Underwater Imaging and Optics: Recent Advances

Frank M. Caimi¹, Donna M. Kocak², Fraser Dalgleish³, John Watson⁴

¹P.O. Box 700367
Wabasso, FL 32970

²Maritime Communication Services / HARRIS Corporation
1025 West NASA Blvd.
Melbourne, FL 32919 USA

³Harbor Branch Oceanographic Institution at FAU
5600 N US 1
Fort Pierce, FL 34946 USA

⁴University of Aberdeen
King’s College
Aberdeen AB24 3FX UK

Abstract – Obtaining satisfactory visibility of undersea objects has been historically difficult due to the absorptive and scattering properties of seawater. Mitigating these effects has been a long term research focus, but recent advancements in hardware, software, and algorithmic methods have led to noticeable improvement in system operational range. This paper is intended to provide a summary of recently reported research in the area of Underwater Optics and Vision and briefly covers advances in the following areas: 1) Image formation and image processing methods; 2) Extended range imaging techniques; 3) Imaging using spatial coherency (e.g. holography); and 4) Multiple-dimensional image acquisition and image processing.

I. INTRODUCTION

Recent advancements in the field of Ocean Optics are, at least, partially attributable to the following developments:

- Affordable, high quality cameras that support a suite of fast, inexpensive specialized image processing software and hardware add-ons;
- Digital holographic cameras that record interference fringes directly onto solid state sensors (i.e., mega pixel charge coupled devices) to produce time resolved, 3-D movies;
- High repetition rate, moderate power lasers and advanced detector designs which enhance performance of two-dimensional (2-D) and three-dimensional (3-D) imaging systems;
- Compact, efficient and easy to program digital signal processors that can execute algorithms once too computationally expensive for real-time applications;
- Modeling and simulation programs that more accurately predict the effects that physical ocean parameters have on the performance of imaging systems under different geometric configurations;
- Image processing algorithms that handle data from multiple synchronous sources and that can extract and match feature points from each such source derive accurate 3-D scene information; and
- Digital compression schemes provide high-quality standardizations for increased data transfer rates (i.e. streaming video) and reduced storage requirements.

The applicability of each of these is unique to general topical areas in undersea imaging covered in the following sections. A more thorough discussion of these topics is found in [1].

II. CONVENTIONAL IMAGE ACQUISITION FOR SURVEY AND MONITORING

Many recent applications are aimed at deploying conventional cameras for long-term monitoring and observation where collection of video imagery and subsequent processing can be used to provide image scaling and measurements, identification and assessment, 3-D reconstructions, and analysis [2-5]. These systems benefit from cameras capable of operating continuously for long, unattended durations with ample resolution and high bandwidth data links. One recent example is a high-speed megapixel benthic imaging system designed to operate from a towed platform for scallop stock assessments [6]. The system features a commercial GigE Vision™ camera (1360 x 1024 pixels), strobe light, and tow vessel that streams images at 4/s for near real-time monitoring. Tow speeds of 5 – 8.5 km/h provide 58 – 27% overlap between successive images with a resolution of greater than 1 pixel per mm, enabling identification of objects less than 50 mm in diameter in most water conditions. The system is also useful for fine-scale habitat mapping, ground-truthing acoustic data, benthic ecology research and fishing damage assessment.

Higher accuracy imaging equates to increased pixel resolution requiring much greater bandwidth and storage capability. High definition (HD) and ultra high definition video (UHD or super hi-vision) produce images having 1920 x 1080 and 7,680 x 4,320 pixels, respectively. Image compression formats such as HDV, MPEG-4 AVC/H.264 and VC-1 reduce data rate and memory demands by up to a factor of nearly 50 [7, 8]. Although these reductions help make requirements more manageable, users are often not in favor of “losing” data. Table 1 provides approximate upper bound bandwidth requirements for various types of cameras, including some with image compression.
III. EXTENDED RANGE IMAGING

A primary goal of extended range underwater imaging is to improve image contrast and resolution at greater distances than what is possible with a conventional camera and underwater lighting. Schemes implemented include:

- Time discrimination / range-gated methods;
- Spatial discrimination / laser line scan (LLS) methods;
- Scattered light rejection using modulation / demodulation techniques;
- Imaging using structured lighting;
- Polarization discrimination; and
- Multiple perspective image construction.

Current emphasis has been to develop wide swath imagers with the potential to be implemented on the common classes of autonomous underwater vehicles (AUVs). Previously demonstrated imaging system architectures include: i) synchronously scanned continuous wave (CW) beam with photomultiplier tube (PMT) detection (Raytheon LS4096 and the Northrop Grumman LLS systems); ii) scanned laser beam and high speed linear CCD camera (LBath system [11]); iii) pulsed fan beam with time-resolved signal detection using streak tube (Arete STIL system); and iv) CW fan beam structured lighting system with a CCD camera (University of South Florida and SRI International Bluefin2).

All approaches provide intensity reflectance maps, while the latter three also provide 3-D topographic information, either via time-of-flight or triangulation. Each system has limitations or operational challenges that result from the design choices, i.e. limited depth of field, leading to difficulty in ‘focusing’ the system variable terrain; large size; high input power requirement; and limited operational range versatility. The operational community however requires reduced operational complexity (size, power, etc.) and enhanced performance.

A. Time Discrimination / Range-Gated Methods

Range-gated systems have traditionally used low repetition rate (< ~100 Hz) green pulsed laser sources to improve image contrast by rejecting backscattered light between the source and target. Previous and recent implemented bench top configurations [12-14] use a spatially broadened laser pulse as the illuminator and a non-scanning gated intensified camera as the detector allowing for the acquisition of a thin temporal slice of the entire (global) scene, over up to 40 degree field of view (FOV). Utilizing suitably high sampling rates, these systems can also allow for 3-D image reconstruction from many short time slices [15].

More recently, an evolved implementation of this variety of system, the LUCIE2 (Laser Underwater Camera Imaging Enhancer) [16], has been packaged and deployed onboard an ROV for propeller blade inspection at 5 attenuation lengths. Optical polarization differencing is used to further enhance target contrast. A compact third generation diver-held version is also under development. Other techniques to extend the performance range and retrieve optical properties of the environment are described in [17].
B. Spatial Discrimination / Laser Line Scan (LLS) Methods

Laser Line Scan (LLS) systems optically scan a narrow instantaneous FOV (IFOV) receiver synchronously with a highly collimated laser source linearly over a large angle. It has been shown that the optical transfer function of such systems can be near the diffraction limit [18].

Although effective at spatially rejecting scattered light, LLS systems can be limited by receiver shot noise resulting from the temporal overlap between the target return and that of the scattering volume, from both laser and solar illumination. To maximize operational range, CW LLS systems use increased source-receiver separation that reduces the detrimental effect of near field multiple scattered light reducing their potential for use aboard small unmanned underwater vehicles (UUVs). Computer models of such systems have been studied [19, 20] indicating that images can be obtained at up to 6 attenuation lengths as has been verified in both field and lab-based demonstrators. Employing high repetition rate pulsed lasers with receiver gating (the Pulsed Laser Line Scan or PLLS imager) reduces the effects of beam overlap and has recently been shown to be effective at over 7 attenuation lengths. This method also offers a reduced form factor for payload operation on small UUVs.

Computer model developers have been concentrating on synoptic model development for LLS and PLLS systems to avert the computational burden of Monte-Carlo ray trace methods and to allow system designers to consider the trade-off of system parameters [21]. Recent studies indicate that the PLLS approach allows temporal separation of the volume scatter and target signals improving the image contrast and the operational range of the system [22, 23]. Electronic gating is particularly advantageous at a smaller source receiver separation [24].

Examples of the improvement possible using the PLLS system has been demonstrated at HBOI/FAU using a custom high repetition rate (357 kHz), high power green pulsed laser (6-7 ns full width at half maximum). Test tank reflectance image comparisons are shown in Fig. 1 indicating improved signal-to-noise ratio (SNR) and contrast of the time-gated PLLS over CW LLS in under near-identical system and operational conditions.

At the limit of detection of the time-gated PLLS, several possibilities exist to extend the performance. The integration of multiple pulses will increase the SNR, albeit with a sacrifice to the achievable image resolution. Physically increasing the size of the receiver aperture can also increase the SNR; likewise the use of coded pulses and coherent detection has also been shown to potentially further extend imaging capabilities [25, 26].

Figure 1. Test tank acquired reflectance image portions taken at 7m stand-off distance.
Left side images of each column: CW LLS image using 1.5W CW laser; Right side: Time-gated PLLS image using 1.5W average power pulsed laser at 357 KHz. For both sets of tests, the scan speed was 100 lines per second, source receiver separation was 23.4cm, instantaneous FOV of the receiver was 15mrad and seabed velocity was 1.5ms⁻¹. Laser divergence for both lasers was 2-3 mrad. Note that all images have been histogram equalized and median filtered (3x3).
C. Scattered Light Rejection using Modulation / Demodulation Techniques

It is well known that coherent modulation/demodulation methods can offer improvement in signal detection for communications systems, but at optical frequencies these schemes are not useful over large distances due to the dispersive nature of seawater [27]. Consequently, coding of modulated waveforms is a more practical scheme to achieve improved S/N at the detector. As an example, earlier underwater coherent detection demonstrations have used a CW laser radiating an amplitude modulated beam of light illuminating an underwater target from an airborne platform or underwater vehicle. A PMT then integrates the backscatter and the target photons together, and by demodulating the AM signal, partially rejects the scattered light signal enabling ranging ultimately limited by receiver shot noise. It has also been recognized by the authors and others, that the non-coherent signal detection methods used by earlier LLS systems might also be improved by using sub-carrier coherent detection at the receiver to separate the temporally dispersive scattered light from the target return and to produce target profile or range information. One such system has been developed by NAVAIR to image targets underwater from an airborne platform or underwater vehicle [28, 29]. This system uses a 3W CW laser sinusoid modulated at up to 100 MHz, with complex (IQ) demodulation to recover magnitude and phase information for enhanced contrast and range imaging capabilities. The system has also been demonstrated as having the potential for hybrid imaging/communications capabilities [30]. Other recent modulated CW imaging demonstrations [31] utilized a 20 mW single mode laser at 405 nm amplitude modulated at 36.7 MHz via control of the current, and scanned in steps by a miniaturized piezoactuator. In laboratory tests at ENEA Research Center, submillimeter range accuracy was reported at a 1.5 meter stand-off distance in clear water. The diode laser wavelength matches the minimum of the pure water absorption spectrum, and hence this system has been designed for (eventual) long range 3-D imaging in relatively clear water.

The goal for any operational imaging system is a full-up demonstration, and progress is being made currently. In 2007, the NAVAIR system was tested with the HBOI bench top LLS system [32]. The results, which demonstrated a noticeable reduction in backscatter and hence improvement in image contrast when compared to CW LLS in turbid water (shown in Fig. 2), were reported in a recent poster [33].

It has been proven in simulation that the use of modulated-pulses, described by [25], as the hybrid LIDAR-radar technique has the potential to further extend the operational range of LLS systems. The simplest method is to impress a high frequency sinusoidal amplitude modulation on the laser pulse. This in turn makes it possible to reject the lower frequency components of backscatter and ambient light, further increasing the range capability of the system. This type of system has previously been investigated by various research laboratories [25, 34] and has been the subject of recent simulations using Metron’s radiative transfer solver [21] and other radiative transfer codes developed specifically for pulsed underwater laser imaging systems [35]. Sub-systems hardware development (lasers and detection means) and verification of simulation codes continue to remain a primary activity for these systems.

![Figure 2. Raw image comparison from HBOI test tank at 7 meters stand-off distance (contrast stretched between min to max) between CW LLS (left side images on each column) versus modulated-CW LLS (right side images on each column). Note: C = beam attenuation coefficient in inverse meters. CL = number of attenuation lengths.](image)
D. Imaging using Structured Lighting

When using structured light, a narrow laser beam or plane is typically projected onto the scene off the center axis of a standard 1-D or 2-D camera to reduce backscatter and enable recovery of the scene's 3-D structure by means of triangulation. One method commonly used with structured lighting is a distance-compensated technique. Fig. 3 shows an example of a distance-compensated structured light system that uses a projector to create structured light patterns, providing improved contrast over wide field illumination [36]. Post-processing compensates for the water attenuation based on recovery of the object distance map.

A second method commonly used with structured lighting is synthetic aperture illumination. While most methods of structured light are based on illumination from a single direction, methods such as that devised by [37] (shown in Fig. 4) use a constellation of illumination sources, each illuminating the scene from a unique position and direction. Multiple frames are acquired while different sets of illumination sources are active, where each combination produces a different illumination pattern. The acquired frames contain backscatter similar to that obtained by floodlighting. When the data are post-processed, the backscatter field is estimated based on the set of frames and then the backscatter component is compensated for to enhance the image quality.

![Figure 3. Example of a distance-compensated structured light system.](image)

![Figure 4. Experimental setup using multiple illumination sources.](image)

Other methods relying on spatial coherence, such as those conceived by [38, 39] using structured illumination for object shape recovery and distance imaging are limited by practical size of the required optics and have not been investigated due to practical implementation problems for scenes of any reasonable size.

E. Polarization Discrimination

An image enhancement technique demonstrated most recently in [40] combines an optical step during image acquisition along with digital post-processing. The optical step uses wide-field polarized light to irradiate the scene, while viewing the scene via an additional polarizer. Two wide-field frames are taken in mutually orthogonal polarization states. Backscatter exists in both polarization states (frames) but in different amounts; hence the optical step modulates the backscatter. A mathematical process is applied using the two raw frames as input, extracts the backscatter field, and then estimates the background free of backscatter.
F. **Multiple Perspective Image Construction**

Imagery of a scene collected from different locations is commonly used to derive size and depth measurements, photo-mosaics, and 3-D reconstructions. This can be accomplished by performing high resolution optical reconnaissance sweeps of a desired area using a single imaging system, or using multiple imaging systems that perform the sweeps in a fraction of the time. When a multiple system technique is employed that separates the illumination from the image formation process, images can be captured at greater distances due to a reduction of the backscatter component. Examples of different configurations of illumination and cameras, such as those shown in Fig. 5, have been simulated by [41].

Woods Hole Oceanographic Institution researchers are implementing a slightly different technique using two AUV’s to cooperatively characterize the Arctic seafloor [42]. Unlike the approach in [41], two AUV’s are launched sequentially with different objectives. The first AUV, Puma or “plume mapper,” is launched to localize chemical and temperature signals given off by hydrothermal vents; while the second AUV, Jaguar, is sent to those locations to use high-resolution cameras and bottom-mapping sonar to image the seafloor.

![Multiple Perspective Imaging Concepts](image1.jpg)

Finally, worth mentioning here are software techniques employed real-time to enhance video quality such as LYYN Visible Enhancement Technology (www.lyyn.com) and tools made available in several image processing software development kits [43].

IV. **Spatial Coherence Methods**

Holography is a technique utilizing the spatial coherence of wavefronts to record 3-D information by forming the interference pattern between waves propagating from an object and a reference wave – usually at optical frequencies. Conventionally, the interference pattern modulus is recorded and played back on film, but today reconstruction can be done digitally by mathematically applying the rules of diffraction on the hologram input data. An electronic hologram can be recorded directly onto a CCD camera and then numerically reconstructed. As a result, 3-D information can be recorded in real-time producing 3-D videos of living organisms and particles in their natural environment.

![Hologram Images of Calanoid Copepods](image2.jpg)

(a) In-line recording at 70m (2.2 x 0.7 mm). (b) An enlarged detail of (a). (c) Off-axis recording at 60m (~ 6mm long).
Holography is possible over short distances in seawater and is used by submersible digital holographic systems, as described in [44-47], to record plankton [48] and other millimeter-sized marine organisms. Examples are shown in Fig. 6. Demonstrated systems operate at stand-off distances of 0.1 mm to several millimeters, are not plagued by backscatter or attenuation effects, and offer considerably greater depth of field than possible with conventional microscopy – allowing a single hologram to provide instantaneous volumetric information.

V. MULTI-DIMENSIONAL METHODS IN IMAGE SPACE

Methods that combine Dual frequency IDentification SONar (DIDSON) and stereo imagery are being investigated for multiple-view 3-D reconstruction of scenes [49, 50]. The intent is to use the sonar to enhance reconstruction in poor visibility conditions, where visual cues become less informative. DIDSON uses high-frequency sonar (1-2 MHz) to produce range and azimuth 2-D measurements that are acquired in a polar coordinate system at operational ranges of 10 to 20 meters in turbid water. Since the geometry of “acoustic imagers” differs drastically from those of “optical lensless pinhole imagers,” the greatest challenge in combining sonar and stereo images is calibrating the system to ensure data model consistency [51, 52]. Not only do the sensors have different areas of coverage, but a pixel in polar coordinates maps to a collection of pixels in the Cartesian coordinate system, further complicating the process of searching and matching feature points in successive images. Other challenges specific to DIDSON include limited resolution, low SNR, and limited range of sight. Ref. [53] developed an algorithm to enhance sonar video sequences by incorporating knowledge of the target object obtained in previously observed frames. This approach involves inter-frame registration, linearization of image intensity, identification of a target object, and determining the “maximum posteriori fusion of images” in the video sequence.

DIDSON as a stand-alone sensor offers numerous advantages and commercial software packages are becoming more available to process this data (e.g., Acoustic View for mosaics, www.acousticview.com). In high-frequency mode, images from the sound beams can show the outline, shape and features of a target object. A common application is fisheries management and assessment, where fish behaviors such as spawning, feeding and migration can be non-invasively monitored and recorded in low visibility conditions [54-56]. However, there still can be conditions that limit the capability of acoustic sensing [57] so a better solution may be a hybrid – combining DIDSON and a video camera [58, 59], stereo camera system [60] or other optical imaging system.

VI. FUTURE TRENDS

Several areas of technology development are particularly notable when anticipating future advancements in underwater imaging technology. Compact high power light sources, data compression and management, energy storage, and realistic recreation of the image space continue to be areas where major strides are being made.

A. LED Technology

White light and single wavelength light emitting diodes (LEDs) have made rapid advancements in the past few years with respect to electrical to optical conversion efficacy and wattage. Single package units are available that can produce output of several hundred lumens – equivalent to 60-watt incandescent sources, but at a fraction of the power. A measure of the output normalized to the wattage is the “luminous efficacy”. This can range from 50 lumens per watt to 70 lumens per watt for fluorescent lamps, and as much as 50 and 150 lumens per watt for arc light and HMI sources, respectively. White light LEDs promise efficiencies of 150 lumens per watt (or greater) and can be powered from low voltage sources without expensive or bulky ballasts. This makes them particularly useful for battery powered applications such as dive lights, small AUVs or other vehicles as well as other restricted power sources (e.g., offshore buoys [61]). As the technology advances, higher power units will become available and will ultimately replace lamps that currently are rated at hundreds of watts or more.

B. Laser Technology and Detector Advancement

New developments in laser technology have achieved short pulses at blue-green wavelengths offering stable, high repetition rate (>500 kHz) and average power (>2W) laser sources (Q-Peak Inc., Aculight Corp.). In an LLS system, this technology enables temporal separation of the backscatter from the reflected target signal and allows primarily the integration of the target photons during the detection process. This tends to increase the SNR allowing a greater distance for detection/imaging of the target. Further refinements in laser technology will be higher speed modulation capability and greater uniformity of pulse-to-pulse energy stability, as well as increases in power, efficiency, and compactness. This will allow more advanced laser imaging systems to be developed and integrated onto small platforms. Development of detectors with extreme speed, high dynamic range, and accurate response to modulated waveforms will also be important.
C. Data Management

Managing data from multiple disparate sensors as well as sensors that produce vast amounts of data is currently a challenge and will continue to be so as technology continues to advance. Data management involves storing, cataloguing, searching, retrieving, interpreting (human in the loop), sharing, editing, reusing, distributing, archiving and thinning. Though specialized systems will need to be developed for custom data types, rich media management and Digital Asset Management (DAM) software for high definition video (for example) and its related data is commercially available (Virage, Artesia). These software packages include features such as automatically capturing, encoding and indexing TV, video and audio content from any source dynamically; automatically generating a comprehensive range of metadata that is immediately fully searchable and accessible by any user; and full screen streaming of HD video content at reduced costs.

Manipulation of the data goes hand-in-hand with data management. Viewing and processing high volumes of data are often cumbersome for both the computer and the user due to memory swapping, slow processing and rendering of the data, and the inability to see all of the data on a typically-sized display. Specialized computers for high-end gaming and video editing are becoming popular, as well as large and multiple widescreen displays. As an example, Alienware presented a “giant ‘curved’ widescreen” display at the 2008 Consumer Electronics Show in Las Vegas. This display is equivalent to two slightly bent 24-inch liquid crystal display (LCD) screens glued together [62].

D. Three-Dimensional Television (3DTV)

A study undertaken by European researchers is exploring the feasibility of 3-D television (3DTV) [63]. Developers of eHoloCam are lending their expertise in holographic imaging to this cause [64, 65]. In another example, deformable meshes and other generic 3-D motion object representation tools found in computer graphics technology provide many of the tools necessary for 3-D scene representation.

ACKNOWLEDGMENT

The authors are grateful for the contributions of Yoav Schechner of Technion – Israel Institute of Technology for his valuable contribution to the original manuscript; Jules Jaffe of Scripts Institution of Oceanography at University of California for providing an update of his recent research; and Marc Levoy, Srinivas Narasimhan and Linda Mullen for data and figures used in this paper.

REFERENCES


