potential markets and pricing

The potential markets where a commercially available TACO system could be used and would be attractive is related to the use cases discussed in Section 5.2. These use cases apply to robots or manipulators where humans are not able to go, or access is restricted. Although the use cases discuss particular focused applications with a relatively small market share, development of the TACO system to be useful here would have applications over many more areas in industry.

- Remote maintenance in the oil and chemical process industries
- Automated manufacture in industries such as metal processing, computer aided manufacture, the automotive industry, food (meat) processing and furniture and textile industries
- Domestic service robots
- Object interaction
- High price marked: medicine, oil
- Medical assistance robots (both for home and hospital environments)
- The Space industry (NASA, ESA) for future unmanned space flights or EVA manipulator operations.
- Manipulation and teleoperation for nuclear decommissioning and hazardous material handling (CERN etc)
- Defense robotics (MOD), counter terrorism, urban policing and urban search and rescue.
- Surveillance applications

It is important that the sensor has a competitive pricing compared to competing sensors in the market. However since the TACO sensor will be adaptable with the ability to either produce high resolution or high frame rate interchangeably and also have added "smart sensor" features in the form of foveation capabilities, a higher price compared to less featured sensors can be acceptable.

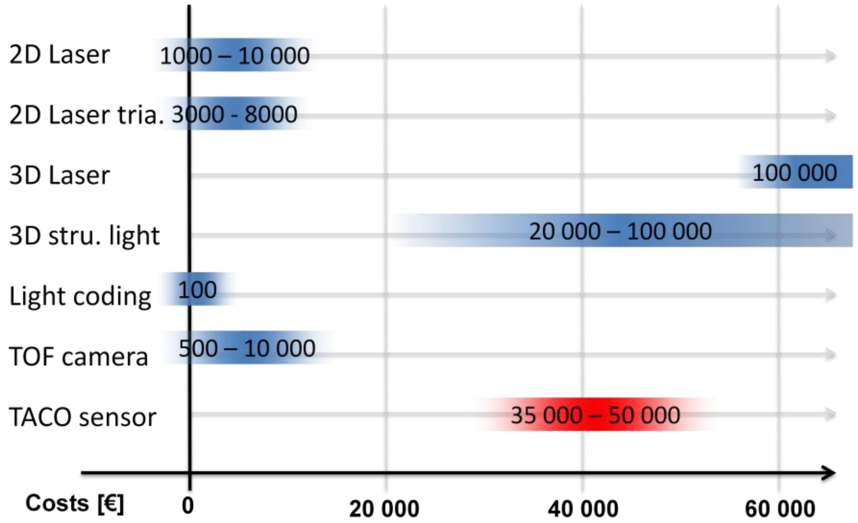



Figure 7 Cost range of different sensors

The expected price for the TACO sensor could be named by approximately EUR 50.000 (it might be reducible by production for the mass market) Compared to the TACO sensor, the TOF is quite cheaper and light coding could be finally named as the cheapest with EUR 100,--. (number of pieces into the millions, since adopted by consumer marked).

# 6 Benchmark definition

In order to quantify the performance of the TACO sensor it is intended to verify that each of the S&T project objectives is reached. The main S&T objective from the DoW is to

> *Develop a 3D sensing system with real 3D foveation properties endowing service robots with a higher level of real time affordance perception and interaction capabilities with respect to everyday objects and environments.*

This main objective is split into the following S&T sub-objectives:

1. Develop a flexible, compact, robust and low cost 3D imaging device providing high resolution 3D data of high quality
2. Achieve 3D measurements of increased spatial and temporal resolution in detected regions-of-interest by developing adaptive and intelligent software for sensor control and 3D foveation
3. Benchmark the 3D sensing system on robots in an everyday environment test bed with interaction with everyday objects

Currently established benchmarks only exist for normal non-foveating sensors. To our knowledge there are no existing standards for benchmarking foveating sensors. In order to handle this three different benchmarking approaches have been chosen:

1. The use of existing benchmarks for similar sensors (i.e. laser scanners) to provide the characteristics of low-level hardware performance (described in section 6.1 below)
2. Establish new benchmarks to characterize the new foveation functionality provided by the sensor (described in section 6.2 below)
3. The above approaches will be performed on synthetic scenes. A number of use cases will also be selected to benchmark the sensor towards real-world applications (described in section 6.3 below)

## *6.1 Existing Technology Benchmarks*

### 6.1.1 Optical Resolution/Beam Profile

The system resolution in the lateral direction depends on the following characteristics:

- The distance to the object of interest from the sensor
- The laser beam profile (which itself depends on the above)
- The motion of the mirror in the emission branch during the measurement of a target point
- The precision of knowledge about mirror position (and motion)
- And for the system as a whole on the algorithms used to match the 3D data along vectorial trajectories to fixed grids expected at the camera system user interface.

To achieve the optimal system resolution the beam profile will be measured statically (without mirror motion) using an appropriate high-resolution camera at predefined distances. Also mirror motion and position measurement will be characterized with respect to absolute precision and repeatability (see below). Finally, integral measurement of sensor resolution will be performed on a star-like pattern of black and white triangles converging to the center point (Siemens star). Alternatively, we may choose a characterization by the response of the scanner to a depth or intensity step.

### 6.1.2 Scan mirror positioning

The assessment of the mirror capabilities with respect to static and dynamic positioning will be based on the mirror characterization by FHG-IPMS Repeatability and Absolute Position Error.

The repeatability of the angular positioning will be assessed from 3D image properties (using a fixed image model at different distances) or from mirror characterization results by FHG-IPMS.

#### 6.1.2.1 Positioning dynamics

FHG-IPMS will characterize and precisely check the mirrors before assembly. Also, the step-response of the mirror will be compared to simulated results.

The TACO sensor will also be tested after construction by verifying the returned mirror position information in response to a step voltage demand to the drive electronics. The temporal resolution is the pixel period (1 µs).

### 6.1.3 Field-of-view

The FOV of the camera in full-frame operation will be checked by analysis of a camera image of the laser illuminated image region, possibly for different regimes of operating parameters which let expect influence on FOV (as, e. g. operating frequency of the quasi-static axis).

### 6.1.4 Distance measurement

The adjustment of the reception branch with respect to the emission branch is ensured by the operating principle, using the same light paths for emission and reception. The distance measurement error will be tested by repeatedly scanning a single line on a simple (black and white) target. This target will be moved radially away from the camera in measured distance steps. The average distance and standard deviation at each measured point (both as a function of distance) will be calculated and compared to the actual distance. Scanning must be restricted to a small angle around the center of the FOV (or a selected number of directions). Some other situations that are known or expected to show artifacts will be checked.

Measurement of a large radial depth change will test the system step response. This is influenced by a combination of the optical resolution and receiver bandwidth.

The measurement error and step response tests will be repeated using combinations of targets of both very high reflectivity (e.g., retro-reflecting foil) and very low reflectivity.

### 6.1.5 Intensity

The camera's output response to intensity of the received signal will be tested using a target with grey steps of known reflectivity. The intensity results must be proportional to the reflectivity of these steps. Tests will be conducted at a number of discrete distances.

### 6.1.6 Standards Conformance/Environmental conditions/Integral Properties

For adherence to safety standards, please see D2.2, Chapter "Optical System, Laser Safety, Distance Measurement."

Protection from humidity, extended temperature range and mechanical vibration will be addressed in future product development and will, therefore, not be tested. Tests of the TOF circuitry with respect to ambient illumination will be performed. From our design computations we do not expect sensitivity to background light.

### 6.1.7 Size/Weight/Mechanical Stability/Power Consumption

The system size, weight and power consumption are consequences of the final mechanical design which will be part of deliverable D2.2. These will be checked after construction.

### 6.1.8 Software/Camera Interface

Adherence to defined protocols is subject to unit testing. Special treatment will be devoted to laser safety issues as may arise if scanning motion is controlled by software.

#### 6.1.8.1 Internal Parameters and Variables/System Monitoring

During the course of the project, we will identify critical parameters and values to be tested; D2.2, D3.3 or D3.4 will contain a list of these parameters.

#### 6.1.8.2 Calibration

Calibration is an integral part of system integration and adjustment the accuracy of which will be tested implicitly 6.1.4 and 6.1.5.

## *6.2 Establishing Benchmarks for foveation*

As far as we know, there are no established benchmarks for foveating sensors. We have therefore chosen to define benchmarks according to the S&T Objective 2 of the Description of Work which provide the objective for foveation. The detailed explanation of the objective from the DoW is as follows:

- **Increased spatial and temporal resolution.** A three to ten times increase in spatial and temporal resolution in regions of interest without increased total data rate.
- **Adaptive and intelligent.**
    - Attention will be focusable towards the robot’s high-level goals
    - All camera parameters will be continuously tuned for optimal scene imaging through feedback from scene analysis
    - Automatic detection of underperforming attention detectors for improved robustness

We believe these points fully explain the meaning of the S&T Objective 2.
We believe that this is most efficiently benchmarked through the established use cases. We therefore only here provide forward links to which use case benchmarks each objective.

### 6.2.1 Increased spatial and temporal resolution

#### 6.2.1.1 Increase in spatial and temporal resolution

This will be benchmarked (amongst others) in OTL's use case on 3D inspection, where the increase in frame rate and temporal resolution will be measured.

### 6.2.2 Adaptive and intelligent

#### 6.2.2.1 Attention focusable towards robot's high-level goals

In each use case, the robot will have different high-level goals. Through the different use cases, we will thus see the camera adapt to each situation individually.

#### 6.2.2.2 All camera parameters will continuously be tuned for optimal scene imaging through feedback from scene analysis

There are primarily two camera parameters that can be tuned at runtime: Spatial and temporal resolution, plus the Field-of-View. The remaining parameters – such as laser intensity and measurement time – are largely fixed due to hardware and laser safety constraints.

Therefore, we interpret this statement as the ability of the system to adapt the temporal and spatial resolution to the scene at hand and according to the high-level goals of the robot. As we see it, this is implicitly benchmarked through the other benchmarks. We do not therefore see the requirement for additional benchmarks to target this specific detail of S&T Objective 2.

#### 6.2.2.3 Automatic detection of underperforming attention detectors

As described in D4.2, the system will switch to becoming a non-foveating system if more than a certain amount of the scene is deemed interesting. This is due to that the system either is unable to focus on anything, or only detecting noise.

This will be monitored for in the use cases, and reported as this occurs.

## *6.3 Benchmarking Use Cases*

The benchmarks above are synthetic, meaning that they do not actually correspond to real-life situations. The following benchmarks are related to the actual use cases, and the experiments that will be carried out in these to quantify the quality of the instrument itself in designated settings.

The experiments per use case are detailed below.

### 6.3.1 3D Inspection & Augmented reality

#### 6.3.1.1 Depth Based Damage Detection

This experiment will test whether the TACO system is able to provide a depth map with a defined resolution and in a time quicker than existing sensors.

The TACO system is positioned in a scene that is representative of a post fusion campaign vacuum vessel. The scene contains features over a large depth range. For distances up to 900mm an optical test breadboard will allow accurate positioning of vessel components. The TACO sensor is used to sample the scene in low resolution, high-resolution (specified minimum Ymm, Zmm resolution across the scene) and foveated mode. In foveated mode the sensor will be used to detect the presence of edges with a depth exceeding some threshold as regions of interest and to sample with the same minimum resolution (Ymm, Zmm) just in these regions (low resolution in rest of scene). Note that the resolution may be greater for some samples in the scene but must not be less than this specified minimum resolution. Averaging of depth measurements will be made to thereby increase the accuracy of foveated measurements.

Metrics Used:

- For ranges 0-900mm: The percentage and absolute differences between the known range values and the TACO range data when the TACO camera is in either low-res, high-res or foveated mode.
- The percentage difference in the times to sample the unfoveated (low and high resolution) and foveated scene data

The TACO advantages displayed are:

- Simultaneous high spatial and temporal resolution – demonstrated in the time to sample the data of interest in the scene compared to the unfoveated scan
- Large focus depth range – demonstrated by sampling a large depth range in the scene with a specified minimum resolution
- Improved range measurements – demonstrated by averaging depth measurements for increased accuracy.

#### 6.3.1.2 Point Tracking

This use case will measure how well the TACO system is able to track a moving object and provide accurate dynamic position information to allow augmented reality overlay of important hidden features.

The development robot will be used in the lab environment with a blank wall behind and will separately hold a number of objects representative of a vacuum vessel components. The CAD model of these objects will be known and made available to the TACO system, which it will use to track each object such that dynamic 6DOF position and orientation data can be returned. The robot will move the objects along a number of repeatable trajectories and orientations. This will be repeated with the TACO camera in unfoveated mode with both low resolution (fast scanning) and high resolution (with minimum resolution equal to that in foveated mode). The known position of the objects from the robot kinematics will be compared to the TACO data.

Metrics Used:

- The percentage and absolute differences of the object position from robot kinematic and TACO data when the TACO camera is in either low-res, high-res or foveated mode.

The TACO advantages displayed are:

- Simultaneous high spatial and temporal resolution – The scanning needs to be performed at such a rate to keep track of the moving item while still providing accurate information about the corner positions.
- Object Tracking

### 6.3.2 Public Safety

#### 6.3.2.1 Swab Localization – amount of background light

The first part of this experiment will be to determine how much background light the TACO sensor can cope with whilst providing acceptable data.). This maximum level of light will be reached when we have less than 6mm accuracy in the FOVEA, using oversampling, at 1m of the sensor, for a maximum std of 2. This is half the width of the smallest object we want to detect (the AAA battery).

We hope to show that the TACO sensor is able to work well with much more background light than the Kinect for example. Once we have this maximum level of light, we'll continue the experience experiments in these conditions.

| | Ground Truth | | Kinect | TACO (Unfoveated) | TACO (Foveated) | |
|---|---|---|---|---|---|---|
| | Position | Orientation | Deviation from Ground Truth | Deviation from Ground Truth | Deviation from Ground Truth | Gain from unfoveated* |
| Normal room lighting | | | | | | |
| Medium photo lighting | | | | | | |
| Full photo lighting or direct sunlight depending on conditions. | | | | | | |

**Gain from unfoveated: difference between the foveated and unfoveated deviation.*

#### 6.3.2.2 Swab Localization – different trajectories

We'll be using the maximum convenient light decided by the first part of the experiment.

The experiment will test how much accuracy is gained by foveating on the swab and compare it with the Kinect sensor and the TACO sensor in non-foveating mode. The experiment will be run using different approach speed and trajectories for the swab.

The scene features a suitcase that contains a suspicious object. Scene data is recorded, with the Kinect, with the unfoveated TACO sensor and with the foveated TACO sensor. The data is processed offline and used to estimate the position of the swab using a developed algorithm. In addition, the time necessary for acquisition is measured. The results are compared to the known position of the swab.

This will give the following table of results:

| | Ground Truth | | Kinect | TACO (Unfoveated) | TACO (Foveated) | |
|---|---|---|---|---|---|---|
| | Position | Orientation | Deviation from Ground Truth | Deviation from Ground Truth | Deviation from Ground Truth | Gain from unfoveated* |
| Simple linear slow trajectory | | | | | | |
| Faster linear trajectory | | | | | | |
| Faster more complex trajectory | | | | | | |

**Gain from unfoveated: difference between the foveated and unfoveated deviation.*

This will demonstrate the increase of accuracy when using the foveation, as well as the robustness of the TACO sensor in different light conditions.

#### 6.3.2.3 Detecting the Objects in the Suitcase

We'll be using the maximum convenient light decided by the first part of the experiment.

This will test the performance of the TACO system in more unstructured environments, where the shape of objects is not known prior.

The scene features an opened suitcase. The scene is recorded using the Kinect and with the TACO sensor foveating on gaps. The data is processed offline to estimate deviation from known position values. In addition, the time necessary for acquisition is measured.

This will give us the following table of results:

| **Ground Truth** | | **Kinect** | **TACO (Unfoveated)** | **TACO (Foveated)** | |
|---|---|---|---|---|---|
| Position | Orientation | Deviation from Ground Truth | Deviation from Ground Truth | Deviation from Ground Truth | Gain from unfoveated* |
| | | | | | |

**Gain from unfoveated: difference between the foveated and unfoveated deviation.*

This will measure the increase of accuracy when using the foveation in a more unstructured environment.

#### 6.3.2.4 Suspicious Object

We'll be using the maximum convenient light decided by the first part of the experiment.

This will test the precision of the foveation regions when doing bottom up foveation.

The scene features an opened suitcase containing folded clothing and a good view on a suspicious object (AAA battery). The TACO system gains foveated data on the suspicious object and the data is processed offline to estimate the correctness of the FOV.

We will quantify this through measuring the quality of the segmentation/foveation region detection. This will give us the following results:

- % FR (foveated region) in object
- % FR outside object
- % object in FR
- % object outside FR

This will measure the capability of the TACO sensor to foveate on the object of interest.

#### 6.3.2.5 Movement detection

We'll be using the maximum convenient light decided by the first part of the experiment.

This will test how fast movement is detected with the TACO sensor, while foveating on something else.

The scene will again feature an opened suitcase containing a AAA battery which the TACO sensor will foveate upon. An object will be moved in the FOV of the TACO sensor at a precisely known time and the time it takes for the sensor to report this movement will be recorded. The test will also use the Kinect sensor to gain comparison data simultaneously. Again, the precision of the object localisation will be measured.

This will give us the following table of results:

| **Ground Truth** | | **Kinect** | | **TACO (Foveated)** | |
|---|---|---|---|---|---|
| Position | Orientation | Time | Deviation from Ground Truth | Time | Deviation from Ground Truth |
| | | | | | |

This will measure the capacity of the TACO sensor to quickly detect movement in the unfoveated region, while foveating on and precisely measuring something else.

### 6.3.3 Home Grasping

#### 6.3.3.1 Self Localisation

There are two hypotheses being tested:

- The (raw) data of the TACO sensor produces localisation results (when fed into a state-of-the-art localisation algorithm) at least as good as from commercially available sensors

- The pre-processing done internally by the foveation software on the raw data (i.e. feature extraction) might help to achieve better results in cluttered environments by highlighting parts of the data that are relevant for localisation, e.g. vertical planar surfaces

The TACO sensor along with commercially available sensors will be mounted onto a mobile robot. The sensors will be spatially as close to each other as possible so that they capture the same part of the scene. The sensors' viewing direction is the motion direction of the robot and tilted towards the ground to capture the space directly in front of the robot.

The robot will be moved along a known trajectory and the odometry data, the TACO sensor data as well as data from the commercially available sensors will be recorded. For one trajectory (total length is about 50 meters) there will be two runs, one with the TACO sensor in non-foveating mode and one in foveating mode. After that, the recorded data (odometry and one range sensor's data at a time) will be fed offline into a self-localisation algorithm. The pose deviation from the ground truth will be evaluated – this is the standard approach for evaluating the quality of localisation in mobile robotics.

3D data can be gained at (approximately) the frame rate of normal cameras, but with the quality of a laser range finder (stereo and the Kinect have decreasing depth resolution with increasing depth, the SwissRanger has either very high noise or no useable data at all at distances above a couple of meters).

The pre-processing done in the TACO sensor helps on the user/robot side to distinguish task-relevant data from irrelevant ones, which saves processing time and increases the quality of the results (e.g. not using data points that correspond to clutter).

#### 6.3.3.2 Object Detection

The foveation capability of the TACO sensor enables more accurate object detection when compared to the fixed-resolution commercially available sensors and sooner object detection from greater distances.

Objects of five types will be placed on a table: cup, bottle, toy as well as two basic geometries (cylinder and cube). A robot with the mounted sensors will move towards the table and stand still in front of it for five additional seconds (to cover the aspects of dynamic and static sensor as well as varying object distances). The same will be done for a door scene, where the interesting object is the door handle.

Of these objects 3D models will be used as ground truth. Each scene will be recorded with the TACO sensor as well as commercially available sensors. The recorded point clouds (of one sensor at a time) will be fed into object recognition software that uses the 3D model as ground truth. The deviation of the geometry fitted into the point cloud data with respect to the 3D model will be evaluated (offline).

The saliency map as well as the scene segmentation done by the TACO sensor to compute the map can speed up processing on the robot/user side as (1) the task-relevant parts of the data are already highlighted and (2) scene segmentation is already available.

The fact that data points representing the objects are available at high(er) spatial resolution makes it easier to detect the geometry of the object at hand.

#### 6.3.3.3 Obstacle Detection

The foveation capability of the TACO sensor enables sooner and more accurate obstacle detection when compared to the fixed-resolution commercially available sensors.

Objects of various sizes and shapes (also including tables with protruding table top) will be put in front of a mobile robot that has the TACO sensor as well as other sensors mounted on top. The robot moves (from far to near) towards the object and the sensor data is recorded. The evaluation will be how soon each object in the robot's path is detected as obstacle (offline). The objects will be static and dynamic.

Given that the reference sensor itself will not be perfect, hand-annotate all obstacles in each sensor data set and use these as ground truth (i.e. individual ground truths per sensor data set).

The saliency map as well as the scene segmentation done by the TACO sensor to compute the map can speed up processing on the robot/user side as (1) the task-relevant parts of the data are already highlighted and (2) scene segmentation is already available.

The fact that data points representing the obstacles are available at high(er) spatial resolution makes it easier to detect them.

## *6.4 Use-Case Demonstrations*

In addition to the above mentioned experiments, we will conduct demonstrations. These are intended to convey a realistic sequence of events where the TACO sensor has a central role. These do therefore not provide quantitative information, but rather demonstrate the sensor in practical use.

### 6.4.1 Public Safety – Full Demonstration

The public safety experiments discussed above will be combined to visually demonstrate the process of using the TACO sensor to allow remote operations to deal with a suspicious object.

A scenario will be setup that features an open case containing a suspicious object. Using TACO sensor data a Shadow robot hand will semi-autonomously find and remove the object. First the sensor data will be used to identify the open bag. It will then foveate on a swab which is being passed to it by a human. The robot will take the swab, return to the case and use this to sample the open case and contents (simulating the checking for explosive materials). The swab will be returned back to the human. The sensor will then be used to analyse contents and locate the suspicious object in this unstructured environment. The suitcase will contain other items, such as clothing, which the robot hand will move in order to locate the suspicious object. Once located the sensor will foveate on the object of interest and the robot hand will remove it from the suitcase.

First we'll determine the acceptable levels of background light, proving the resistance of the sensor to varying lighting conditions. While foveating on the swab, we'll focus on testing the accuracy and speed of the TACO sensor. The TACO sensor will continually monitor the full FOV and will detect movement around the suitcase as an additional safety measure demonstration. Any detected movement will immediately pause the robot motion.

### 6.4.2 Home grasping – Full Demonstration

The TACO sensor can be meaningfully used in the context of a home robotics scenario. Due to its foveation capability and large field of view, it can replace several conventional sensors that would be used simultaneously otherwise.

A typical scenario from a home robotics setting will be showcased. A mobile robot with the mounted TACO sensor moves through a home-like environment. The sensor data will be used for navigation and object detection. It will be shown that the sensor is capable of foveating on relevant regions of the scene depending on the current task.

The robot starts in one room of this environment, moves along a narrow corridor, encounters a closed door and waits until it opens. After that it continues its tour into the entry hall and into the kitchen, where it will stop in front of a counter with several objects (among them a cup) on it. In the course of this tour the task based on which foveation is done will be switched between navigation (in free space) and object detection otherwise (objects of interest are "door handle" in front of closed door, "cup" in front of the kitchen counter). Switching between these two tasks is done by the robot/user side.

The demonstration should result in a video showing the robot with the TACO sensor mounted onto it moving through the environment as well as the results of the foveation software's processing results (saliency map, label map and 3D data).

The idea of the demonstration is to show that the TACO sensor can be used for the various tasks of a home robot i.e. navigation and object detection, especially, that a single sensor can do what usually requires at least two different sensors.

### 6.4.3 Augmented Reality & Task Safety – Full Demonstration

This will demonstrate that the TACO system and robot system can allow improved and safe performance using augmented reality

The experiment will take place within a fusion vessel mock-up and will include an operator controlled development robot for task manipulation. The operator will be performing a simple assembly task, installing a component onto the vessel wall. The four manufactured mock-up vessel tiles (discussed in 6.3.1.1 Texture Based Damage Detection) will be mounted to the vessel wall mock-up with the assembly site located in the centre. The component will have assembly features on the rear which must interface with features on the vessel wall (i.e. dowels and holes). Due to space constraints and camera positions the assembly features are not always visible but are necessary to allow the operator visual feedback of the relative position of the component with respect to the vessel wall. The TACO camera will be mounted at a fixed location and will be used to track the object and with the robot system will overlay an AR representation of the features. As in the 6.3.1.2 Point Tracking use case, markers will be positioned on the robot end-effectors to allow effective tracking. The TACO system will scan the full field of view in low resolution. This will be done both to determine the regions of interest (where a high resolution scan will take place) and also to detect when a possible collision may occur and both inform the operator (using AR overlay) and halt the robot motion.

The TACO advantages displayed are:

- Simultaneous high spatial and temporal resolution – The TACO system will provide accurate data on the moving object while also monitoring the full scene to detect motion.
- Adaptability – When a possible collision is detected the TACO system will adjust its foveation to obtain more accurate data as required.
- Object Localisation and Tracking

# 7 Implementing, Disseminating and Exploiting

## *7.1 Implementing the Strategy*

The TACO project engages three internal Bodies (General Assembly, Executive Board, Project Management Team) which are supported by an additional management body, the Industrial Advisory Board, consisting of 3 selected experts, not directly involved in the project as partners.

These three selected European organizations assist and advice the consortium in order to strengthen the influence and engagement of public corporations and organizations within the TACO project. These organizations possess, however, considerable know-how on the matter and will help to set targets and act as a feedback group to the project consortium. They provide an external unprejudiced view without receiving significant funding from the European Union (only minor effort and travel costs will be reimbursed). The final selection of members for the Advisory Board has taken place during the Kick-off-Meeting. The following organizations have confirmed their interests to guide and support the TACO consortium with advice and expertise throughout the project duration:

- Robert Bosch GmbH – Autonomous Systems and Robotics (Germany)
- Leuze Electronic GmbH & Co. KG (Germany)
- Tordivel AS (Norway)

The TACO project therefore gathers a multi-disciplinary perspective that brings together the relevant disciplines in the fields of robotics and sensors as well as process design experience.

## *7.2 Disseminate the findings*

Within the first few project months the TACO consortium elaborated a plan of concrete dissemination activities. An overview of the potential **dissemination** contributions of the project partners is given below:

**TEC**: TEC provides the TACO-project IT-infrastructure – more precisely the whole set of tools which will foster the project cooperation, communication and dissemination, whereby the project website will serve as the most versatile external information and communication tool for a worldwide audience. In addition TEC elaborated a TACO-project leaflet as well as a press releases together with the other partners.

**OTL:** In terms of dissemination, papers will be submitted on the results of the testing to suitable conferences and journals. The results and the project will also be discussed with suitable contacts in both industrial and academic fields.

**TUW:** Main dissemination is scientific through conferences (e.g., ICRA, IROS, CogSys, SciCog, ICVS, ISVC, ECCV) and journal publications (e.g., MVA, IVC, CVIU).

**FHG:** Conferences and Publications: Research results will be submitted to scientific

- FHG plans to present the results related to the MOEMS technology at the following scientific conferences: SPIE Photonic West Symposium and Exhibit (San Jose, USA), SPIE Photonics Europe Symposium and Exhibit (Strasbourg, France), Transducers, Eurosensors, IEEE MEMS, IEEE/LEOS Optical MEMS
- IEEE/ASME, Journal of Micro-Electro-Mechanical Systems: Micromechanics, microdynamical systems, microfabrication technologies useful for MEMS, modelling and design issues for MEMS, MEMS characterization and reliability.

- IOP, Journal of Micromechanics and Microengineering: Covers all aspects of microelectomechanical structures, devices and systems, as well as micromechanics and micromechatronics.
- SPIE, Journal of Micro/Nanolithography, MEMS and MOEMS: Development of lithographic, fabrication, packaging and integration technologies for microoptoelectro mechnical (MEMS and MOEMS), and photonics industries.
- Elsevier, Sensors and Actuators A: Focuses on the research and development of solid-state devices for transducing physical signals.

**SINTEF:** SINTEF will disseminate project results through publishing project information on the company web site, scientific articles relating to research within 3D foveation algorithms and communicating the project's concept and results to new and existing customers. Project results will further be exploited in current and future research projects undertaken by SINTEF.

**CTR:** CTR will present results of the TACO project at different conferences and scientific workshops with relevance to MOEMS based sensor devices; in particular, it is planned to participate at SPIE Photonics West, covering the latest enabling technologies and applications for micro- and nanofabrication and Photonics Europe, an appropriate European conference of SPIE with one focus on Micro/Nano Technologies Metamaterials, Devices, MEMS, MOEMS, Nanometrology. CTR will present results of the TACO project in line with the planned presence on international fairs with topics on Microtechnology and sensors.

**SHADOW:** Shadow is already in discussion with end-users as to the possible applicability of the TACO sensor to their commercial problems, and will disseminate as appropriate TACO results through their customer mailings and contacts.

## *7.3 Exploit the results*

The **exploitation** of the project results is clearly defined in the objectives of TACO. As the project consortium consists of major European players in both science and industry the usage of the results will be exploited in both the science and commercial sector. The main exploitation will be through each partner's own organization.

**TEC**: Experience gained will be funneled into our industrial services on requirement engineering. As an emerging SME, the reputation gained from the project will positively influence our future acquisition activities. TEC provides workflow based management support systems for cooperative research efforts on national and European level. Project experience will trigger improvements of TEC's "Trusted knowledge Suite". Any novelties introduced will elevate the market position of this IT tool.

**SHADOW:** Shadow is developing advanced robotic technologies based around manipulation (see e.g. FP7 Project HANDLE) and is keen to bring high-grade 3-d vision technologies to bear in this area. Shadow will be looking at ways to integrate the TACO sensing technologies into their existing product developments, as well as reaching out to current and new partners to explore routes to market for the TACO sensing system.

**OTL:** Given a successful outcome of the TACO project OTL will integrate TACO systems into their own products where suitable, offer the solution to clients wherever benefits could be seen and pursue the supply, sale and integration of the TACO system itself.

**TUW:** In terms of exploitation our main strategy is to exploit the gained knowledge to stay at the forefront of vision research for robotics in Europe and to exploit specific results for spin offs.

**FHG:** The results of the TACO projects will be published in relevant scientific and technical journals as well as via Internet. Furthermore the results will be presented on several conferences linked to technology and application of MOEMS devices. In agreement and cooperation with our partners first of all SINTEF, IPM and CTR the results (e.g. 3D sensing system) will be presented on international fairs (e.g. Photonics West, MEMS/Micromachining (Tokyo), Laser).

Furthermore, effective implementation of the strategy will include:

- Good communication between all partners, particularly between those of different technical backgrounds, to produce good quality specifications
- Enhancing the IAB by inviting more members
- Integrating a user forum on the TACO web site
- Releasing software demonstrators and open source software libraries for further R&D
- Information gathering about the industry and academic research groups
- Direct contact with industry
- Demonstration of results of scenario testing to clients operating in areas of application

Clearly, two aspects are of immanent importance for the TACO strategy implementation. Firstly, the communication and cooperation within the consortium; and secondly, the contact to the relevant industry, especially potential end-users of the TACO sensor.

## 7.4 Risks and Contingency Measures

For an exact breakdown of risk categories, please refer to the risk matrix on page 39.

| Risk Nr. | Description | Probability (SM \| MD \| LG) | Impact (SM \| MD \| LG) | Contingency plan | Risk category |
|---|---|---|---|---|---|
| 1 | Sensor does not interface with robot software. | SM | LG | Create well defined sensor requirements fully understood by all parties. | Yellow |
| 2 | Sensor not really usable due to bad specifications. | SM | LG | Good communication and putting good efforts in the use cases and the specifications. Update the specifications in light of the new information. | Yellow |
| 3 | Sensor not as good as expected/WP5. | MD | LG | Try to have a first sensor as soon as possible. | Red |
| 4 | TACO sensor delayed/WP3. | MD | SM | Algorithm development will be based on data from commercially available scanners. | Yellow |
| 5 | Delays in scenario implementation due to unforeseen problems with implementation. | MD | MD | Testing of robot software with some representative existing hardware during development (i.e. far in advance of TACO system delivery to end users). | Yellow |
| 6 | Difficulty of integrating the software with current hardware. | MD | MD | Good communication with WP2 – software team. | Yellow |
| 7 | Delays in programming the tests for TACO hardware. | MD | MD | Have access to the hardware as soon as possible, or in a first time to a simulated stream of data. | Yellow |
| 8 | Late discovery of inconsistent system specification from WP1/WP2. | MD | MD | Design studies planned in WP2 have to be performed early to make sure that | Yellow |

| | | | | | |
|---|---|---|---|---|---|
| | | | | competing specifications are not contradictory. If necessary WP1 results must be revised to find acceptable compromises. | |
| 9 | Relevant laser scanner for making simulation data is too far from the developed camera in specifications/WP4. | SM | MD | Thorough review of specifications in WP1 before choosing simulation scanner and scenes to use for test sets. | Yellow |
| 10 | Time consumption in real-time control loop /WP4. | MD | MD | Separate in a fast data acquisition control algorithm and more complex and time consuming cognitive foveating algorithm. Implement parallel processing (multi-core or GPU). | Yellow |
| 11 | Usability of the data for end users, both human programmers and machines/WP4. | SM | MD | Deliver a traditional 3D video stream from the sensor data. | Yellow |
| 12 | Insufficient detection performsance of 3D feature extraction algorithms. | SM | MD | Through literature review and continuous benchmarking of features. | Yellow |
| 13 | MEMS mirror device delayed. | LG | MD | a) close follow up of Fraunhofer IPMS pushing for early results and an early test of critical parameters.<br>b) use commercially available scanners to collect high resolution data that can be used for alternative testing. | Red |
| 14 | MEMS mirror device with lower capability than promised. | MD | MD | a) perform tests with lower capability (lower speed/resolution than intended).<br>b) identify alternative applicable bidirectional scanning mirror designs. | Yellow |
| 15 | Unsynchronised I/O scanning optics. | MD | LG | a) close collaboration with an open dialog and a focus on quality control between Fraunhofer IPMS, IPM, CTR .<br>b) perform optical I/O through the MEMS mirror with reduced aperture. | Red |
| 16 | MOEMS design, since coaxial design is impossible, synchronization between reception and emission mirror (if applicable). | MD | LG | Problems are minimized when identical designs are used for the MOEMS (or when the reception branch is mirror free, synchronization must be achieved using electronic control in both branches. IPMS must provide sufficient control methods (electronical/optical). | Red |
| 17 | MOEMS design delay/WP2. | MD | MD | IPMS must commit to green electronic/geometric properties for work in WP3 to proceed in time. | Yellow |
| 18 | MOEMS design: dynamical mirror medium properties are not met in final design. | MD | LG | Perform appropriate mechanical simulation (up to M17), iterate system specifications necessary. | Red |
| 19 | TACO camera system unit delayed. | MD | LG | a) close contact and rapid feedback between partners on possible delays.<br>b) use commercially available scanners to collect high resolution data that can be used for alternative testing. | Red |
| 20 | Disturbing measuring artefacts from 3D camera/WP4 | SM | MD | Correction of image using knowledge of the measurement principles. | Yellow |
| 21 | Mismatch between protocol capabilities and camera input/output needs, e.g. clock and data streams. | SM | MD | Detailed planning of input and output needs before interface implementation starts. | Yellow |
| 22 | Delays in HANDLE project. | MD | MD | Keep in touch with our HANDLE partners and do our best to keep up with the planning. | Yellow |
| 23 | Targeted power consumption or size of TACO camera are significantly exceeded. | MD | MD | For demonstrator, use more powerful supply unit or split system into several units. Review of specifications. | Yellow |
| 24 | TACO camera consists of two units: optical head, electronic unit. | MD | MD | Review of specifications, redesign. May be advantageous. | Yellow |
| 25 | Costs of implementation efforts for sufficiently performing TOF measurement method significantly exceeds budgetary restrictions. | MD | MD | Using commercial components and modules. Must possibly revise target specifications or fallback to a design with larger effective aperture. | Yellow |
| 26 | Targeted price range for TACO camera are significantly exceeded. | LG | MD | Using lower-priced components. Cost-effective implementation of TOF | Red |

| | | | | | |
|---|---|---|---|---|---|
| | | | | measurement method. Must possibly revise target specifications. | |
| 27 | Laser Source: target specifications on beam quality. Laser power (measurement range), necessary control of pulse timing and pulse form can only be reached with fiber amplifiers. Cooperation with fiber amplifier producer is required to provide appropriate light source. Risks: target price range of TACO exceeded. | LG | MD | First priority: obtain a laser to get a working demonstrator. Second: the laser producer should agree that laser source prices will be reduced to an acceptable level (several 100 EUR) if the number of units ordered increases to the anticipated level. Fallback option can only be to negotiate with alternative suppliers. | Red |
| 28 | TOF performance lacking in close-up range. | MD | SM | Careful design with optical simulations, especially concerning measures to increase S/N in close-up range. Fallback design using separate emission mirror, prism in optical window (possibly further special optical components). | Yellow |
| 29 | Insufficient performance of synchronized driver. | MD | LG | a) close dialog between IPMS, IPM, CTR.<br>b) tests at early stages with available devices.<br>c) evaluation of different implementation possibilities for fall back options. | Red |
| 30 | Delay of synchronized driver. | SM | MD | Close dialog between IPMS, IPM, CTR in order to have detailed specifications, simulations and tests at an early stage. | Yellow |
| 31 | Too low overall yield of 2D MEMS array related to synchronizing. | MD | LG | a) MEMS design is focused to enhance the frequency band width at nominal deflection to compensate frequency tolerances caused by the DRIE etching process. The influence of DRIE fabrication tolerances on frequency bandwidth is more relevant for the resonant fast axis scanning. In case of the slow axis scanning tolerances can be compensated due to the quasistatic driving principle.<br>b) Minimization of frequency tolerances by optimization of DRIE etching processes. For TACO DRIE processes with alternating RF plasma are used enabling reduced tolerances of the etched shallow trenches in comparison to a DRIE etching with DC-coupled plasma resulting in lower frequency yield of MEMS scanner technology. DRIE process and MEMS yield can be enhanced further by using a low frequency plasma generator.<br>c) Alternatively to a monolithic MEMS array realization of this 2D mirror array by hybrid system integration of single 2D-MEMS dies. Selection of individual dies is based on wafer level characterization enabling a fault tolerant operation point of synchronization<br>d) Reduction of nominal FOV will significantly enhance frequency bandwidth for synchronized MEMS operation, optical FOV can enhanced again by passive optical elements but will reduce optical resolution | Red |

Table 3 Risks and Contingency Plan

*SM=Small, MD=Medium, LG=Large

## RISK MATRIX

| | | | | | | |
|---|---|---|---|---|---|---|
| IMPACT | **Very serious** | E1 | E2 | E3 | E4 | E5 |
| | **Serious** | D1 | D2 | D3 | D4 | D5 |
| | **Moderate** | C1 | C2 | C3 | C4 | C5 |
| | **Minor** | B1 | B2 | B3 | B4 | B5 |
| | **Minute** | A1 | A2 | A3 | A4 | A5 |
| | | **Very small** | **Small** | **Medium** | **Large** | **Very large** |
| | | **PROBABILITY** | | | | |

| **Colour** | | **Description** |
|---|---|---|
| Red | | Unacceptable risk. Immediate steps to reduce the risk. |
| Yellow | | Under constant consideration. Revision of risk at fixed intervals. |
| Green | | Acceptable risk. Annual revision of risk. |

Table 4 Risk Matrix

## 7.5 Work Package Detailed Plans

The consortium has started with an implementation plan as described in the proposal. The times foreseen for the different WP and tasks were derived during discussions with all partners in the preparation phase. Once started, the project plan (Figure 8) was updated according to new knowledge (proposal preparation and proposal start differs by 9 month) all plans were evaluated and again approved.

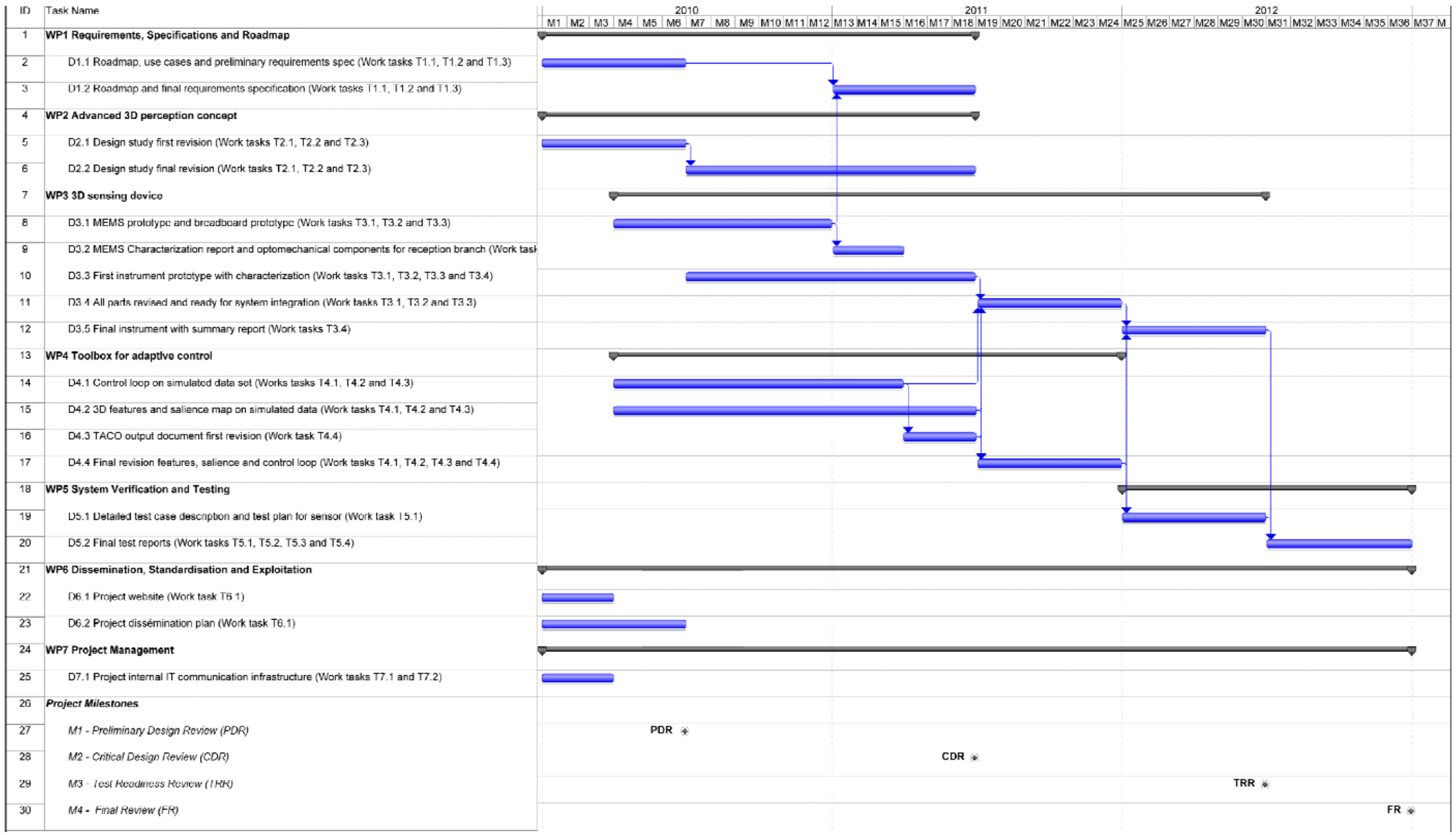



Figure 8 TACO Project Plan

At the first review meeting in April 2011, it was indicated that due to the challenging hardware requirements, the TACO sensor would be significantly delayed with regards to the original plans. This meant that the time originally scheduled for evaluating the actual hardware in end-user scenarios would be strongly limited. The consortium could justify, that with a project extension we would be able to draw final conclusions on sensor performance, report better results, reduce the risk and increase impact and ease exploitation. Therefore a project extension was

In Month 21 the Amendment No. 1 to the GA was officially approved by the EC, which included a project extension of 6 month. A new implantation strategy was derived and set forward in a new project plan. This currently active project plan is shown below (Figure 9).

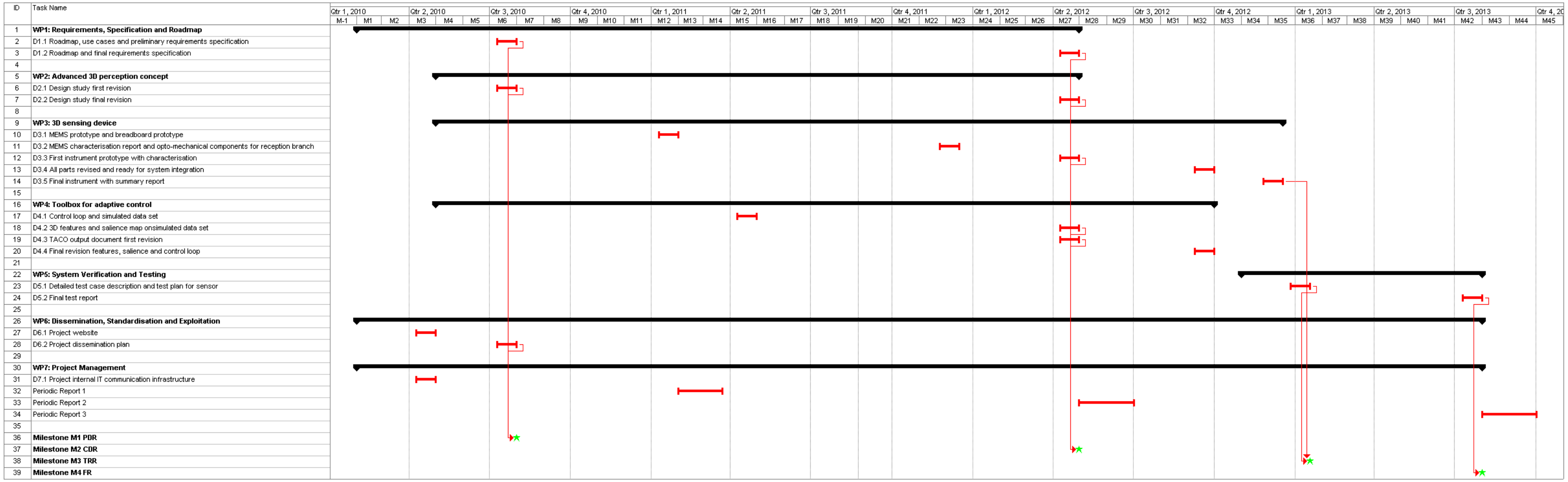



Figure 9 Updated TACO Project Plan after Amendment No.1

# 8 Technology Roadmap and Future Implementation

This chapter focuses on the interface region of technical TACO sensor capabilities and market needs and acceptance. It reveals that the TACO technology has the potential to fill an emerging market gap and shows clearly the need of further technical refinements.

Our main conclusions and suggestions for future technology rollout are:

- **Major technology gap is within outdoor object interaction:** Especially after the KINECT[1] sensor was rolled out, many of the sensor requirements for indoor interaction are now be solved at a reasonable price and performance. However, with the exception of existing laser range measurement devices, there are no sensors that can provide quality 3D range image data outdoors.
- **System alternatives must be considered for a mass-market version:** The current sensor prototype, being in a prototype and proof-of-concept stage only, is too big and expensive for an immediate market rollout. We have identified alternatives that can be feasible for mass-market commercialisation, but these require system-level redesign. These alternative technical solutions and new system level concepts exist and will foster TACO sensor successor developments.
- **Foveation makes clear sense for object interaction,** but **requires continued focus on ease-of-use.** Numerous applications of foveation have been evaluated through the project, and object recognition and interaction is the application where it seems to bring most benefit. However, to achieve wide-spread usage, it is important that it is easy to configure what kind of objects that are interesting, and that the data output by the sensor easily can be used.

## *8.1 Methodology and Objectives*

In the initial phase of the work done for this chapter relevant questions were developed, which were discussed by all participants of the project on both the operator's as well as the developer's side.

The discussion contained important and critical questions concerning the market situation, cost and performance of the system and components. Possible improvements and application fields relevant for the TACO sensor were discussed. A reflection in respect to competitive 3D sensors was done throughout of the whole process.

During the project period, the KINECT sensor hit the market. As a consequence of its high penetration and success in the mass market sector it was added to our area of investigation.

Our aim is to provide aid and mechanisms to forecast technology development and to provide a crude framework to access and support future technology evolvements. The work is significant to understand the current situation of costs and performance and also to define future cost and performance goals. High stakes in the 3D sensor area are the reduction of costs, size and weight as well as robustness, high performance and capability.

TACO introduces 3D foveation as an important concept for service robot interaction with their natural environment. This new concept will allow robots to interact with everyday environments in a more natural and human-like manner, increasing the level of detail whenever needed for interaction between the robot and everyday objects and humans.

[1] KINECT: Trademark from Microsoft for 3D gaming sensor

The 3D sensing system aims to be fast, small, lightweight and relatively energy-efficient to facilitate use in real-time operations and efficient and practical mounting on a service robot or a robot arm.

In the figure below, the envisaged goals are displayed in respect to each other. The heights of the different columns (costs, performance and size) indicate the relative improvements strived for.

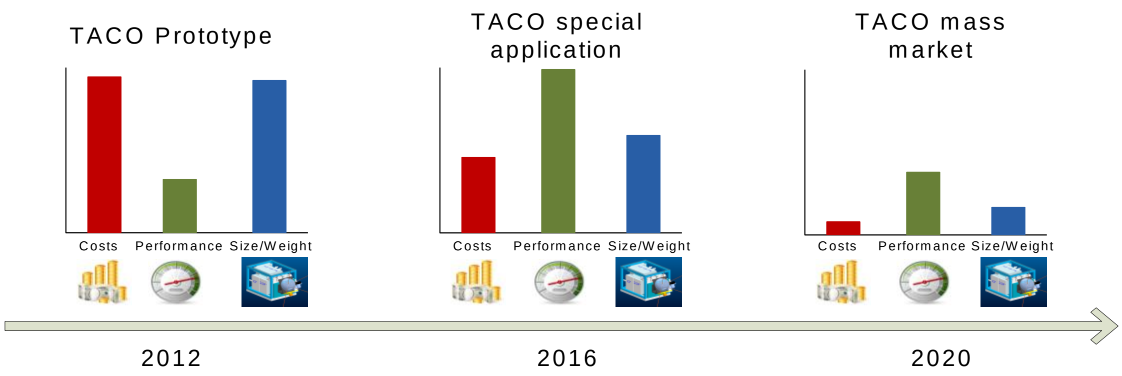



Figure 10 Illustration of the investigated timeline for further developments

The centre of gravity of our work was to investigate application fields for the 3D TACO sensor, to research the current and future market situation with their barriers/limitation for 3D sensors, to identify realistic cost and performance goals at component and system level, and to identify opportunities for technology improvements.

To sum up, the chapter was developed to describe improvement opportunities giving the current status of the TACO sensor. This includes opportunities that exist for technology and specific activities needed to reach the targets of the TACO sensor for providing better spatial and/or temporal resolution and being smaller and cheaper than existing 3D laser scanners. The results will be used for identifying, selecting and developing technology improvements/alternatives to satisfy the targets of the TACO sensor and also identifying relevant future trends for 3D laser scanners.

This section is structured into the following main sections:

- Current Market situation and future prospects
- Sensor System
- Market Opportunities
- Technology improvement opportunities
- Implementation beyond the project lifetime

## *8.2 Current Market situation and future prospects*

This chapter addresses the market situation for different types of commercially available 3D sensors and we will analyze the opportunities and challenges for 3D foveation technology.

### 8.2.1 Sensor Technology and Commercial Availability

Nowadays, the market of 3D sensors is emerging and growing very fast. The state of the art of 3D sensors can be divided into two main tracks: Triangulation and/or Time-of-Flight (TOF) measurements.

**(A) Stereo cameras** require the presence of texturing targets and proper lighting conditions to work. Depending on the stereo algorithm, very high computational power is required to achieve camera frame rates. There is a wide range of possible implementations: two or multiple camera setups, a wide range of resolutions, with or without colour, close to far range, narrow or wide angle, conventional cameras plus a conventional PC to run the stereo algorithm or embedded, pre-calibrated stereo systems where everything is in a common enclosure. Depending on the solution, prices vary from a few hundred to many thousands of Euros, but can still be lower, judging from the pure hardware costs. The sensor size mainly depends on the stereo baseline required by the application. The interface and power requirements are low (USB or FireWire), the weight is (typically) low as well. Stereo cameras are sensitive to lighting conditions, dirt and mechanical stress. An advantage is the availability of colour information as well as 3D data. Also the fact that stereo systems are passive sensors can be of importance for some applications. Stereo cameras can be combined with structured light approaches to deal with texture-less environments (see below). Due to their sensitivity to lighting conditions stereo cameras are usually not usable for safety-critical applications. The market for stereo cameras became harder after the introduction of the KINECT, but with cheaper camera modules and increasing (and cheap) embedded computational power their attractiveness will rise.

**(B) Imaging structured light sensors** consist of a projector that projects one or more known patterns of light onto the object to be measured and a camera to register the projected light. The object surface is reconstructed from trigonometric considerations similar to those on which stereo cameras are based however using the advantage of a priori knowledge of the projected pattern. The accuracy of structured light systems depends on the baseline between projector and camera and their respective resolution. Under controlled conditions the depth resolution and accuracy can be down to 10-5 of the extent of the measurement volume. Depending on the application and requirements, these systems are portable or stationary and may be priced up to several 10 000 EUR.

Structured light sensors scale well up to measurement volumes of several meters extent, the Microsoft KINECT being the latest example of a mass-market device for such measurement volumes. The KINECT uses structured light, but there is only one dot pattern projected (using diffracted infrared laser light rather than various stripe patterns of visible light). The typical depth resolution of the KINECT is in the centimeter range, but with its low price tag of around EUR 100, the price/performance ratio is currently unbeatable.


Figure 11 KINECT sensor from Microsoft

Other advantages of the KINECT are the high frame rate, small size and low weight. A drawback is the limited FOV of only 58° x 45°. Altogether, the KINECT is suitable for mobile/service robotics (cleaning, navigation), object grasping (relatively big objects) and detection, human-robot interaction and also motion capture.

**(C) Time-of-flight cameras** emit modulated or pulsed (infrared) light to illuminate the scene and detect the returned light using a focal plane array (FPA) similar to regular digital cameras, which enables then to measure the time the emitted light needs to reach each pixel of the detector. Since they illuminate the whole scene, the detected light is of low intensity and the measurement noise is large. TOF cameras suffer from background light interference, and depending on the operating principle, the measurement distance is either limited or far objects are mapped into a smaller uniqueness interval. The FPA sensors require careful calibration to eliminate temperature drift or to compensate non-linear intensity dependent effects. Additionally, parasitic reflections in the image space may cause the measured distance to be greater than the true distance. Due to the specific sensor IC, the bottom line price for TOF camera hardware is typically above that for triangulation based systems.

**(D) Line laser scanners** consist of a laser range finder mounted onto a rotary axis or a stationary range finder with intervening rotating or oscillating mirror. We obtain a 3D-imaging device by rotating this unit around an independent perpendicular axis or by deflection using a second mirror. The range finder uses active illumination typically based on a laser beam which illuminates only a small spot of the object surface. Laser scanners have long ranges and they are very robust with respect to lighting conditions. The prices range from about 1.000 EUR to several 10.000 EUR for high-precision surveillance-grade instruments and device size ranges from rather small devices (e. g. Hokuyo, < 1 kg) to larger ones (e. g. SICK LMS, several kilos). Laser scanners are commonly used in industry, so they are designed to deal with harsh and unfavourable environments. There has been a development towards cheaper and smaller scanners, so this trend will go on. The disadvantage of these macroscopic scanners tilting laser scanners is the comparably low rate for full frames, since tilt periods are of the order of several seconds. However, laser scanner data has been the data of choice in the robotics community and for reliable 3D data outdoors there is no technology of equal performance. The use of laser scanners indoors decreased when the KINECT became available, but they are still widely in use (e.g. on WillowGarage's PR2[2]).

**(E) MEMS based laser scanners** have only recently entered the market, driven by advances both in MEMS availability and TOF technology. They address the main shortcoming of mechanically driven laser scanners by introducing fast and robust micro mechanical mirror elements to achieve sub-second or faster frame rates. Pulse TOF based scanners have been realized based on two dimensional resonant scanner modules (e.g. Nippon Signal, http://www.ecoscan.jp/). However, the available mirror aperture is the limiting factor for depth accuracy and beam divergence. The TACO sensor is such a micro-mirror based scanner whose mirror elements are optimized for sufficiently fast scanning. It is able to provide video frame rates with images of 50000 voxels each. In one direction the mirror motion can be controlled in a vectorial manner, thus permitting to freely specify voxel density. Sufficient aperture is achieved by combining several mirror elements into one array. TACO 3D point measurement precision will be in the 2 to 5 mm range. The precision is comparable to high end survey scanners.

The following table is an (incomplete and simplified) representation of the technical merits of the competing technologies:

[2] www.willowgarage.com/pages/pr2/overview

| | TOF camera | TOF scanner | Stereo vision | Structured light (projection + camera) |
|---|---|---|---|---|
| Example | MESA | Sick / **TACO** | | KINECT, z-snapper |
| Cost | low – medium | (medium -) high | Low | low (– high) |
| image data | Intensity (b&w) + depth | Intensity (b&w) + depth | colour + depth, external light req. | depth, colour by supplemental camera, colour requires external light |
| (B&w) image available | Lighting independent, up to measurement range | Lighting independent, up to measurement range | If background light level suffices | If background light level suffices |
| Depth image available | up to measurement range | up to measurement range | If background light level suffices and object structure permits up to a distance corresponding to base line | If background light level sufficiently low for projection wavelength up to a distance corresponding to either base line or detection limit for illumination |
| Measurement range (3D) | (low-) medium | High | medium | medium |
| voxel rate | up to 1..5 x $10^6$ | $10^4$ – $10^6$ | up to 1..5 x $10^6$ | up to 1..5 x $10^6$ |
| Data repeatability/ accuracy | 1-10 cm range, distance dependent / difficulties controlling additional systematic effects | 1-10 mm range, weakly distance dependent / systematic effects mostly under control | No depth information in occluded regions / data quality depends on identification of corresponding points in different images, strongly distance dependent | No depth information in occluded regions |
| environment | indoor/(outdoor) | indoor/**outdoor** | indoor/outdoor | indoor |

Table 5 Technical merits of competing technologies

Within the table above we placed our development within the current 3D sensor market. Its intention is to give a crude guideline to set the technological baseline for our proceeding chapters.

From a technical point of view, it is important to know which application fields are interesting or relevant for a product to be well positioned. Thus, the main application fields were discussed and identified by the project team and listed in the table below.

| Fields of application | Suited 3D sensors |
|---|---|
| Remote operation | Stereo vision |
| (support by augmented reality: | TACO sensor, laser scanners) |
| transport and logistics | TOF: Sick LMS / Mettler Toledo Cargoscan |
| Industrial automation | Stereo vision |
| Mobile / service robotics<br>(navigation, cleaning) | KINECT<br>**TACO sensor**<br>Combination of lasers / 3d cameras<br>TOF: MESA |
| Object detection | KINECT<br>Stereo vision<br>**TACO sensor**<br>Tilting laser scanner |
| Human-robot interaction | KINECT<br>Stereo vision<br>**TACO sensor** |
| Autonomous operation / interaction<br>(indoor and outdoor) | **TACO sensor**<br>Velodyne 360 x 30-50°<br>Lidar |
| 3D inspection | Structured light<br>**TACO sensor** |
| Object grasping | Stereo vision<br>**TACO sensor**<br>Structured light<br>Laser-scanning triangulation<br>TOF-cameras (low-res)<br>KINECT<br>Tilting laser scanner |
| Motion capture (humans) | KINECT<br>TOF-camera<br>**TACO sensor** |
| Unknown/dynamic environment application | **TACO sensor**<br>Combination of laser scanners and 3D cameras |
| Outdoor object interaction | **TACO sensor** |

Table 6 Fields of application of different 3D sensors

The TACO sensor is represented in most of the application fields but also competitive 3D sensors were considered, as it can be seen in the following table. Within each field, approximate rankings based on suitability and performances have been given by the project team.

### 8.2.2 Gap Analysis: Opportunity and Risk for Foveation

Each 3D technology has its specific merit and each technology spans a large price range depending on the sector of the application. The cheapest widely available 3D sensors are based on triangulation principles. These sensors have already entered the home consumer mass market (e.g. KINECT) and strong growth rates are expected in the near future. The weakness of these sensors is in data quality. For stereo vision systems, on the one hand, characteristic points must be identified in both images which does not work on featureless (technical) surfaces or if parts of objects are occluded in one of the images. On the other hand, structured light systems do not suffer from problems on featureless surfaces, but only work as long as the projected pattern can be identified which competes with other light sources. The accuracy of all triangulation techniques decreases rapidly with distance from the sensor pair.

Demand for higher quality, increased reliability, robustness, lifetime and the option of outdoor operation arises in industrial applications as autonomous motion and object interaction in open spaces. Further fields of application, possibly outdoors, include

- robot-robot and human-robot interaction;
- object manipulation;
- visual inspection;
- dynamic obstacle detection; and
- advanced scene understanding.

TOF technology has the potential to deliver high quality data for outdoor (and indoor) applications. Rotating arrays of pulse range finders have been successful in environment mapping and autonomous vehicle control (Velodyne HDL-64/32). Their resolution and precision is however not sufficient for object interaction. TACO adds foveation to TOF technology, thus the possibility to trade between precision, resolution and speed as required and is thus clearly suitable for object interaction or for dynamically changing situations where high speed is important.

Though technically superior in distance determination, scanned TOF technology suffers in important aspects: cost, size, weight and imaging speed of the sensor. These issues must be addressed to improve the marketability of scanning sensors. The costs are coupled to or scale with the offered functionalities and the quality of the acquired data. Market-dominating brandings are frequently challenged by low volume competing technologies. For large volume sales there seem to be a virtual price limit at a few hundred Euros.

Such price target is not currently feasible for scanning solutions. In order to close on such a price segment several tracks should be followed up:

First of all, the laser source cannot be eliminated without sacrificing performance. From the electronic point of view (using ASICs) costs will be saved, but only at the expense of heavy development costs. A price tag of the KINECT or two webcams with software will never be reached by these sophisticated sensors. The mass market, like gaming applications in 3D, is under full control of cheap technologies.

Another track for reducing the system costs is to employ a single mirror device with diode-pulse TOF at lower precision (beam size 5 mrad and 1-2 cm stdev). Reduction of range (1-2 m) is in line with certain applications (e. g. agriculture), but a price target of 1.000 EUR seems very tough here as well.

Regarding robustness, performance and capabilities it is preferable to target the demands of the larger part of potential customers than of only a few ones with special requirements for a 3D sensor. Robustness will always be an important factor for industrial applications. If the system cannot survive vibration, noise, dust and dirt, as well as electrical interference, it will not be used in this sector.

Further concerns to 3D sensors are physical limits on mirror size determining the smallest size for a certain resolution. High precision scanners will always have a certain size and volume.

A reduction of the system volume can be reached by partly or fully integrating the electronic components: mirror controller, embedded digital electronics and TOF electronics. In the course of such a shrink process the volume will come down, but currently cannot be reduced by more than 50%. Even so, for the 'industrial' market segments, this is probably acceptable (cf. Sick's scanners).

A question regarding foveation is how much existing software for 3D data has to be modified to gain any benefit from foveated data. If the power requirements are high or special interfaces are required, potential users will refrain from buying the sensor. Concerns of eye safety and active sensing principles also lower the attractiveness of the sensor for large variation of applications.

In contrast to stereo imaging or projection techniques, TOF sensors provide 3D data without an intermediate processing step. However, for focal-plane TOF sensors (MESA, IFM), reflections and mode jumps are common problems that potentially generate wildly erratic data.

The main stake is to specify in which sectors the sensor has to be used. As discussed before not every sensor fits to any application field. For instance, weight and size is not an issue for industrial-scale AGVs but for flying or small mobile platforms it is naturally.

## *8.3 The Sensor System*

The TACO sensor is a micro-mirror based TOF scanner whose mirror elements are optimized for fast scanning to provide QVGA resolution images at video frame rates. In one direction the mirror motion is controlled in a vectorial manner, thus permitting free specification of voxel density. The mounted elements are very robust with respect to environmental effects, including shock. Satisfactory optical aperture, sufficient for reliable imaging up to at least 7 m, is achieved by combining several synchronized mirror elements into one array. System alternatives with single mirrors will be discussed in Section 8.5.

TACO 3D point measurement precision is in the range of 2..5 mm which is comparable to high end pulse TOF survey scanners. TACO provides detection of natural surfaces at least up to 7 m (10 % reflectivity) independent of lighting conditions and features a FOV of almost 90° in the horizontal and 50-60° in the vertical direction which compares favourably with 3D/TOF camera solutions. Within limits, TACO permits to trade real time image resolution, frame rate and precision as its application demands. The downside of the pulse TOF technology within TACO is its size, volume and price, due to modularity and current commercial unavailability of appropriate integrated circuits.

Speed, precision, FOV and robustness qualify the TACO sensor for application in the robotics area and more precisely in outdoor operation, (within limits) navigation, object detection/grasping, human-robot interaction and motion capture.

The main advantage of the TACO sensor in comparison to traditional tilting laser scanners is its high frame rate and flexibility of configuration. TACOs main disadvantage is the expected system price, which probably limits application of the unmodified system currently to technological niches.

### 8.3.1 Sensor System Hardware

The design of the TACO sensor has been based on the intended broad use in robotics. The TOF technology was used based on a throughout competitive analysis, where TOF cameras set the standard for the FOV and frame rate.

A wide field of view has been considered necessary to support robot navigation or collision avoidance, which is a typical task for 3D sensors. Video frame rates in combination with reliable data give TACO a competitive advantage with respect to TOF cameras where detailed object recognition and interaction is desired, possibly supporting dynamically changing environments.

The following figures depict the TACO optical head with the single components and the TACO overall system design. The graphical representations will help the reader to understand the discussion in the following sections.

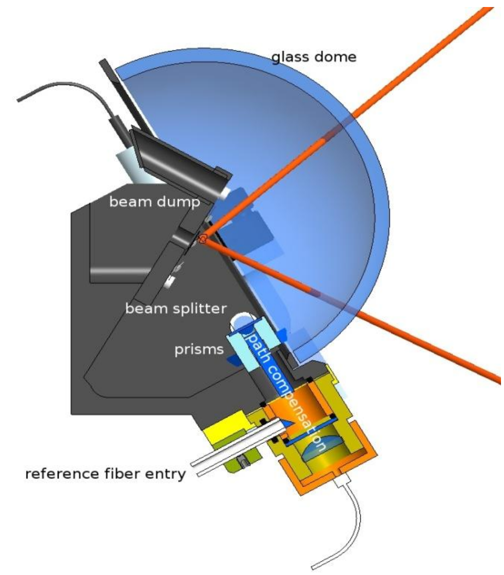



Figure 12 TACO optical head

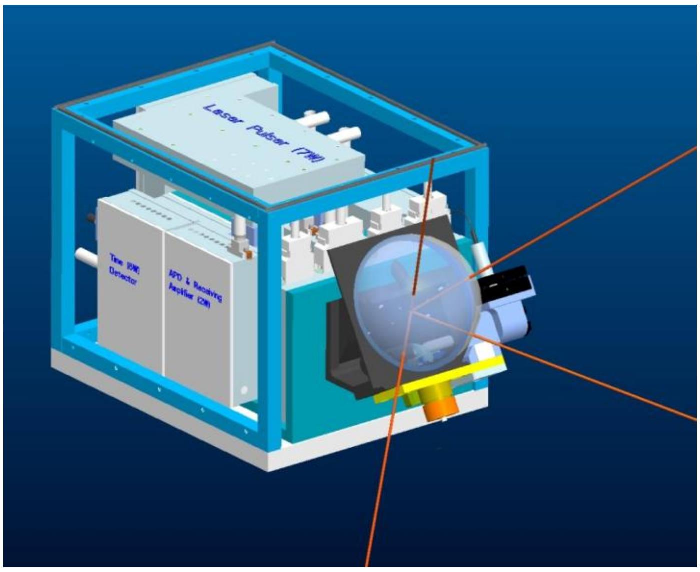


Figure 13 TACO overall system design, covers removed

The main components of the TACO sensor are considered separately to allow more detail concerning the status quo of the cost and performance goals. The main components are:

1. Time-of-flight unit: comprising embedded controller (EmCon, double 100 x 160mm PCB for CPU and FPGA base board), pulse generator (LP, top of Figure 13), laser amplifier (LA), detector and low noise analogue electronics (RA), pulse detector (TD), time-to-digital converter (TDC), see Figure 13.
2. Optical head: comprising optical elements to introduce the signal laser, beam splitter and prism assembly to guide light onto mirrors and the photodetector, respectively (see Figure 12).
3. Mirror control electronics (see behind optical unit at the front of the sensor, Figure 13)
4. System case with cooling utilities (see Figure 13).

5. Foveation system hardware (not part of the sensor prototype and not shown here)

In conjunction with the micro-mirror array aperture, the properties of the **Time of Flight** (TOF) unit translate into the beam properties of the system, its depth precision and accuracy and the measurement range. It contains very low jitter (less than 15 ps/2 mm, excluding TDC component) and high frequency electronics for time delay determination.

On the one side the optical pulse generator, the laser amplifier is used and on the other side, for signal detection, the analogue front end, a pulse discriminator board and the time-to-digital converter (TDC) is build in. The electronic interfacing on the embedded unit utilizes an FPGA interface to the timing circuitry, coupled by PCIexpress to an embedded CPU board (Q7 standard).

The **Optical Head** consists of different components like the **cover, dome, case, beam feed, beam dumps, the optical adjustment assembly, reception optics** and **the fibre coupling to photo detector**. The optical head is quite specific for the purpose at hand and it is hard to imagine use cases other than the integration with components similar to those that exist in TACO to provide an imaging scanner device. It uses some novel techniques (patent applied for) to reduce parasitic reflexes within the glass dome and to maximize the permissible mirror array synchronization error.

The foveation software currently requires an external computing device which, due to power requirements, will be placed with the sensor or the robot and uses the aforementioned I7 CPU based on recent Intel quad core technology.

The most important factor for broad market acceptance is the sensor system price. Below we list the estimated component costs (production, assembly, component test and tuning) for the sensor subsystems.

| **Part** | **Sub-Part** | **Subtotal EUR** |
|---|---|---|
| **Time-of-flight** unit and **system housing** | Optoelectronical components (Lasers, APD) EUR 10,000<br>RF electronics, power supplies, safety PCBs EUR 12,000<br>Component tuning and system calibration EUR 10,000<br>Embedded Controller EUR 3,500<br>Interconnects, mechanical parts EUR 3,000 | 38,500 |
| **Optical head** | Optical and mechanical components EUR 3,500 @ for engineering samples of ≤ 10 units.<br>MEMS array, mounting, test, tuning:<br>EUR 25,000 @ for engineering samples of ≤ 10 pcs.<br>EUR 2,000 @ 100…1000 units / a<br>≤ EUR 500 @ ≥ 5000 units / a | 28,500<br>≤ 10 units |
| **Mirror control electronics** | PCB and PCB housing | 1,500 |
| **Foveation system hardware** | External computing device with Intel quad core technology | 1,000 |
| **TOTAL** | | **69,500** |

Table 7 Summary TACO sensor component costs

### 8.3.2 TACO Foveation Software

A strong feature of the TACO camera is the integration of the sensor with a software layer between the hardware and the (robot) user which abstracts away the complex details of the implementation of the foveation mechanism.

Quoting D2.1, the primary rationale for foveation is that the current design point of the system hardware would allow approximately an optical resolution of 1600x1000 pixels for the full field of view, and that there is no sampling hardware or mirror design that allows this to be captured at full frame rate (25 Hz).

For the sampling hardware, we would require sampling hardware capable of a 40 MHz sampling rate. This is far beyond what is possible - 1-2 MHz is a more likely estimate. Eye safety constraints also put limits on the maximum possible sampling rate.

The mirror would have to be capable of providing a horizontal resonance frequency of 12.5 kHz. Such a mirror cannot be realized without significant size reductions, which would again have detrimental effects for range and precision of measurements.

Foveation acts as a "bridge" between the maximum optical resolution, the optical part of the sensor can provide, and the resolution that the rest of the system (sampling hardware and mirror) is actually capable of providing.

By detecting regions-of-interest and selectively increasing resolution temporally and spatially only in such regions, we can provide a sensor that autonomously provides simultaneously high spatial and temporally resolution. The software supports the user in addressing the challenges of the specific sensor application. As part of its foveation strategy the software features several ways to detect possibly interesting objects in the scene.

The primary reason for choosing foveation operators have been the use cases of TACO. Secondly, we have chosen to focus on operators that provide predictable performance, and that are easily understandable by the robot operator. We believe the TACO sensor will be used in a context where a robot has the main intelligence and brain in the system, and which performs overall judgement of what is important at any particular time.

For a commercial sensor, this means several design goals for the attention system:

I. The attention system must be modular: We will never be able to cover all possible use cases, simply due to that the more knowledge one has about a use case, the better performance can be obtained. By providing a modular architecture where attention operators can be plugged in, this opens the system for future innovation.

II. The foveation system must be easy to use: This relates to both the attention operators themselves – they should be very easily configurable – and the actual output of the sensor. While 3D data has become a more common source of data, most developers work with uniform resolution images, and not the variable resolution images that the TACO sensor can provide. It will be important to provide both a technical and mental bridge to ease the use of a foveating sensor.

III. The attention operators included should be applicable for many different use cases: The system should be equipped with rather generic operators that allow for a reasonable amount of foveation with little effort from the end user.

Currently, the attention system has been implemented in C++/CLI and Matlab, and further attention plugins can be developed in either of these languages. This meets design goal 1.

Regarding design goal 2, the attention operators that have been implemented in general require very few parameters to operate successfully. With regards to data output, we will continue to focus on providing data in an easy-to-use form.

Towards design goal 3, we have developed several object detection operators. In general, we would comment that our experience is that the more specific and top-down the operator is, the more useful it is in providing good foveation.

The following table shows the developed object detection operators and envisioned application fields related to them:

| Object detection operator | Example fields of application |
|---|---|
| Detecting objects based on low-level characteristics. | Agriculture<br>Object inspection<br>Manipulation/grasping<br>Mobile robotics (navigation) |
| Detecting objects of particular size and (rough) shape | Agriculture<br>Object understanding in public safety scenario<br>Natural user interfaces (for man-machine interaction)<br>Surveillance<br>Object grasping<br>Object inspection<br>Human-robot interaction<br>Object detection and classification |
| Detecting moving objects | Agriculture<br>Surveillance<br>Human-robot interaction<br>Manipulation/grasping<br>Mobile robotics (navigation) |
| Detecting moving objects even when camera itself is moving | Agriculture<br>Obstacle avoidance<br>Mobile robots<br>Visual tracking/grasping |
| Tracking of objects | Agriculture<br>Grasping moving objects<br>Tracking human hands<br>Natural user interfaces<br>Surveillance<br>Object grasping<br>Object inspection<br>Exploration (after fire, in collapsing infrastructure, rescue) |

Table 8 Object detection operators and their corresponding application fields

**Detecting objects based on low-level characteristics**

This involves detecting objects based on low-level, local characteristics like curvature, intensity and texture, and position in space, and also detecting planar surfaces and objects being attached to these planar surfaces.

Depending on the complexity of the task this can be sufficient to provide foveation, i.e. when the objects to be detected have a significantly different shape or remissivity than surrounding objects. The primary benefits of these types of operators are that they can be computationally inexpensive, which is beneficial for embedded operation.

The application fields based on low-level characteristics are certainly mobile robotics (navigation), more precisely manipulation, grasping of objects, object inspection, and agricultural applications (e.g. plants from the ground, row and groove detection). For obstacle detection the detection of the ground is required, so that (for example) objects in holes can be determined. Using low-level characteristics to foveate on the objects normal to this plane would produce good accuracy in detecting objects placed on a table.

**Detecting objects of particular size and (rough) shape**

This relates to being able to detect objects whose shape and size is known in advance. This could either build on the previous object detection (through classification of segmented objects), or be independent operators.

Regarding this class of object detection operator, there are many different application fields as can be seen in the table above. Examples for objects of particular size and rough shape are cups (on the table) or the knob of a cupboard door.

In such a case there is a need of providing high/multi resolution on objects to grasp and also of the imagery of user’s gesture and hand movements. The foveated image allows spending less computational time on the surroundings but more on the object of interest.

Concerning agricultural applications, more specific object detection operators will have to be developed. However, these could easily benefit the detection, recognition and picking of both weed and produce.

**Detecting moving objects**

This detection operators use motion as its primary cue to detect regions of interest. This can be used for multiple purposes:

- **Surveillance:** Typically, interesting objects will move around (or be moved into some position). By using motion as a cue, such objects will receive automatic attention.
- **Recognition of moving objects:** By foveating on moving objects, more details can be revealed which in turn allow moving objects to be recognized.

Especially the possibility to provide more details on moving objects, and thus allow for better recognition, can provide benefit to several application domains. Examples could include agriculture, human-robot interaction, and visual tracking for grasping.

**Tracking of objects**

By this, we mean that the sensor will be able to provide continued foveation on an object that is tracked by the sensor itself. In itself, tracking is a useful operator, as it allows for eased recognition and interaction with moving objects.

By also applying tracking to enable foveation, we can again enable better object recognition and positioning. Example application fields include grasping of moving objects, man-machine interaction and behaviour analysis.

**Other relevant object-detection operators**

To make the system more relevant to other application fields there are other operators which should be considered. The foveation operators in order to discriminate what can be accomplished in software later using the recovered 3D data. More precisely the object-detection using a CAD model would be one possibility, while it will require significantly more CPU/configuration.

For surveillance/security applications face detection TACO will provide quality data on people´s faces. Furthermore, foveation should not only focus in one dimension but in 2-dimensions, hence 'real' foveation of a x,y box. Another important factor is the identification of 'thick' straight lines, tracking these over repeated 'images' and for the 'intrusion' detection, e.g. any object within a defined zone.

**Detecting moving objects even when camera itself is moving**

In case both, the camera and the object of interest are moving the 3D vision becomes an elaborate task. The TACO sensor is set out to master such challenges and to improve both the identification of passing by objects and the secure path finding of the carrier. The primary application of this is related to obstacle avoidance/recognition during robot navigation.

## *8.4 Market Opportunities*

We believe that the primary market for the TACO sensor is within object interaction outdoors. The reason for this is that in this scenario, several of the strong benefits of the sensor are necessary: Tolerance towards background light, quality range data with mm-precision and providing range area images (not line scans).

As we currently see it, no other sensor alternatives are capable of providing this. The closest competitors are TOF cameras and laser triangulation.

TOF cameras have inherent problems in precisely measuring complex geometries. Laser triangulation can be made quite robust towards background light, but not to the level of laser ranging. For all TOF distance measurement the light paths from source over object surface to detector must be well defined, implying a very small extent of the light source.

For all TOF measurement and for laser triangulation where it is applicable to the light source, laser safety considerations will thus limit the amount of available signal light to similar levels. Best signal-to-noise ratio is achieved if all this light is concentrated on a single spot in the instant of measurement. Scanned TOF must thus be expected to deliver the most robust performance with respect to background light.

To maximize the benefit of micro-mirrors commercially, we believe that it is necessary to determine a mid-scale market, such that volumes can be sufficient. One such market could be the automation of agriculture, which is an application where automation is being rolled out more expansively, and which could have a significant volume.

For more high-end markets, examples could include oil industry (where the sensor could be applied as part of a solution for remote maintenance and operation of offshore oil fields) or nuclear industry (where remote operation again is necessary). To the extent that performance requirements are met, these could be interested in a (ruggedized) version of the current sensor system.

## *8.5 Technology improvement opportunities*

Now we will look in some details of possible TACO technology improvements. First we will introduce briefly the TACO sub-components to facilitate the discussions about possible technological changes and their implications on specification and sensor cost, the possible applications (cf. see Table 8) and the ensuing marketing opportunities.

### 8.5.1 Giving up the quasi-static micro-mirror axis

Given the small size of the micro-mirror aperture in TACO a pulse TOF distance measurement to provide measurement distance of at least 7 m for dark grey objects is needed. Phase TOF in comparison causes smaller measurement distance, is more precise at small distances, but less precise farther away, and requires a very clean separation of emitted and received light.

However, it can be implemented using significantly less expensive components. Rethinking TACO as a phase TOF scanner implies either (i) increasing aperture significantly in order to maintain measurement distance and precision or (ii) accepting a performance penalty using comparable optics.

Addressing (i), the array of 2D-mirrors may be replaced by an array of resonant 1D-mirrors, mounting the optical unit on the axis of an external motor which provides a steerable second axis. An aperture of 2 cm, increasing available signal light by a factor of about 16 with respect to TACO, is achievable.

Foveation will still be possible by changing mirror amplitude and controlling the motor path, but imaging conditions cannot be changed as fast as in the current TACO solution. A motor driven system may not provide a single frame as fast as TACO can, but it may implement up to 360° FOV, which has been proving valuable in object detection.

Using phase instead of pulse TOF (option ii) has been studied in D2.1, with the result that phase TOF precision will drop below that of pulse TOF at distances larger than 1 m from the sensor. A dedicated emission mirror will be necessary for light path separation, requiring redesign of the mirror array and the optical unit. However, from the perspective of the desired object interaction capabilities, the performance hit at larger distances may be acceptable.

We consider the cost of a phase-TOF system to be about 1/3 of a pulse system for similar point rate and detector. In particular, a laser amplifier will not be required.

### 8.5.2 One mirror replacing the mirror array

From the perspective of the optical unit, using a single larger ‘micro’ mirror with similar FOV and operating frequencies as the current array TACO mirrors will simplify the optical unit and eliminate the parallelized mirror control electronics.

A suitable mirror with about 5 mm aperture has been designed at IPMS and a prototype is being built. The single mirror must be shared for emission and reception though, because the design is characterized by a very high Q-factor and does not permit an economic possibility to combine two or more mirrors in an array.

The resonant axis frequency is lower than TACOs by a factor of 3. Matching horizontal and vertical point densities will thus entail slower frame rates, while the reduced system complexity will certainly be advantageous for its robustness.

The TOF options have not been changed, so we refer to the proceeding section.

### 8.5.3 Pulse TOF options

Apart from the option to employ phase-TOF to reduce cost for the TOF unit, there exist some options for pulse-TOF implementations that have not yet all been discussed in D2.1.

Using a semiconductor laser source at red or near infrared wavelengths will lead to reduced beam quality and a larger spot size because diodes suitable for high peak power are multi-mode edge emitters which cannot be collimated to small diameter circular beams without losses.

Reasonably short pulses (ns-duration) from single emitter diodes are possible with several 10 W, probably up to 50 W peak power. Laser safety will impose either slower repetition rates or less peak power than at 1550 nm, where permissible average laser power is at least one order of magnitude higher. Visible and near-IR options have been studied in D2.1 and were abandoned in favor of 1550 nm operation wavelength for reasons of better signal-to-noise (measurement range) and beam quality.

One could harvest on cost advantages from eliminating the TACOs fiber laser amplifier and using silicon in lieu of InGaAs detector devices. This would possible reduce 10 – 15% the estimated system cost.

As well interesting, but with some conceptual question marks, would be the use of picosecond (20-60 ps) laser sources in combination with a fast Geiger-mode avalanche photodiode (silicon solid state photomultipliers). The timing precision from the rising flank of the pulse will likely match current specifications but no pulse processing circuitry will be necessary on the receiver side.

Depending on dark count rate of the diode, however, multiple measurements per point will be required and must be accommodated by appropriate time digitizers. To achieve TACO point measurement rates, a small degree of parallelization seems to be necessary to compensate repeated measurements. Possible savings are again about 10 - 15% of estimated recurring system cost.

### 8.5.4 Cost reduction of system electronics

There is no single price driving component within the system electronics. Optical unit and MEMS electronics, the optoelectronic, in particular the power laser and the TOF electronics components each contribute significantly to the system cost budget.

In a further development step towards system production, some inefficiencies due to modularization (interconnects, PCB size and production) and component cost can be eliminated or reduced, but a high system cost baseline will remain, probably at a level of 70 to 80% of the total listed in the cost table. We also estimate that by such an improvement step, cannot reduce the system costs by more than 50%. This would include the integration of current components on a single main board and an ASIC development for critical TOF components but the Baseline costs for Emitter, APD, TDC, or EmCon will stay.

Even with mass-market sales numbers in mind, which may permit targeted ASIC developments, the system will remain complex. Comparing with Sick's line laser scanners, the TACO price will likely not drop below 2 or 3 times the price of a line scanner.

### 8.5.5 Software Outlook

We believe that the primary use case for foveation is object interaction and recognition. In these cases, only a small part of the field-of-view requires higher data quality than the remaining image.

We see two primary ways forward for the foveation software: Extension by including new object detection algorithms, and application of the software towards other hardware combinations.

**New object detection operators**

To provide a high degree of foveation, it is necessary to have precise object detection algorithms. This means that for targeted markets, it makes sense to develop specific algorithms for object detection within that market.

This could be i.e. plant/produce detection (for agriculture), person detection (for surveillance), or hand/face detection (for man-machine interaction).

**Use towards other hardware concepts**

As we see it, foveation has merits beyond the current hardware implementation. Both the current foveation architecture and its underlying object detection operators could provide benefit towards other (3D) sensor combinations to provide foveation and increased data quality in regions-of-interest.

Examples of other hardware combinations that can enable foveated data acquisition based on 3D:

- TOF area camera (overview image), scaled down version of the TACO laser scanner sensor (foveated image). This has the benefit of reducing demands on the micro-mirror, as it is not longer required for acquiring the overview image.
- Kinect-similar sensor (overview image), tiltable laser triangulation setup (foveated image). This can provide a rather inexpensive setup for foveated imaging, albeit primarily towards indoors application.

Looking beyond 3D, one could envision looking into using normal cameras, possibly in a stereo configuration, and applying a more classic attention-based approach.

The software system is designed such that the foveation architecture largely can be reused independent of the underlying sensor modalities. The same is the case of many of the foveation operators, which can be adapted to work on other hardware sensors.

## *8.6 Implementation beyond the project lifetime*

The technology roadmap of Three-dimensional Adaptive Camera with Object Detection and Foveation (TACO) has been developed to shade light on a set of needs and technologies required to satisfy specific needs. The TACO 3D sensing system itself lends an opportunity not only for robot developers and academic institutions, but also for an instrument supplier willing to develop an industrial version of a new 3D sensing system.

It might become a highly competitive product not only in the robotics market, but also for a wide range of industries. Examples include logistics (package size measurement), surveillance (person counting and tracking), medical (measurements for manufacturing orthopedic inserts) and part-producing industry (part inspection and quality assurance).

Given the importance of the results derived from the TACO project, possible areas of implementation beyond the project lifetime are manifold. The individual components can be industrialized and commercialized separately:

4. Industrialization and commercialization of the TACO sensor, either through collaboration with established industry or by a spin-off from the project itself
5. Bidirectional MEMS mirror (Fraunhofer IPMS)
6. Electronics components for time-of-flight units (Fraunhofer IPM and CTR)
7. Software licensing towards industry (beyond the open source software on foveation which is available for further research and development) SINTEF will handle these issues through **Sinvent AS**, their Technology Transfer Office[3] in Norway.

Mirrors, drivers and camera intended for robot vision will be used for very different applications (e.g. surveillance cameras). Also experience from the development of the foveation software toolbox will be useful in completely unrelated application fields.

Conversely, there already exist some ideas for commercializing the entire technology. A potential implementation of the TACO sensor would be to mount it on an autonomous maintenance robot for use in large-scale physics experimental systems, like the CERN Large Hadron Collider.

There the TACO sensor would help to localize circuit boards and how to un-mount them from the racks they are screwed onto. Another implementation would be to use the TACO sensor on the ESA rover, for it to perform autonomous tasks in space. It would grant the rover the ability to localize precisely different objects in real-time, hence to manipulate them achieving different envisaged goals.

Lastly, a possible implementation for this sensor would be to provide data for a visual data glove for teleoperation control.

We had bilateral exchange of ideas and thoughts with projects where TACO technology could deliver valuable input:

- The Jet Fusion Project
- The ITER fusion project[4]
- Other fusion projects such as MYRRHA[5]
- Dounreay site decommissioning project[6]
- Other decommissioning projects

The TACO sensor offers a unique platform for further research on perception strategies, control strategies, robot-environment interaction modes, cooperative robotics, human-computer interaction which will be used in both academic and applied research to reach a higher level of understanding of how to develop new solutions to many of the problems currently encountered within robotics area.

[3] http://www.sintef.no/Home/Technology-transfer/
[4] www.iter.org
[5] www.sckcen.be/en/Our-Research/Research-projects/Internal-projects/MYRRHA
[6] www.dounreay.com

# 9 List of Abbreviations

| | |
|---|---|
| AGV | Advanced video graphics array |
| APD | Photo detector |
| ASIC | Application-specific integrated circuit |
| CPU | Central Processing Unit |
| ECL | Emitter coupled logic |
| ETX | End of transmission |
| EUROP | European Robotics Platform |
| FOV | Field of View |
| FPGA | Field programmable gate array |
| GPU | Global Processing Unit |
| HANDLE | Developmental pathway towards autonomy and dexterity in robot in-hand manipulation |
| HF | High frequencies |
| I/O | Input / Output |
| InGaA | Indium Gallium Arsenide |
| IR | Infrared |
| MEMS | Micro-Electro-Mechanical Systems |
| MMI | Man-machine interface |
| MOEMS | Micro-opto-electromechanical systems |
| PCB | Printed Circuit Board |
| PMD | Photonic mixing device |
| SDK | Software Development Kit |
| Si-Ge | Silicon-Germanium |
| SME | Small and Medium sized Enterprise |
| SoC | System on a chip |
| Stdev | Standard deviation |
| TACO | Three-dimensional Adaptive Camera with Object Detection and Foveation |
| TDC | Time-to-digital conversion |
| TOF | Time-of-flight |
| VGA | Video graphics array |
| VR | Virtual Reality |
| ROS | Robot Operating System |

# 10 References

[1] P. Einramhof and M. Vincze, Stereo-based Real-time Scene Segmentation for a Home Robot, 52nd International Symposium ELMAR-2010, Zadar, Croatia, September 15-17, 2010.

[2] Sebastian Thrun, Dieter Fox, Wolfram Burgard, and Frank Dellaert. Robust monte carlo localization for mobile robots. Artificial Intelligence, 128(1-2):99–141, 2000.

[3] Kai O. Arras, Jose A. Castellanos, Martin Schilt, and Roland Siegwart. Feature-based multi-hypothesis localization and tracking using geometric constraints. Robotics and Autonomous Systems, 1(44):41–53, 2003.

[4] P. Liljebäck, T. Kavli, and H. Schumann-Olsen, "Robotic technologies for an unmanned platform", SINTEF, Report STF90 F05405, 2005.

[5] B. Graf and K. Pfeiffer, "Mobile robotics for offshore automation", in Proceedings of the IARP/EURON Workshop on Robotics for Risky Interventions and Environmental Surveillance, Benicssim, Spain, 7-8 January 2008.

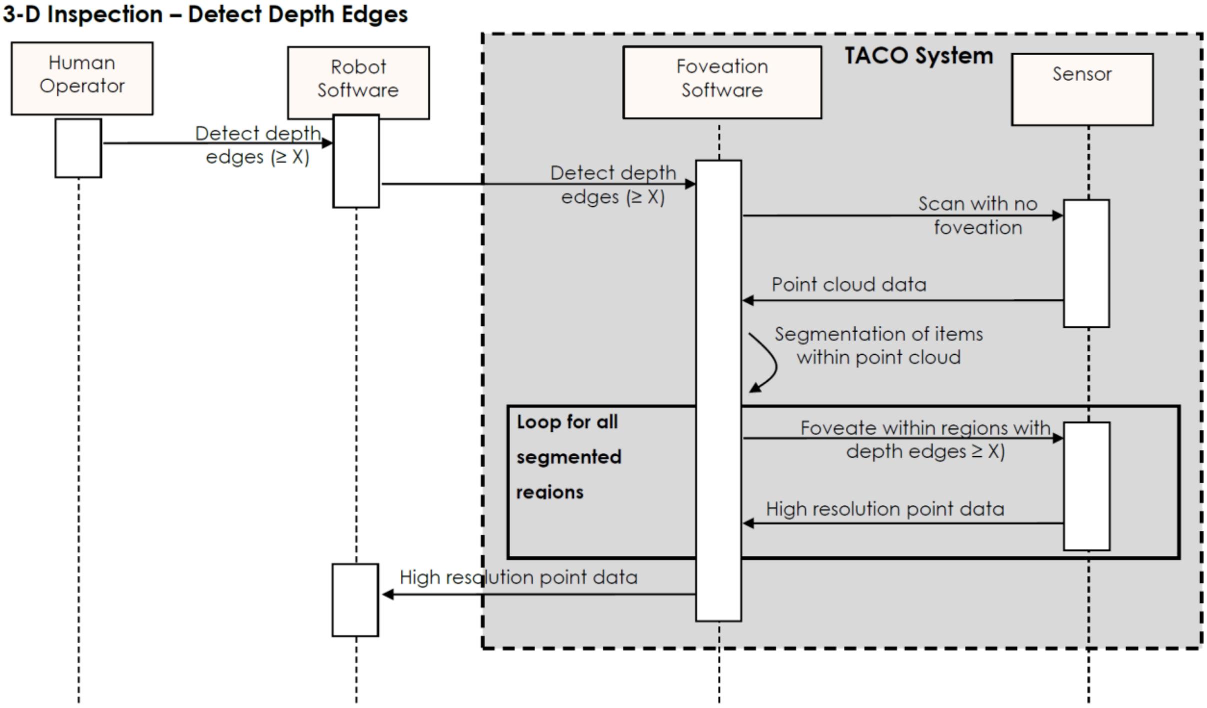
[6] B. Graf, K. Pfeiffer, and H. Staab, "Mobile robots for offshore inspection and manipulation," in Proceedings of the International Petroleum Technology Conference. Dubai, U.A.E.: SPE, December 4 - 6 2007.

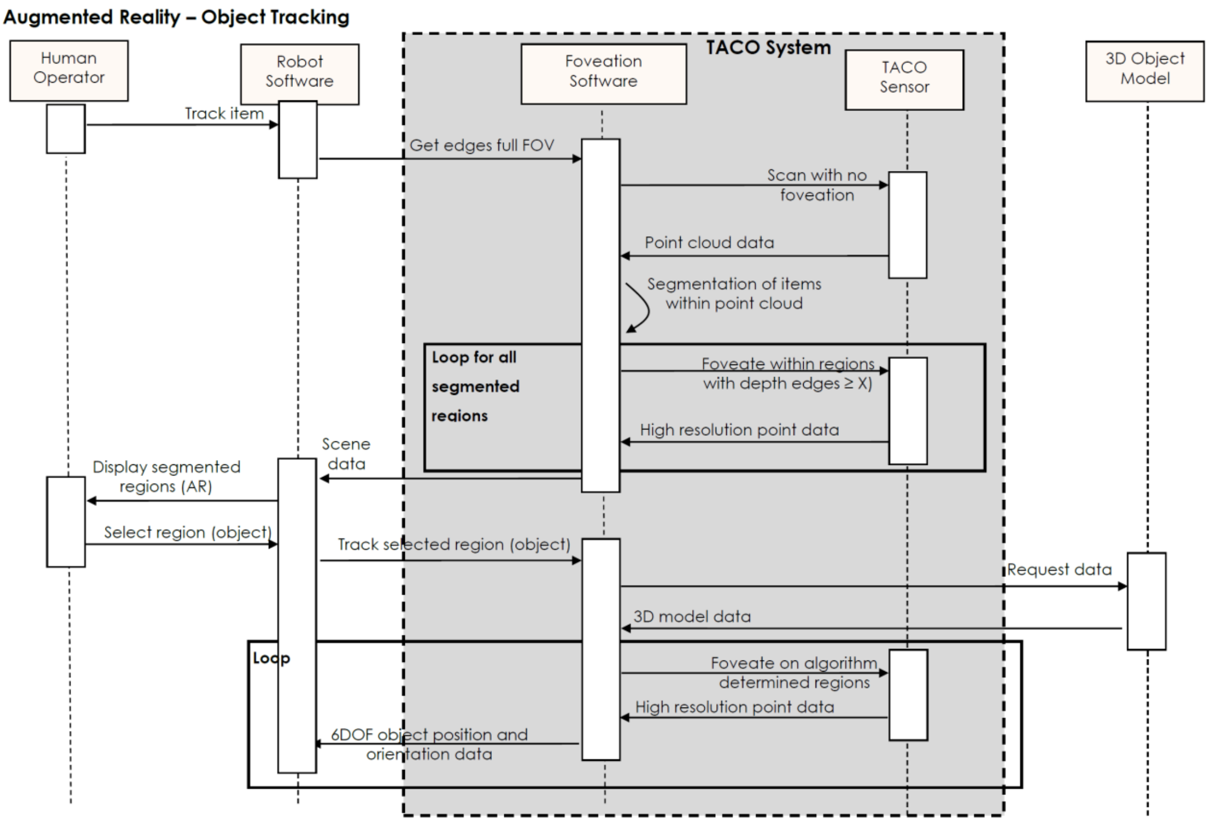
[7] E. Kyrkjebø, P. Liljebäck, and A. A. Transeth, "A robotic concept for remote inspection and maintenance on oil platforms," in Proc. ASME 28th Int. Conf. on Ocean, Offshore and Arctic Engineering (OMAE 2009), Hawaii, USA, May 31 - June 5 2009.

[8] A. A. Transeth, Ø. Skotheim, H. Schumann-Olsen, G. Johansen, J. Thielemann and E. Kyrkjebø. "A Robotic Concept for Remote Maintenance Operations: A Robust 3D Object Detection and Pose Estimation Method and a Novel Robot Tool", International Conference on Intelligent Robots and Systems 2010, Taipei, Taiwan. Paper submitted.

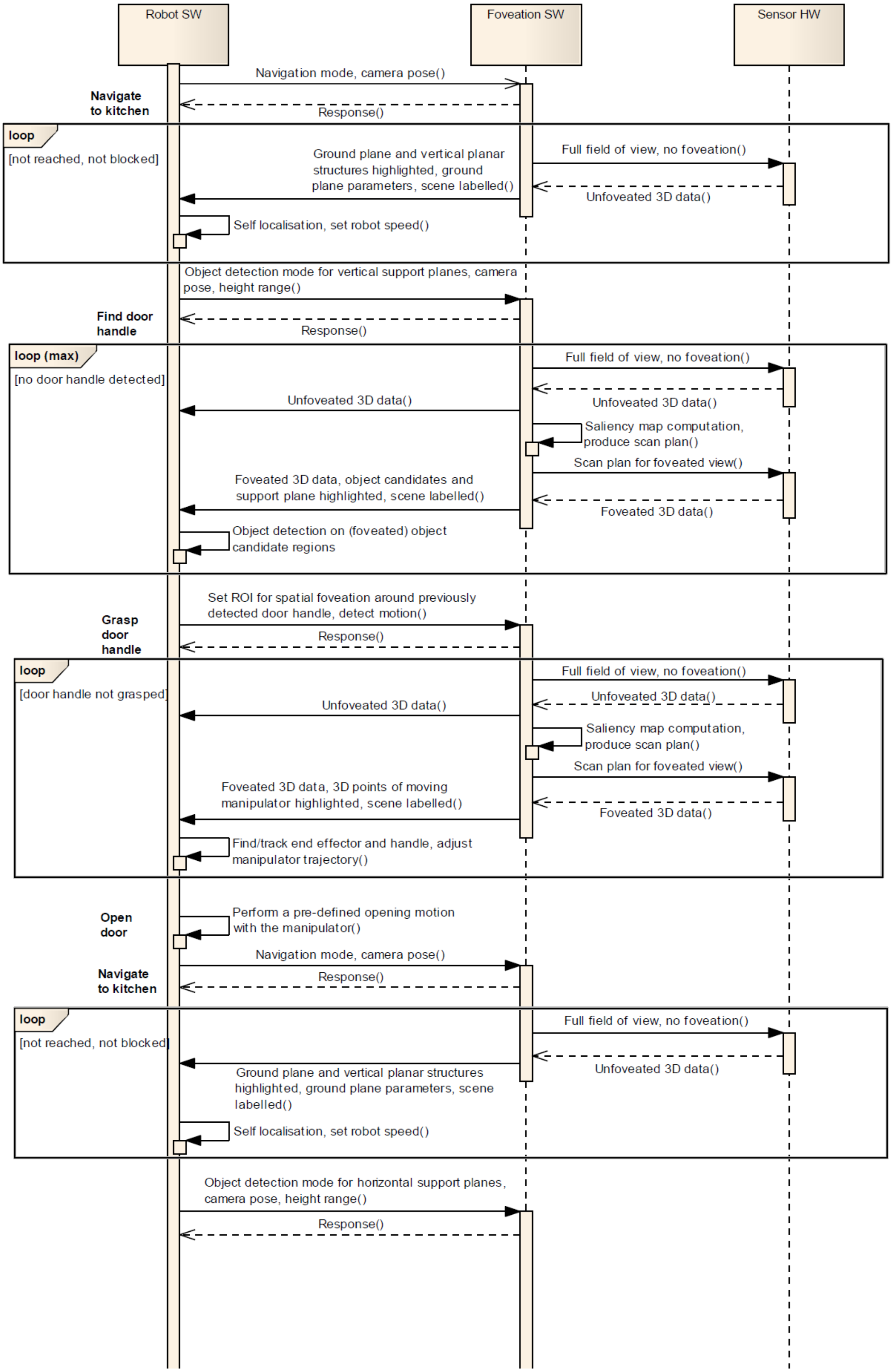
[9] An autonomous towel folding robot;
http://www.willowgarage.com/blog/2010/04/02/towels-uc-berkeley

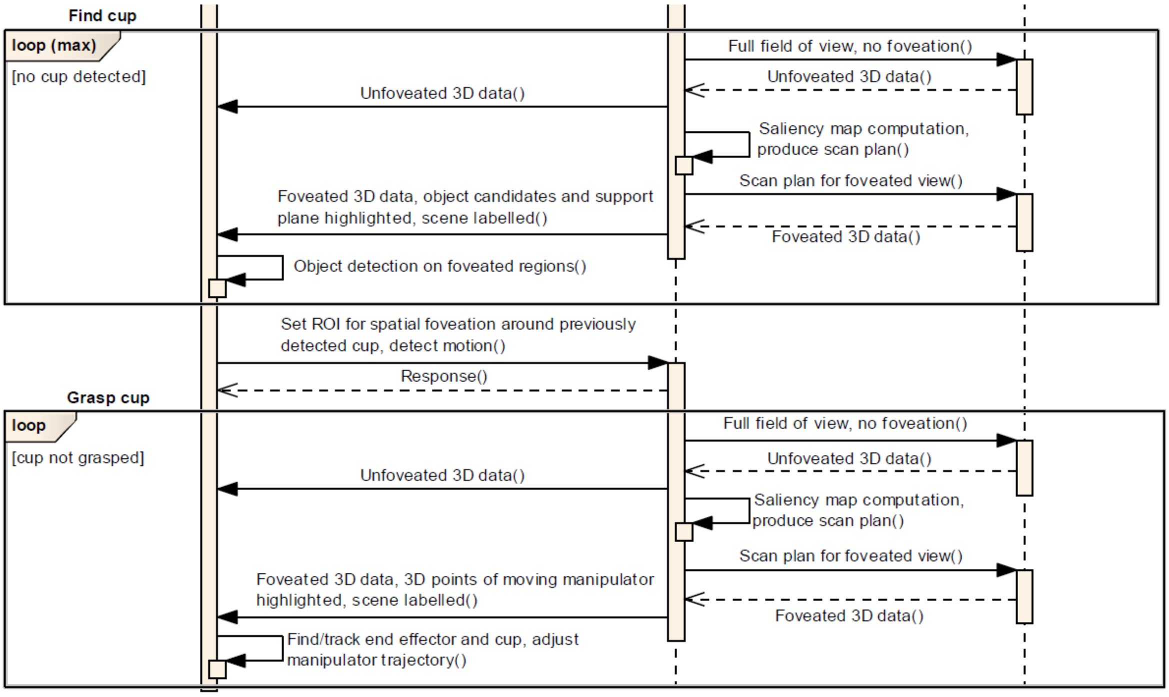
# A.1 UML Diagrams



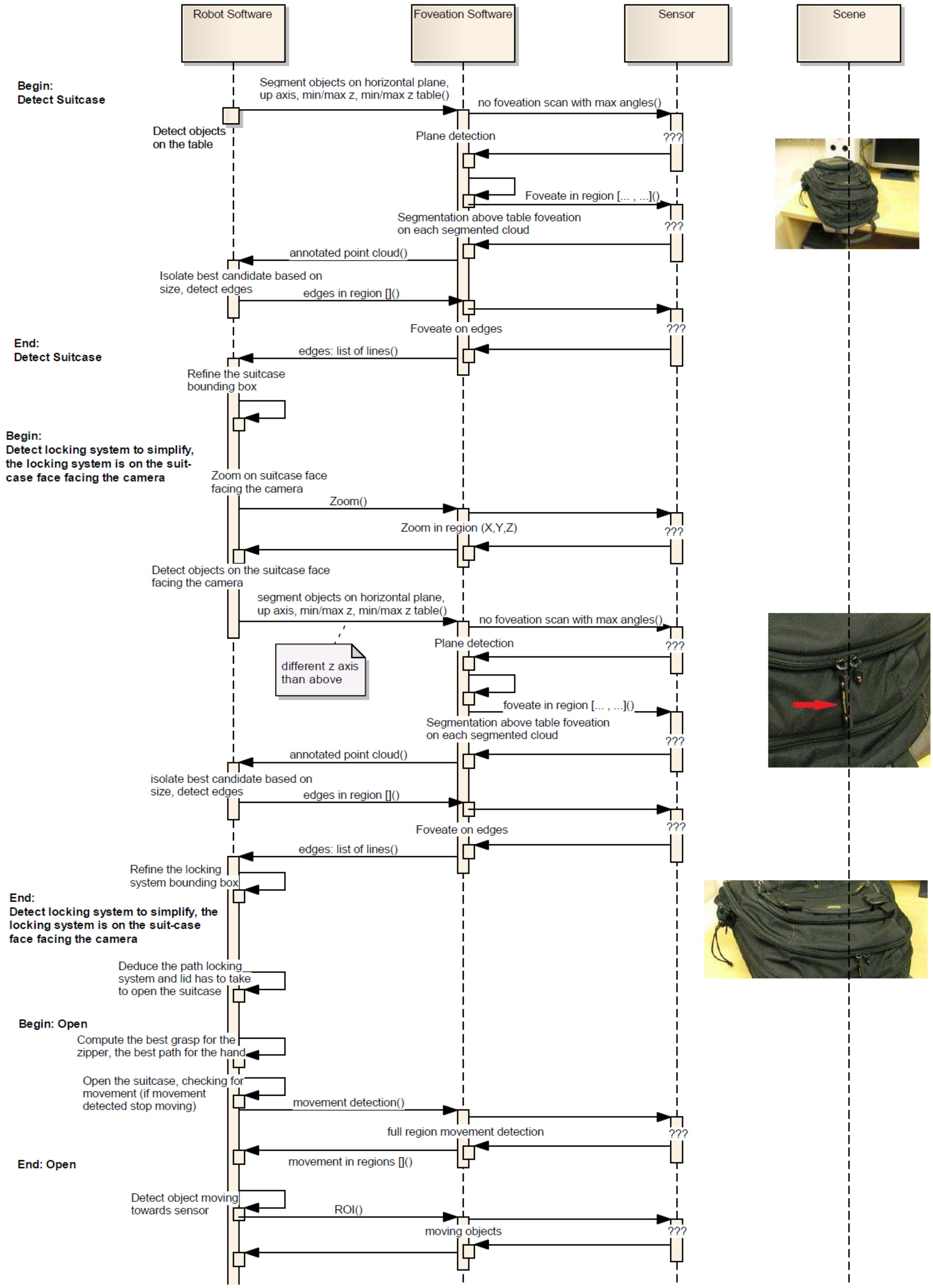



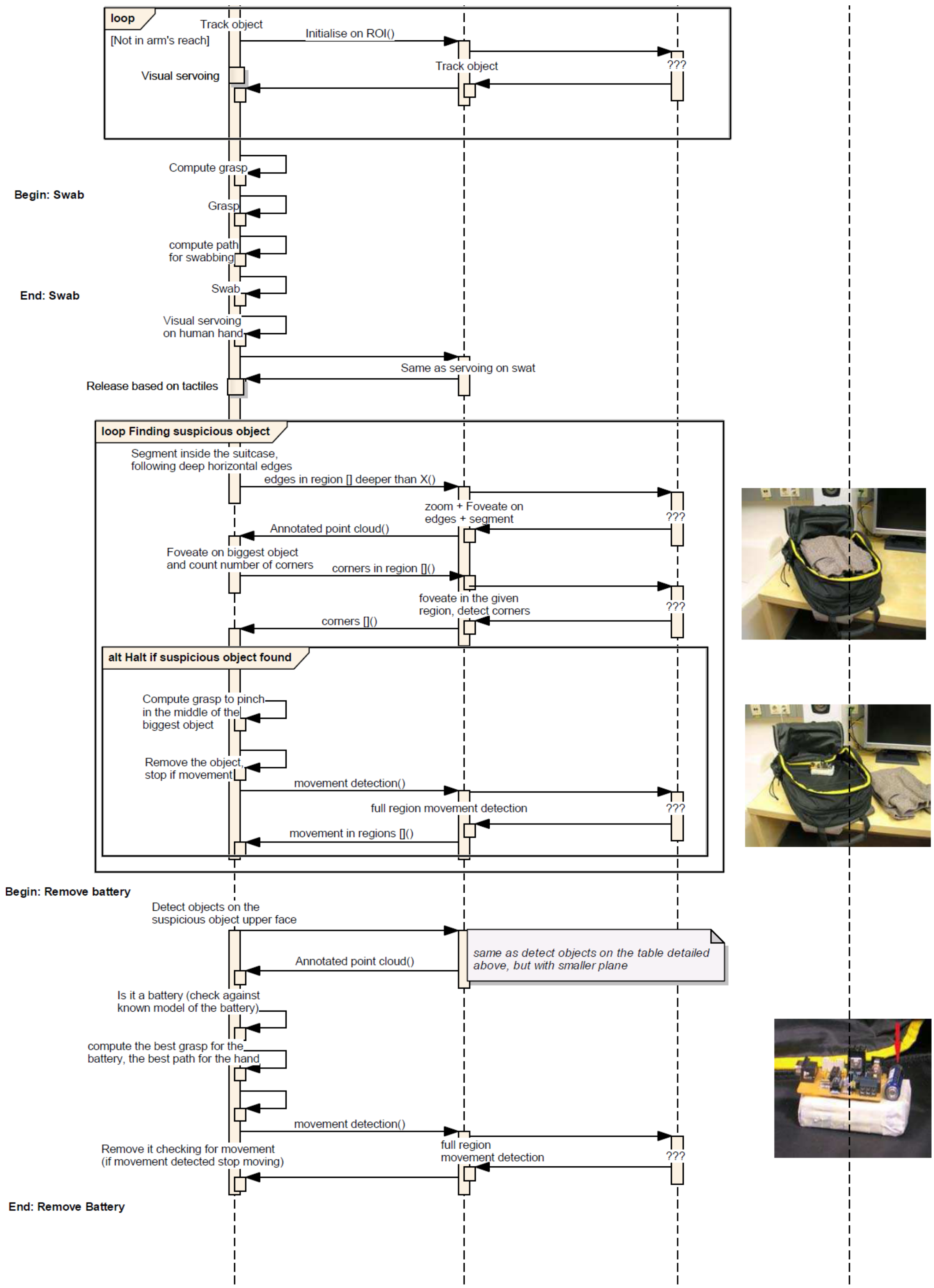
## Home Grasping




Find cup
loop (max)
[no cup detected]
Full field of view, no foveation()
Unfoveated 3D data()
Unfoveated 3D data()
Saliency map computation, produce scan plan()
Scan plan for foveated view()
Foveated 3D data, object candidates and support plane highlighted, scene labelled()
Foveated 3D data()
Object detection on foveated regions()
Set ROI for spatial foveation around previously detected cup, detect motion()
Response()
Grasp cup
loop
[cup not grasped]
Full field of view, no foveation()
Unfoveated 3D data()
Unfoveated 3D data()
Saliency map computation, produce scan plan()
Scan plan for foveated view()
Foveated 3D data, 3D points of moving manipulator highlighted, scene labelled()
Foveated 3D data()
Find/track end effector and cup, adjust manipulator trajectory()

## Public Safety





loop
[Not in arm's reach]
Track object
Initialise on ROI()
Track object
???
Visual servoing
Compute grasp
Begin: Swab
Grasp
compute path
for swabbing
Swab
End: Swab
Visual servoing
on human hand
Same as servoing on swat
Release based on tactiles
loop Finding suspicious object
Segment inside the suitcase,
following deep horizontal edges
edges in region [] deeper than X()
zoom + Foveate on
edges + segment
???
Annotated point cloud()
Foveate on biggest object
and count number of corners
corners in region []()
foveate in the given
region, detect corners
???
corners []()
alt Halt if suspicious object found
Compute grasp to pinch
in the middle of the
biggest object
Remove the object,
stop if movement
movement detection()
full region movement detection
???
movement in regions []()
Begin: Remove battery
Detect objects on the
suspicious object upper face
Annotated point cloud()
same as detect objects on the table detailed
above, but with smaller plane
Is it a battery (check against
known model of the battery)
compute the best grasp for the
battery, the best path for the hand
movement detection()
Remove it checking for movement
(if movement detected stop moving)
full region
movement detection
???
End: Remove Battery