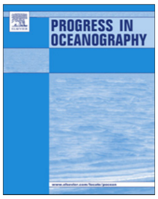
Progress in Oceanography 149 (2016) 106–120

Contents lists available at ScienceDirect

Progress in Oceanography

journal homepage: www.elsevier.com/locate/pocean



Review

Current and future trends in marine image annotation software

Jose Nuno Gomes-Pereira[a,x,*], Vincent Auger[w], Kolja Beisiegel[b], Robert Benjamin[c], Melanie Bergmann[d], David Bowden[e], Pal Buhl-Mortensen[f], Fabio C. De Leo[g], Gisela Dionísio[h,i,x], Jennifer M. Durden[j,k], Luke Edwards[l], Ariell Friedman[m], Jens Greinert[n], Nancy Jacobsen-Stout[o], Steve Lerner[p], Murray Leslie[g], Tim W. Nattkemper[n], Jessica A. Sameoto[c], Timm Schoening[n], Ronald Schouten[g], James Seager[q], Hanumant Singh[p], Olivier Soubigou[r], Inês Tojeira[s], Inge van den Beld[t], Frederico Dias[u], Fernando Tempera[a,v], Ricardo S. Santos[a]

[a] MARE – Marine and Environmental Sciences Centre, IMAR, Department of Oceanography and Fisheries, University of the Azores, 9901-862 Horta, Azores, Portugal
[b] Leibniz Institute for Baltic Sea Research, Warnemünde, Department of Biological Oceanography, Seestraße 15, 18119 Rostock, Germany
[c] Fisheries and Oceans Canada, Bedford Institute of Oceanography, PO Box 1006, Dartmouth, Nova Scotia, Canada
[d] HGF-MPG Group for Deep-Sea Ecology and Technology, Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung (AWI), Am Handelshafen 12, D-27570 Bremerhaven, Germany
[e] Coasts and Oceans Centre, National Institute of Water and Atmospheric Research, Wellington, New Zealand
[f] Institute of Marine Research, N-5017 Bergen, Norway
[g] Ocean Networks Canada/Biology Department, University of Victoria, BC, V8W 2Y2, Canada
[h] Departamento de Biologia & CESAM, Universidade de Aveiro, 3810-193 Aveiro, Portugal
[i] MARE – Marine and Environmental Sciences Centre, Laboratório Marítimo da Guia, Centro de Oceanografia, Faculdade de Ciências da Universidade de Lisboa, Av. Nossa Senhora do Cabo, Cascais, Portugal
[j] Ocean and Earth Science, University of Southampton, National Oceanography Centre, University of Southampton, Waterfront Campus, European Way, Southampton, UK
[k] National Oceanography Centre, University of Southampton, Waterfront Campus, European Way, Southampton, UK
[l] iVEC, 'Advancing science through supercomputing', Edith Cowan University, Mt Lawley, Building 13, Western Australia, Australia
[m] Australian Centre for Field Robotics, Rose St Bldg J04, University of Sydney, Sydney, NSW 2006, Australia
[n] GEOMAR Helmholtz Centre for Ocean Research, Kiel, Germany
[o] Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA 95039, United States
[p] Woods Hole Oceanographic Institution, 266 Woods Hole Rd., MS# 07, Woods Hole, MA 02543-1050, United States
[q] SeaGIS Pty. Ltd., PO Box 1085, Bacchus Marsh, VIC 3340, Australia
[r] Institut Français de Recherche pour l'Exploitation de la Mer, Département Infrastructures Marines et Numériques, Unité de recherche Navires et Systèmes Embarqués, Ifremer Brest, ZI Pointe du Diable, 29280 Plouzane, France
[s] Portuguese Task Group for the Extension of the Continental Shelf (EMEPC), Rua Costa Pinto 165, 2770–047 Paço de Arcos, Portugal
[t] IFREMER, Centre de Bretagne, Laboratoire Environnement Profond, CS10070, 29280 Plouzané, France
[u] Task Group for the Extension of the Continental Shelf (EMEPC), Rua Costa Pinto, n°165, 2770-047 Paço de Arcos, Portugal
[v] European Commission, Joint Research Centre, Inst. for the Environ. and Sustainability, Water Unit, 46/107, 21027, Italy
[w] Canadian Scientific Submersible Facility, 110-9865 West Saanich Rd., North Saanich, BC, V8L 5Y8, Canada
[x] Naturalist, Science & Tourism, 9900-029 Horta, Azores, Portugal

ARTICLE INFO

Article history:
Received 5 September 2014
Received in revised form 9 July 2016
Accepted 11 July 2016
Available online 28 September 2016

Keywords:
Underwater visual surveys
Image analysis
Image annotation
Data collection
Data storage
Monitoring
Marine imaging

ABSTRACT

Given the need to describe, analyze and index large quantities of marine imagery data for exploration and monitoring activities, a range of specialized image annotation tools have been developed worldwide. Image annotation – the process of transposing objects or events represented in a video or still image to the semantic level, may involve human interactions and computer-assisted solutions. Marine image annotation software (MIAS) have enabled over 500 publications to date. We review the functioning, application trends and developments, by comparing general and advanced features of 23 different tools utilized in underwater image analysis. MIAS requiring human input are basically a graphical user interface, with a video player or image browser that recognizes a specific time code or image code, allowing to log events in a time-stamped (and/or geo-referenced) manner. MIAS differ from similar software by the capability of integrating data associated to video collection, the most simple being the position coordinates of the video recording platform. MIAS have three main characteristics: annotating events in real time, posteriorly to annotation and interact with a database. These range from simple annotation interfaces, to full onboard data management systems, with a variety of toolboxes. Advanced packages allow to

---

* Corresponding author.
E-mail addresses: gomespereira@mare-centre.pt, zn.multimedia@gmail.com (J.N. Gomes-Pereira).

http://dx.doi.org/10.1016/j.pocean.2016.07.005

input and display data from multiple sensors or multiple annotators via intranet or internet. Posterior human-mediated annotation often include tools for data display and image analysis, e.g. length, area, image segmentation, point count; and in a few cases the possibility of browsing and editing previous dive logs or to analyze the annotations. The interaction with a database allows the automatic integration of annotations from different surveys, repeated annotation and collaborative annotation of shared datasets, browsing and querying of data. Progress in the field of automated annotation is mostly in post processing, for stable platforms or still images. Integration into available MIAS is currently limited to semi-automated processes of pixel recognition through computer-vision modules that compile expert-based knowledge. Important topics aiding the choice of a specific software are outlined, the ideal software is discussed and future trends are presented.

Contents

1. Annotation of video and still images from the marine environment

Underwater photography has been used extensively to quantify the marine environment since the middle of the 20th century, and the amount of underwater digital media being captured has grown exponentially in the last decades (Vevers, 1951; Jaffe et al., 2001; Solan et al., 2003). In parallel, an increasing number of research and monitoring teams became interested in describing and analyzing large quantities of imagery data collected underwater. Recent examples of mass imagery collection for scientific purposes include the regular remotely-operated vehicle (ROV) operations in Monterey Bay (Schlining and Jacobsen Stout, 2006), the National Oceanic and Atmospheric Administration live under sea video feeds from the ship *Okeanos Explorer*, or the seabed video continuously streaming to the world-wide web from Canadian observatories (Leslie et al., 2010). Underwater imagery is used to log events related to technical operations, or to answer scientific questions within the disciplines of biology and geology. For example, still images and motion video have been used in ecology to count and track individuals in different habitats and at different spatial and temporal scales (e.g. Benoit-Bird and Au, 2006; Huffard et al., 2016; Porteiro et al., 2013). The imagery type, platforms used and camera orientations can be varied to suit the scientific aims, and both discrete and continuous events may be recorded, including size and area measurements.

Annotation is a wide concept (from text documents to digital databases), and what is less ambiguous is the *process*, which we call annotation, of adding data to something that already exists. In the digital imaging community, annotation is often accomplished by superimposing the annotated data on an image. Image annotation also includes translating objects or events in an image to the semantic level, representing the actual content recorded in the image or video (e.g. describing an object or an animal behavior). Therefore, image annotations can be stored with the imagery, facilitating data management and the accessibility of information.

Scientists have endeavored to develop annotation tools for fast labelling and data retrieval since the early 1990s. This early work focused on analogue photographs and videotape, and later moved on to computer-supported investigation (Kipp, 2001). The emergence of technology to aid image annotation increased reliability,

Table 1

List of marine image annotation software packages under comparison in this paper; Acronym and full name, year of software release and year of recent updates, developing institutes and contact person; *consortia of Spanish and French institutes, see Appendix I for list of contributing Institutes (Within contacts, D = Developer, Us = User Support).

Acronym of name	Name	Release/updates	Developing institutes	Contact	Website
ADELIE	Underwater vehicle data post-processing software	1998/ongoing	Ifremer (FR)	adelie@ifremer.fr (D) Olivier Soubigou	Link
BIIGLE	Bio-Image Indexing, Graphical Labelling and Exploration	2007/ongoing	Center of Biotechnology (CeBiTec), Faculty of Technology, Bielefeld University and Alfred Wegener Institute for Polar and Marine Research (AWI) (DE)	info@biigle.de (D) Tim Nattkemper, tim.nattkemper@uni-bielefeld.de, (Us) Melanie Bergmann, Melanie.Bergmann@awi.de	Link
CATAMI	Collaborative and Automation Tools for Analysis of Marine Imagery and Video	2012/2013	National eResearch Collaboration Tools, iVEC and Resources/Australian National Data Service (AU)	catami@ivec.org, (project manager/Us) Luke Edwards	Link
ClassAct Mapper	ClassAct Mapper	2002/2008	Fisheries and Oceans Canada (CA)	(D) Robert Benjamin, Robert.Benjamin@dfo-mpo.gc.ca, (Us) Jessica A. Sameoto, Jessica.Sameoto@dfo-mpo.gc.ca	Link
COVER	Customizable Observation Video imagE Record	2010/2010	Ifremer (FR)	(D) Cyril Carré, (Us) I. van den Belt	Link
CPCe	Coral Point Count with Excel extensions	2005/ongoing	National Coral Reef Institute, Oceanographic Center, Nova Southeastern University (US)	ncri@mail.ocean.nova.edu	Link
DIAS	DIAS Image Annotation Software	2015/ongoing	Bielefeld University (Tim W. Nattkemper), developed under the JPI Oceans project Ecological Mining Impact	(D) Timm Schoening, tschoening@geomar.de	Link
EventMeasure	Event logging & 3D measurement	2008/ongoing	SeaGIS (AU)	info@seagis.com.au, (D) Jim Seager	Link
FISH_ROCK	A Tool for Identifying and Counting Benthic Organisms in Bottom Photographs	2006/ongoing	Woods Hole Oceanographic Institution, NOAA (US)	(D) Hanumant Singh, hsingh@whoi.edu	Link
Frame-Grabber	The Alvin Frame-Grabber System	2002/ongoing	Woods Hole Oceanographic Institution, NOAA (US)	(D) Steven Lerner slerner@whoi.edu, Dan Fornari and Barrie Walden	Link
ImageJ	Image Processing and Analysis in Java	(1987) 1995/ongoing	National Institute of Mental Health, Bethesda, Maryland (US)	(D) Wayne Rasband, wayne@codon.nih.gov	Link
IRLS	Integrated Real-time Logging System	1999/ongoing	Canadian Scientific Submersible Facility	(D) Vincent Auger, auger@ropos.com	Link
NICAMS	NIWA Image Capture and Management System	2010/ongoing	National Institute for Water and Atmospheric Research (NIWA, NZ)	(Ds) David Bowden (David.Bowden@niwa.co.nz), Brent Wood Brent.Wood@niwa.co.nz	Link
OFOP	Ocean Floor Observation Protocol	2000/ongoing	Scientific Abyss Mapping Services, Rodenbek (DE)	ofop@ofop-by-sams.eu, (D) Jens Greinert	Link
photoQuad	Photo quadrat analysis software	2012/ongoing	University of the Aegean (GR)	photoquad@hotmail.com, (D) Vasilis Trigonis & Maria Sini	Link
Seascape	Segmentation and Cover Classification Analyses of Seabed Images	2011/ongoing	Large Consortia* (ES, FR)	(D) Núria Teixidó	Link
SeaScribe/ Seatube	NEPTUNE Canada's real-time georeferenced library of deep sea video	2010/ongoing	Ocean Networks Canada, University of Victoria (CA)	(D) Ronald Schouten and Eric Guillemot (user support), Fabio C. De Leo, fdeleo@uvic.ca	Link
Squidle	SQUIDLE: a centralised web-based framework for management, exploration and annotation of marine imagery	2014/ongoing	Australian Centre for Field Robotics, University of Sydney (AU)	(D) Ariel Friedman a.friedman@acfr.usyd.edu.au	Link
TransectMeasure	Single camera biological analysis tool	2007/ongoing	SeaGIS (AU)	(D) Jim Seager info@seagis.com.au	Link
VARS	Video Annotation & Reference System	2005/ongoing	Monterey Bay Aquarium Research Institute (US)	(D/user support) Nancy Jacobsen Stout, jana@mbari.org; Brian Schlining, brian@mbari.org	Link
VIDEOMON	VIDEOMON MONitoring Software	2014/ongoing	Leibniz Institute for Baltic Sea Research Warnemünde (IOW) (DE)	(D) Kolja Beisiegel, kolja.beisiegel@io-warnemuende.de, Steffen Bock	Link
Video Navigator	VideoNavigator	2009/ongoing	Institute of Marine Research (NO)	(Us) Paul Buhl-Morthensen, paalbu@imr.no	Link
VirtualVan	The Jason II Virtual Control Van System	2001/ongoing	Woods Hole Oceanographic Institution, NOAA (US)	(D) Steven Lerner slerner@whoi.edu, Andrew Maffei amaffei@whoi.edu	Link

repeatability and workflow optimization (Burr, 2006). Software packages used during manual annotation consist of a Graphical User Interface (GUI), an interface that allows users to interact with electronic devices through graphical icons and visual indicators, including a video player or thumbnail viewer. Annotation takes place by identifying objects in time by using shapes or polygons and recording descriptors with customizable buttons or lists.

The development of image annotation software has been oriented to film and video production and archiving, particularly for education, sports, and home entertainment purposes (Hauptmann, 2005), but also for scientific purposes, such as studies of human behavior (Rosenblum et al., 2004). There are several software available for manual video content annotation such as ANVIL (Annotation of Video and Language; Kipp, 2001), ImageJ (Schneider et al., 2012) or VATIC (Video Annotation Tool from Irvine, California; Vondrick et al., 2013). While these software can be used to annotate virtually any type of imagery, specialist software to annotate images collected from underwater environments often allow the integration of additional data related to the image capture that assist in interpretation, such as position, depth and camera settings.

A series of specialized underwater image annotation software have been developed in the last two decades. These have evolved from stand-alone players (i.e. software that runs individually and not as a part of a bundle or package), to software with GUIs that run within more complex environments, such as with relational databases or integrate logging systems (e.g. Integrated Real time Logging Systems; Juniper et al., 2000). Current uses range from simple annotation to assisting with platform deployments, logging sensors and events at sea, to processing data from sensors and annotations.

Here we present the first comparative review of marine image annotation software (MIAS), providing users and software developers an overview of the functionalities and capabilities of 23 commonly used existing programs. Textual and graphical explanations are given on the general functions, structure, customization, annotation interfaces, output files, as well as costs and availability. We describe the capability of these software to operate during imagery capture (in real time) and posteriorly to data collection, and how these software interact with databases to simplify workflows and data management. The current status of computer-assisted annotation is also presented. Finally, cues to aid the choice of software are discussed, and the ideal software package is described with a look to developmental trends.

2. Software functions and specifications

2.1. Marine image annotation software (MIAS) under comparison

We compiled information on 23 software systems currently used in the annotation and analysis of underwater imagery (listed in Table 1) to understand their origin, main components and functions. Software not exclusive to underwater image analysis, but known to be used for the analysis of underwater images were also included. A short description of each software package including the developing institution, general functions and selected references can be found in Appendix A, Supplementary data I.

This review is written with no intention to favor any particular software, or to select the best software available. In this sense, (i) all software packages referred to in tables and text are presented in alphabetical order; (ii) the list of software is not exhaustive, including a sample of the software tools frequently in use today, and therefore some may have been unintentionally omitted. Other solutions have been developed for management and distribution of video content (e.g. Japan Agency for Marine-Earth Science and Technology E-Library of Deep Sea Images; Kitayama et al., 2012), and were not included by falling out of the scope of this work.

General and advanced specifications of each MIAS are compiled in Tables 2 and 3. Table 2 includes information about the type of imagery accommodated (i.e. video or still images); operating modes, database connectivity, costs (€) and software characteristics. Table 3 lists advanced specifications, such as data formats, capability of multiple or automated annotation, and features for data management such as ingestion of metadata and annotation data and querying/editing annotation files.

The MIAS under review were developed and released during the last two decades in European, North American and Oceanian countries (Table 1). Large state-funded academic and governmental institutes have developed the majority of MIAS. Software has also been made available from private companies, and in both cases, these are designed for marine research or environmental monitoring purposes. Many packages were initially developed at a particular institute for a particular annotation purpose, and have eventually become freely or commercially available, often after further development.

A bibliometric analysis was performed to obtain an indication of the contribution of selected MIAS for content production. Despite not being a true quantitative measure, because MIAS are not always listed in publications, hampering founded comparisons (as explained below), it provides a rough approximation on the actual number of publications (Datta et al., 2005), outlining the importance of MIAS for marine research. The acronym of each software was used as a search term in Google Scholar (www.scholar.google.com; consulted 10th of June 2014). All resulting references were checked individually. Confounding terms were removed and additional descriptors were added every time these resulted in an increasing number of positive results and reduced the number of negative results (e.g. developing institute, software developer, etc.; the list of search terms and results is presented in Appendix B). Over 500 citations were identified, including conference proceedings, dive reports and academic theses. Results are presented in Fig. 1 by year of software development.

An evaluation of the number of bibliographic references that employed each specific annotation software system shows an expected trend where the oldest software (i.e. by year of development) appeared in a larger number of publications (Fig. 1). Two freely available packages – ImageJ and CPCe – were not represented from Fig. 1 as these summed several hundred publications. ImageJ is an open-source, community-driven generalist image processing software that supports all common image manipulation. From the exceedingly large number of results (>500), it was impossible to determine the accurate number referring to imagery collected underwater. CPCe, a tool largely used by the research community interested in monitoring shallow water tropical coral reef systems, was the first software designed for area-based analysis using still images that became freely available. The large number of publications referring the software packages Adelie, OFOP, VARS and EventMeasure reveal that popular choices among the scientific community can include both commercial and open-source solutions. The adoption of a software package by a research group or a research community is driven by a matrix of factors, which are not always straightforward, as discussed below.

2.2. General structure and operation mode

Three main features can be found in MIAS which determine their overall structure and function, including: (i) the capacity to operate during data collection, hereafter designated to be operating in ‘real-time’ mode (e.g. onboard vessels), (ii) the ability to operate following data collection, the posterior annotation mode,

Table 2

General specifications of marine image annotation software packages under comparison. Video/Images refer to type of image data input; Real-time, indicates if the software allows video annotation during platform deployment; Posterior annotation, video annotation not developed in real time; Database connection, if the software interacts with a database with data from different deployments; Costs (€) software availability costs in €uros, ± given values are only indicative and include initial support and training; Code availability, if software programming language is available, i.e. Open source; Language, programming language for annotating interface and other modules; Output file, refers to format; Interface customization, the possibility of customizing the interface layout (besides field codes), including three customization levels: basic, intermediate (e.g. XML) and expert (more complex languages, e.g. Java, Matlab); OS, Operating System; for additional tools see Appendix I. Features not available on each software are indicated with a minus "–" sign. (A) is single user database; (B) possible but not developed for that purpose; (C) it does not interact by default but it can exports data strings to databases in a specified, NMEA style format (NMEA is a combined electrical and data specification for communication between marine electronics); (D) interface and other software (html5, JQuery, Bootstrap, Backbone.js), other software (Python, Django, PostgresSQL, GeoServer, Lire); (E) the server is written in Perl and is available for Unix platforms, for server plotting capabilities, GMT 3.x, Matlab 5.x, and Fly v1.6; (F) can be made available by contacting developers; (G) EIC, Electronic Index Cards is a simple self-documenting ASCII data format.

Software acronym	Imagery input	Real time annotat.	Posterior annotation	Database connection	Language	OS	Costs (€)±	Code availability	Output file
ADELIE	Video	Yes	Yes	Yes (A)	C++, C#, Visual Basic	Win	From free for research, to license agreement	–	dbf
BIIGLE	Images	–	Yes	Yes	Flex 3 SDK by Adobe	All. Web based	Free	–	csv
CATAMI	Images	–	Yes	Yes	CSS, HTML, Javascript, Python (D)	All. Web based	Free	Yes	csv
ClassAct Mapper	Video	Yes	Yes	Yes (A)	Delphi v6	Win; Uses ODBC to connect to database	N/A	–	.mdb
COVER	Video/Images	Yes (B)	Yes	–	Java	Win	Free	Yes	csv (hierarchical)
CPCe	Images	–	Yes	–	Visual Basic	Win	Free	–	xls
DIAS	Images	–	Yes	Yes	JavaScript, PHP	All. Web based	Under development	–	csv
EventMeasure	Video/Images	–	Yes	–	C++	Win	640€ (student) to 2000€	–	txt
FISH_ROCK	Images	–	Yes	Yes	Matlab	Mac, Linux, Win	Free	Yes	txt
Frame-Grabber	Video/Images	Yes	Yes	Yes	Perl	All. Web based	N/A	–	ASCII (EIC)(G)
ImageJ	Images	–	Yes	–	Java	Mac, Linux, Win	Free	Yes	txt and other
IRLS	Video/Images (up to 6 video feeds)	Yes	Yes	Yes	Java, PHP	All. Web based	Free for ROPOS ROV users	–	HTML, CSV, KMZ, IRLS server
NICAMS	Images	–	Yes	Yes	Java	Mac, Linux, Win	Free	Yes	Direct to PostGIS Database
OFOP	Video	Yes	Yes	Yes (C)	Visual Basic6	Win	1st license 450€ 10 years; next 400€	–	txt
photoQuad	Images	–	Yes	–	Matlab	Win	Free	–	csv
Seascape	Images	–	Yes	–	C++	Mac, Linux, Win	Free	Yes	XLS or ASCII
SeaScribe/ Seatube	Video	Yes	Yes	Yes	Java	Win	N/A (SeaScribe) Free (Seatube)	–	HTML, CSV, KMZ, other.
Squidle	Images	–	Yes	Yes	Javascript & python	All. Web based	Free for research	Yes	CSV, excel, HTML, other.
TransectMeasure	Video/Images	–	Yes	–	C++	Win;	From 620 €±	–	txt
VARS	Video/Images	Yes	Yes	Yes	Java	Linux, Mac, Win	Free	Yes	csv, kml
VIDEOMON	Video	Yes	–	–	C#	Win	N/A (F)	–	.csv
Video Navigator	Video	–	Yes	–	Visual Basic	Win	N/A	–	txt
VirtualVan	Video/Images	Yes	Yes	Yes	Perl (E)	Linux User Interface: WebBrowser	N/A	–	HTML, csv, kml

Table 3

Advanced features of marine image annotation software under comparison. The fields are described as follows. Annotating fields, Real-time includes Video Input/Capture still image from video (i.e. if software receives video signal in real-time and if it allows capturing a still image from video in real-time); Other inputs/formats, refers to types of inputs, and how the information is imported to the software; Display Data refers to types of data display; Posterior annotation fields include data display (e.g., ship and platform position, depth, temperature, etc.); Browse annotation file/edit (re-annotate), if the software allows browsing annotation file and edit current annotations; Multiple Annotation, allows annotating the same video content from different computers (simultaneously); Capture still image from video, allows capturing a still image from video; Browse pictures, allows browsing through images; Image Analysis, includes tools to analyze images (e.g. annotate, measure, generate points, etc.); Annotation Data Analysis, calculate basic statistics and other analysis from annotated data; Process navigation data, tools to process raw navigation data (e.g. remove outliers, spline); Integrate annotation file, integrate annotation file with data from other sensors; Automated Annotation (does not including annotation of sensors or cameras metadata) – Event Detection, Algorithms for automated event detection; Object Identification, Algorithms for automated object identification; Database, Functions and Format, Software interaction with database; Browse database, browse through database displaying results; – refers to feature not available. Features not available on each software are indicated with a minus "–" sign. (Note that all features mentioned refer solely to the specific software, and in some cases other software is used for a particular task not available in the UIAS.)

Specifications/ acronym	ADELIE	BIIGLE	CATAMI	ClassAct	COVER	CPCe	DIAS	EventMeasure	FISH_ROCK	Frame-Grabber	ImageJ	IRLS	NICAMS	OFOP	photoQuad	Seascape	SeaScribe	Sqiudle	TransectMeasure	VARS	VIDEOMON	Video Navigator	VirtualVan
Descriptors ofimage ann otation interface	Buttons Lists (tree) Rulers	Lists	Lists	Buttons Lists	Buttons Lists and Rulers	Buttons	Buttons Lists	Buttons	Buttons Lists	Text Entries	–	Buttons Lists	Buttons Lists – Database search	Buttons Lists	Buttons Lists Rulers	Lists	Buttons Lists	Buttons Lists	Buttons	Buttons Lists	Buttons Touch-screen	Buttons Lists	Buttons Lists
Real-time Video input/capture still image	Y/Y	–	–	–	–	–	–	–	–	Y/Y	–	Y/Y	Still image	Y/Y	–	–	Y/Y (G)	–	–	Y/Y	–	–	Y/Y
Receive other data inputs	–	–	–	Y (N)	–	–	–	–	–	Y	–	Y (U)	Y (P)	Y (P)	–	–	Y (O)	–	–	–	Y(X)	–	Y
Display data	Y	–	–	Depth Navigation	–	–	–	–	–	Depth Nav Sci Inst	–	All ROV Data	–	All Ship & Sub + customizable	–	–	Sub, +other metadata	–	–	Y	Y	–	Depth Nav Sci Inst
Posterior annotation Display data	Y	Depth, Nav, Temp, etc.	Nav.	–	Depth	–	Nav.	–	Date Time Depth Nav. Alti Area	Nav. Attitude other	–	All ROV Data, logger name	–	Ship & Sub – Nav.	–	–	–	Depth, nav, meta data, terrain, mosaics	–	Y	–	Y	Depth Nav Sci Inst
Browse annotation file/edit	Y/Y	Y (B)/Y	–/–	Y (J)/Y	–/–	–/Y	Y/Y	Y/Y	–/Y	Y/–	–/–	Y/Y	Y/Y	–/Y (D)	–/Y	–/Y	–/Y	–	Y/Y	Y/Y	–	Y	Y/–
Multiple annotation	–	Y	Y	–	–	–	Y	–	–	–	–	Y	Y	Y	–	–	Y	Y	–	–	–	Y	–
Extra tools Capture still image from video	Y	–	–	–	Y	–	–	Y	–	Y	–	Y	–	Y (Y)	–	–	– (G)	–	Y	Y	–	–	Y
Browse pictures	Y	Y	Y	–	Y	Y	Y	Y	Y	Y	–	Y	Y	–	–	–	–	Y	Y	Y	–	–	Y
Other tools for image analysis	–	Y	Y	–	Fixed and random point count	Random point count	Length or area, point count	Y	Length or area meas.	–	Y	–	Y	–	Y (M)	Image Segmentation	–	Y	Y	Y	–	–	–
Annotation data analysis	Y	Y	–	–	–		–	Y	–	–		–	–	–			– (B)	Y	Y	– (B)	–	–	Y
Process navigation data	Y	–	–	–	–	–	–	–	–	–	–	–	–	Y	–	–	Y	.	–	–	–	–	Y
Integrate annotation file	Y	–	–	–	–	–	–	–	–	Y	–	Y	–	Y	–	–	Y	Y	–	Y via database	–	–	Y
Automated annotation Event detection/ object identification	–/–	–(H)/Y (I)	–/Y	–/–	–/–	–/–	–/–	–/–	–/–	–/–	–/–	–/–	–/–	–/–	–/–	–/–	–/–	–/Y	–	–	–	–	–/–
Database Functions and format	Y (T)	MySQL server	–	MDB (E)	–	–	PostgreSQL	–	Matlab	GeoBroswer/ EIC	–	MySQL	PostGIS or Spatialite	–	–	–	Y (R)		–	Y (S)	–	–	GeoBroswer/ EIC
Browse database	–	Y (B)	–	via MDB	–	–	Y	–	Matlab	Y	–	Y	db search tools or image-by-image review	–	–	–	Yes (via database application)	Yes	–	Y	–	–	Y

(A) There are currently three distinct versions of the Annotation system, for: (1) video tapes, (2) still images, and (3) digital video files. The latter two are currently only available for post-dive editing, not in real-time; (B), through multiple tools that query the Database; (C) ongoing developments; (D) it does not allow browsing through annotation file, but as the annotation output is displayed, particular lines can be deleted; (E) See text for more explanation; (F) Automation is for whole image and points; (G) Frame grabs are taken each second; (H) Under development; (I) In the connected iSIS, Delphi can automatically detect laser-points and calculate the imaged area (camera footprint); (J) Two modes available: (1) Synchronizes with time stamp and overwrites, (2) Synchronizes with time stamp and adds to (Example: process video for substrate classification then re-process for Fish.); (L) Option 1: up to 8 unique events with associated counter buttons as well as manual entry count and size input fields. Option 2: up to 45 customizable buttons – when button is pressed it records a single count of this event; (M) See Appendix I for more details; (N) Serial connection to GPS data (Geo-stamp)/NMEA; (O) Sub position, and other metadata; (P) SHIP, SUB, ROV, attitude, chemical, etc. via COM or TCP-IP/UDP, various formats (e.g. NMEA, user specific); (R) Uplink system in real time (Apache, MySQL and PHP); (S) Real-time data transferred on land to Microsoft SQL Server databases (also compatible with MySQL, Apache, PostgreSQL, and others); (T) Management of several tables (navigation, attitude, sensors, etc.); U, ROV data (NMEA string); X, ship data (NMEA); Y, customizable overlay on image (position, depth, time, etc.).

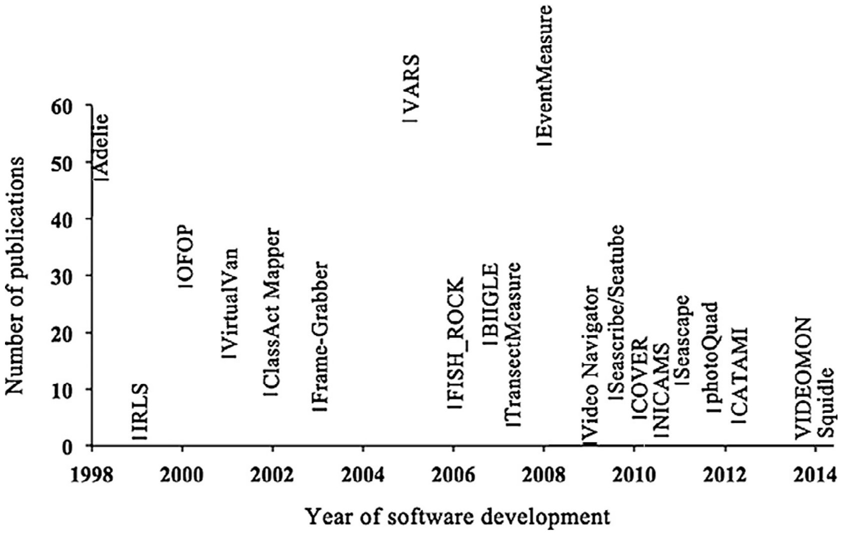


Fig. 1. Total number of publications using each underwater image annotation software system, presented per year of software development between 1998 and 2014 (see Table 1 for software development year); CPCe and ImageJ are not represented as these were cited by more than one hundred publications each; and DIAS was not developed at the time of search.

and (iii) the capacity to interact with a relational database for data management, including archiving, querying and exporting annotation outputs (Fig. 2; Table 2). Most software packages were developed to operate in either or both of real time or posterior to acquisition mode. A schematic of the workflow relating the capture of imagery to the MIAS based on their operating mode appears in Fig. 3. Detailed software specifications (Tables 2 and 3) are discussed comparatively throughout the paper with reference to these two operating modes, whenever applicable.

In the process of analyzing underwater imagery for quantification purposes, two types of data are recorded: (i) the image annotations, and (ii) metadata captured along with the imagery to assist in its use (e.g. time, position, depth/altitude, and other scientific data). For clarification purposes, the former is denoted as 'annotation data' throughout the manuscript, while the latter is denoted as 'metadata'. In the process of ecological or geological analysis, for which most MIAS are used, image annotations are often referred to as 'data'.

Real-time annotation of video from underwater platforms can save a significant amount of time on posterior analysis by indexing the major events observed during a field survey, and contribute to short deployment summaries. An organized and informed dive summary also aids in detailed planning of subsequent surveys, particularly with ROVs, towed-vehicles, or drop-cameras where live video is fed to the surface. MIAS operating in real time include an annotation window and, in some cases, visual displays or panels showing metadata like deployment information and input data from multiple sensors, such as the geographic positions of the ship and surveying platform (e.g. depth, altitude, heading, latitude and longitude, etc.), or environmental data captured concurrent to the imagery (e.g. temperature, salinity, dissolved oxygen, turbidity, current velocity, etc.).

The great majority of video and/or image annotation and analysis is done posterior to the acquisition, hereafter referred to as "posterior annotation". In contrast to real-time mode, it allows a more in-depth analysis, and can be used to detail, improve, validate or re-annotate previous annotations. One benefit of this mode is the capacity to reduce spatial or temporal correlation bias by sub-sampling and randomization of images or video clips prior to annotation. This posterior annotation mode is currently the only option for platforms and equipment set-ups that do not provide live video feed to the surface, for example imagery captured with

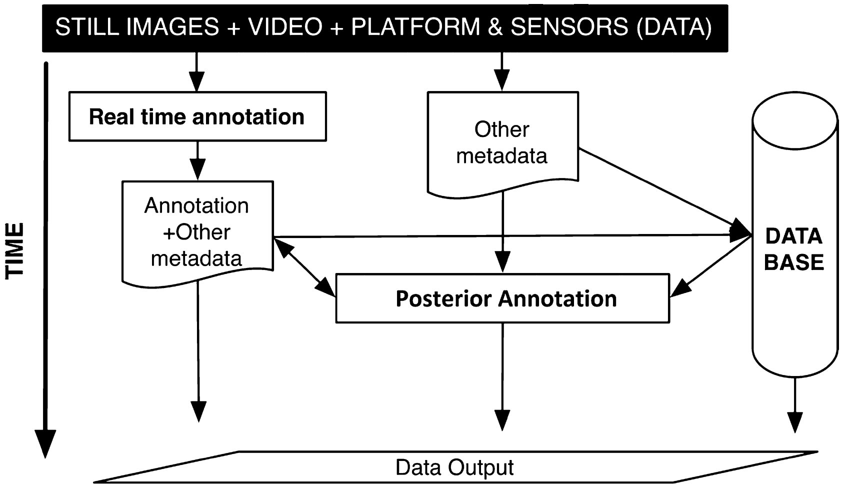


Fig. 2. Simplified data workflow in underwater imagery surveys with image annotation in real time, in posterior annotation mode, and interaction with a database, as well as possible combinations of operation modes (shown below in black bars); ROV, Remotely Operated Vehicle; AUV, Autonomous Underwater Vehicle; BRUV, Baited Remote Underwater Video.

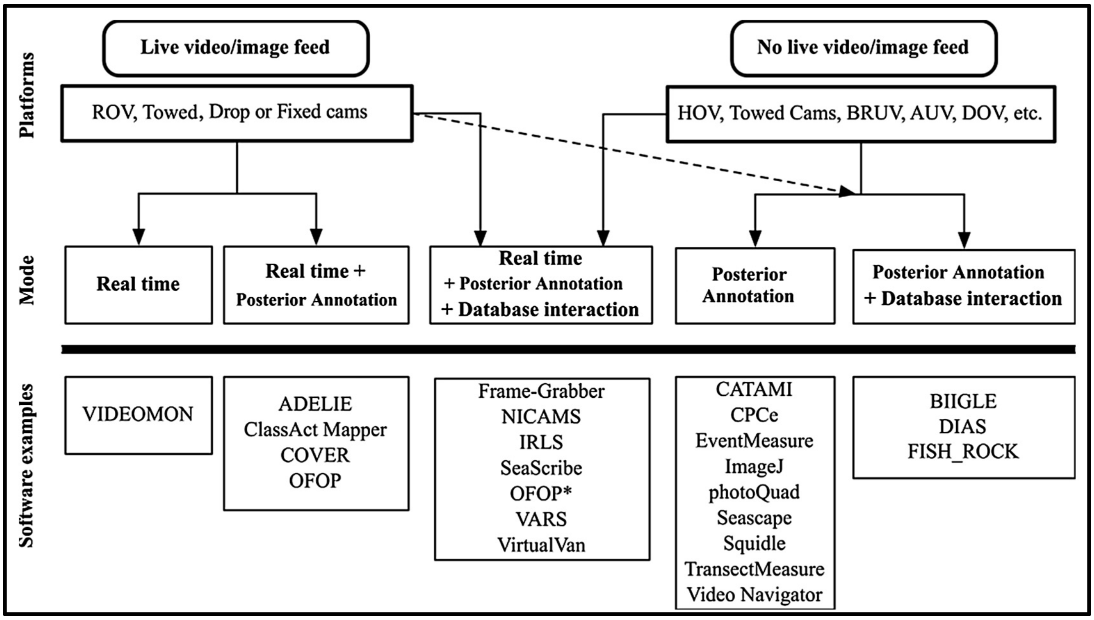


Fig. 3. Survey platforms, operational mode (real time, posterior annotation and database interaction) and examples of underwater video annotation software available (ROV = Remotely Operated Vehicle, HOV = Human Occupied Vehicle, BRUV = Baited Remote Underwater Video, AUV = Autonomous Underwater Vehicle, DOV = Diver Operated Video; analyzing OFOP* indicates that the package can be customized to associate to a database although it is available as a standalone package).

Autonomous Underwater Vehicles (AUVs), Baited Remote Underwater Video (BRUV), and some Human Operated Vehicles (HOVs; Fig. 2). MIAS operating in the posterior annotation mode often include a video player interface or image browser where the user locate a still image or video clip using specific video timestamp or filename information. Similar to MIAS operating in real time mode, other metadata collected *in situ* during video/image collection can often be displayed. In both annotation modes, users will find the annotation window with multiple menus, lists and buttons to create specific annotations related to pre-determined objects or situations (e.g. biological, geological, etc.) or even to add user-specific comments.

Software for annotating in real time normally allow annotating in the posterior annotation mode, with the exception of VIDEOMON, which was design only for real-time operations (Fig. 3; Table 2). However, not all MIAS that operate in both modes can interact directly with a database. The majority of MIAS operate only in the posterior annotation mode and do not interact directly with a database. Other packages such as NIWA Image Capture and Management System (NICAMS), Integrated Real-time Logging System (IRLS) and VirtualVan fall more adequately in the classification of data management system.

2.3. Programming language and operating systems

Most MIAS have been developed in Java (nearly 40%), working as stand-alone applications. Other programming languages include Delphi, Matlab, Visual Basic and C++ (see Table 2). Video Annotation and Reference System (VARS; Schlining and Jacobsen Stout, 2006) is a stand-alone application, but has an online interface to access and query data from the Monterey Bay Aquarium Research Institute's (MBARI) internal database, or it can run on personal database servers (with databases written in PostgreSQL). The online web applications such as BIIGLE (using Flex 3 SDK by Adobe, JavaScript, PHP and, MySQL; Schoening et al., 2009), Data Integration and Annotation Services (DIAS, using JavaScript, PHP; Purser et al., 2009), Squidle and CATAMI (CSS, HTML, JavaScript and Python), run over web browsers, enabling online image annotation in a centralized repository with access to remote databases or upload of data to a remote server. Other MIAS operating over web-browsers are Seatube Pro and VirtualVan (Lerner and Maffei, 2002), where video is already on a server. With the exception of web-based software packages, which run on internet browsers, most software run only in Microsoft Windows environments, although ImageJ and VARS run also in MAC OS X and Linux/UNIX (VARS with some limitations), and Seascape runs in Linux. FISH_ROCK is written in Matlab and should run on different operating systems.

2.4. Availability and costs

MIAS range from open-source to packages that are currently not available outside of the developing institution (e.g. IRLS, SeaScribe, Video Navigator, VIDEOMON, VirtualVan). MIAS can vary widely in costs: commercial licenses range from 450 € to 2000 € (for several years), while several are non-profit packages (Table 2). Web-based solutions include non-profit software and are often based on crowd sourcing (e.g. BIIGLE, DIAS, CATAMI, Squidle), following the Web 2.0 concept of enabling easy data sharing and collaboration, sometimes referred to as Science 2.0 (Nattkemper, 2012; Shneiderman, 2008; Waldrop, 2008). Crowd-sourced annotation is one method for annotating large volumes of imagery. Open-source packages have benefitted from community driven developments (e.g. ImageJ).

Commercial packages include initial support and training and, in some cases, provide client-specific customization (e.g., integrating sensors, output data), increasing the software capabilities and reducing the implementation time. Free and open-source software have no official programming support, often requiring programming skills for installation and maintenance. Some of the software listed are not being further developed, but can be obtained by contacting the developers/developing institution (e.g. CATAMI, COVER, FISH_ROCK).

2.5. Data inputs

Real time annotation is done on video data, but video recording is often independent of the MIAS. A useful feature involves feeding the video stream into the software. Some software allow both annotations to be recorded and the extraction of associated frame grabs that can be archived and post-processed as still images, for example COVER, OFOP, and VARS.

During posterior annotation, MIAS have different capabilities and specifications for data import. A video or still image file is normally recorded and compressed using other software before it is imported. Both resolution and frame size may be reduced to facilitate storage, often at the cost of reduced image quality. GUIs normally rely on widespread freeware video players or image browsers, and depend on third party decoding capabilities to deal with the different video or image compression formats (e.g. COVER).

The facility to receive external data inputs besides the imagery is what distinguishes MIAS from other image annotation software. However, not all MIAS designed to operate in real time can receive external data inputs (e.g. Frame-Grabber, SeaScribe, VARS, VirtualVan; Tables 2 and 3). External metadata are typically transmitted through customizable COM ports (or Ethernet connection) and include parameters, such as ship position (latitude, longitude), platform position (latitude, longitude), depth (m), chemical sensors and other (Table 3). Alternatively, some MIAS log the information from the ship, platform or sensors separately, such as in a database or text file, for example ADELIE (Aide au DEpouiLlement Interactif des données des Engins sous-marins) and VARS. This metadata can be viewed, integrated, or analyzed during the posterior annotation. In many cases, these can be imported just after data collection (e.g. VARS).

Time (timestamp) is often the key indexing field during underwater video surveys allowing combining data from different sensors. The computer where the imagery and annotation logging takes place, as well as for all other deployed sensors, must be synchronized with the ship's clock and all other equipment being deployed. Other identifiers (e.g. textual, numeric) are also used to index imagery, such as file name for still images. Indexing ensures successful integration of data from different platforms and sensors. When time is not registered, time code generators may be used to label video frames, providing a reference for synchronization (e.g., by recording the time code in the video track; Reynolds and Greene, 2008; Sameoto et al., 2008).

Importing metadata (or other annotation data) for display or for compilation with annotations in the output file (e.g. OFOP) is done according to MIAS specifications, but is often input using text files. The most common are the time code (e.g. yyyy-mm-dd hh:mm:ss.000) and camera platform position, but other metadata is often input (e.g. depth, temperature, etc.). Data that have originated in different sensors may require formatting. This is a time-consuming task, and packages are evolving to accept more types of data formats, to allow metadata input during data collection or *a posteriori*, and to interact with databases where these data are already standardized. Importing a previously annotated file including the associated annotations is only available in some packages (see Sections 2.5 and 3.1).

2.6. Annotation interfaces

Most MIAS provide intuitive menus and interfaces that allow annotators without a computer science background to annotate images and/or video. The specifications of annotation processes depend on the survey context, and can include a search/detection task, a classification task and/or a quantification task, and each specification calls for a specific interface design. Annotation interfaces comprise a variable number of buttons (often >10) and lists, appearing as drop-down menus, where customizable lists of descriptors can be included for classification purposes (e.g. lists of taxa, seafloor descriptors). Radio buttons, track bars, or count grids are also available in some software tools, often associated with quantification (Table 3). Some software packages allow additional customization of the interface itself, such as the number and type of buttons, or data type input and display. The difficulty of interface customization varies among systems, with some choices built into the software, while others require programming. A large screen, or the use of two screens, is sometimes recommended, as additional windows will also be open during annotation sessions (e.g. ancillary data, sub-samples images). Interface aesthetics and functionalities is one of the aspects that will benefit from further developments (Hagedorn et al., 2008; Vondrick et al., 2013).

2.7. Image descriptors

MIAS commonly allow a list of descriptors to be added in the annotation window/interface (e.g. uploaded as a text file). Custom-generated lists of descriptors/classifications (e.g., taxonomical) can be changed according to the type of survey and are often based on existing classification schemes.

Reference classification systems offer solutions to organize and interpret data about the marine environment, allowing coherence in the annotation between different experts, promoting consistency and a more rapid selection of commonly used descriptors. Some of these classifications are organized hierarchically and can be used at different spatial scales, spanning from environmental informations concerning the water column, to seafloor types, habitats and biotopes.

Several comprehensive schemes have been developed in recent years. At an European level, the broadest framework is the European Union Nature Information System (EUNIS; Davies et al., 2004). This system features a large coverage of the Northeast Atlantic habitats, extending from the Artic into the Mediterranean and Black Sea. It is currently undergoing a revision that should include major updates to its deep-sea habitat section (Beuck et al., in press; Davies et al., in press).

For North America, the Coastal and Marine Ecological Classification Standard (CMECS; FGDC-STD, 2012) covers marine benthic habitats from coastal to deep waters of the United States, with application to other maritime regions. Another example is the Australian classification scheme for scoring marine biota and substrata in underwater imagery, the Collaborative and Automation Tools for Analysis of Marine Imagery and Video (CATAMI; Althaus et al., 2013). For the classification of marine species the World Register of Marine Species: WoRMS database is generally used (Costello et al., 2013).

2.8. Data display

Real-time operating MIAS can display several types of information, including the imagery, annotations and metadata. In some cases, video is displayed on a different screen (especially with high definition) in order to avoid overloading the image-recording computer (see Section 2.11). The most frequently available ancillary data item is the position (latitude, longitude) of the imagery platform, often with depth and ship position also available in some cases (see GUI examples in Supplementary Material A). The annotation log allows users to delete or correct entries during imagery analysis, either in real time or during posterior annotation (note that only OFOP, VARS and VIDEOMON display the annotation log in real time). Some software packages allow clearance of previous entries (e.g. COVER, VARS). The capacity to overlay metadata on the image file is set up during data collection. It is important to ensure the synchronization of data and logging of events for posterior analysis.

The display of expedition metadata and annotation data allows annotators to have a better understanding of the environment and the "behavior" of the camera platform recording the imagery.

While a camera location can vary considerably, real-time annotation often occurs near the platform control room where all sensors are displayed. Alternatively, video and sensor displays can be streamed to other rooms where annotation takes place.

MIAS operating in the posterior annotation mode display much of the same information as those operating in real time, including the imagery and metadata in some cases (Table 3; examples in Supplementary Material A). During posterior annotations, still images are normally presented as an overview image, with zooming and panning functionality to browse the image at multiple resolutions. The standard image file formats (JPG, PNG) are commonly supported, whereas others (TIF, PPM, RAW) can require additional software libraries or prior conversion in other software. Video display during posterior annotations is typically done through video controls integrated in the GUI, including buttons for pause/play, fast-forward, rewind, and speed control. Display of associated metadata range from systems displaying only one type (e.g. depth or position) to those that display all data collected (e.g. VirtualVan).

Surprisingly, the display of annotations during posterior annotations, such as the annotations taken in real time, is not a common feature (available in Coral Point Counter (CPCe), DIAS, EventMeasure, VARS), with packages having different capabilities in terms of browsing and editing such data (Table 1). This feature has a major relevance for data validation and revisiting datasets and has long been adopted in video annotation software systems that are not dedicated to underwater video (e.g. VCode; Hagedorn et al., 2008). In a recent review, Hagedorn et al. (2008) outlined the lack of visual functionality for reviewing annotations in context with the corresponding video; specifically, a lack of context and precise synchronization to correlate resulting data output files with a given annotation in a video.

2.9. Multiple annotators

The interest of having simultaneous annotations from multiple sources on the same image content (i.e. human domain experts, citizen scientists or algorithmic tools), has driven the development of such functionality in recent years, mostly for MIAS developed for the posterior annotation mode. Collecting annotations from a group of experts for an underwater image collection may facilitate: (i) ground truthing data, (ii) integration of broader taxonomic expertise, and (iii) investigation of the inter- and intra-observer agreement of the users (Howell et al., 2014; Schoening et al., 2012, 2016).

Traditionally, logs for each source are collected separately and compiled *a posteriori*. However, there are MIAS that support annotation via intranet during data acquisition (e.g. IRLS, OFOP, SeaScribe). The need to aggregate expert knowledge from a geographically spread community benefited from the advent of the Web 2.0 concept, where web-based technologies are seen as a major driving force for the collection of user-generated content (Ontrup et al., 2009). Nowadays, IRLS software, in particular, allows several loggers to enter data simultaneously from various stations with a conversation-style interface.

Real-time multiple annotations on board by intranet facilitate the compilation of data logged by different experts onboard into a single file. Multi-user annotations can also occur via the internet, a technologically demanding solution requiring uplink and downlink from the ship/deployment station. The NOAA ship *Okeanos Explorer* (Bell et al., 2016) uses remote communication between scientists onboard and onshore; one person onboard is dedicated to annotating video based on comments made by scientists onshore in near real-time, while video, communications and other data are streamed to shore via the World Wide Web. Ocean Networks Canada has incorporated the underwater telepresence technology on its yearly sea operations while maintaining NEPTUNE and VENUS cabled observatories (http://www.oceannetworks.ca/expeditions). Similarly, scientists ashore and the public can interact with the scientists onboard. Scientists ashore can take part during ROV dive events, either during sampling routines or for carrying on manipulative experiments, communicating with the ship and ROV pilots via a satellite link, and can even annotate videos using the web-based software SeaScribe. Telepresence is ensured by advanced broadband satellite communication. Some of the major challenges include the costs associated to host the uplink and downlink, the amount of metadata being streamed, and maintaining communication with the chief scientist onboard and between annotators (Leslie et al., 2010).

Most MIAS developed for simultaneous multiple-user annotation use still-images (e.g. BIIGLE, CATAMI and Squidle). Only recently, Seatube Pro and DIAS have evolved to allow multiple annotations on video files via intranet or internet. Real-time streaming is becoming more common, particularly for cabled observatories, offshore oil and gas companies and a few large research initiatives.

Finally, crowd sourcing and citizen science, where members of the public assist in large data collection projects, are now common using the internet (e.g., Zooniverse.org). Such programs are expected to increase in the near future to help tackle the ever-increasing amount of image and video being collected.

2.10. Computer-assisted annotation

For over 50 years, research groups worldwide have worked in the areas of automated image and video processing and information retrieval, and recently a slowly growing number of applications being developed specifically for underwater image data have been reported in the mid 2000s (Kämpfe et al., 2004; Smeaton, 2005). The overall research aims of the community are "…achieving machine understanding of video, including all aspects of search, retrieval, visualization and summarization in both contemporaneous and archival content collections." (Hauptmann, 2005). Current research in automated annotation aims to minimize the involvement of the human annotator. Algorithmic approaches are used for automatic image enhancement and interpretation to support manual annotation, or even to automate the entire annotation process. Most automated annotation is focused in operating during posterior annotation, but efforts to implement automated annotation of underwater imagery in real time area underway, a tool of particular use for cabled observatories or any high-volume image capture study (Aguzzi et al., 2012; Spampinato et al., 2008; Strachan, 1993). The recent massive increase in marine image data captured presents a major application of automated annotation, where annotation of large data sets manually would be too time expensive and consistency in annotations difficult to maintain.

Algorithms replicate human visual cognition using automated learning of semantic concept models from a large number of human-annotated image samples, to then detect and identify objects in new images (Zhang et al., 2012). Research continues to work on the difficult problem of understanding images at the necessary level of detail, and is building increasingly sophisticated machine-learning models for the relationship between low-level feature vectors (an n-dimensional vector of numerical features that represent some object) and the content represented in the video or image (Hauptmann, 2005). With sufficient annotation data with which to train the algorithm, increasingly sophisticated machine-learning approaches may converge on the right models needed to understand video or image collections. However, algorithms are limited in several ways: their capacities limit annotation tasks, they depend on human experts to provide sufficient annotation data with which to test the algorithm, and by technical aspects of underwater imaging (e.g. inadequacies in perspective, video resolution, lighting, and other optical properties).

Automated event detection has been applied successfully to videos from the water column and on benthic surveys (Edgington et al., 2006). Several groups are doing research on event detection (e.g. Cline et al., 2007), aiming to reduce observation times. Dynamic events captured in video, from cabled observatories or oil platforms, can also provide cues for analysis, including event-based video indexing, browsing and image segmentation (i.e., partition of a digital image into sets of pixels to simplify the representation into something that is more meaningful and easier to analyze). The detection algorithm is developed by running video frames through an attentional selection algorithm (Itti and Koch, 2000), where objects or events tracked over several video frames are flagged as potentially interesting events. MBARI developed the event-based computer vision system Automated Visual Event Detection for the automated detection of midwater organisms, which mimics human attention while visually browsing images (by combining a saliency-based attentional module and a recognition module; Edgington et al., 2006; Itti et al., 1998). Ocean Networks Canada and partners have also developed algorithms for the automatic detection of events of interest in video imagery from static cameras, such as simple animal motion, benthic species shapes (e.g. squat lobsters), and bacterial mat areal coverage (Aguzzi et al., 2011; Gebali et al., 2012). However, the process of developing algorithms that generically apply to a multitude of applications and situations, and to increasingly large data sets, represents a bottleneck to be overcome.

Automated image analysis protocols have also been created for specific taxa and platforms. The automation of object characterization and their interrelationships in image scenes remains the most challenging research issue (Cox et al., 1998; Zhang et al., 2012). The primary limitations of the automated analysis of benthic images are habitat variations in background (e.g., sediment texture) and water column turbidity. The application of filters for background corrections is a required preliminary step for efficient recognition. Algorithms to classify/identify objects are developed by extracting morphometric indices and analyzing their variations in different species by multivariate statistics (Aguzzi et al., 2011).

In recent years there have been promising results from automated object/species detection and segmentation employing machine-learning algorithms, such as for segmentation/quantification of cold-water corals (Purser et al., 2009) or measuring polymetallic nodule coverage on the seabed (Schoening et al., 2015b). In case of the detection, classification and quantification of species, computational assessment of shrimp distribution (Osterloff et al., 2016) and the megafauna (Schoening et al., 2012) have been considered. The aim of this development is an increase in the speed of image analysis and with similar identification accuracy to that of human annotators (Purser et al., 2009; Schoening et al., 2012). In the future it may be possible to implement some specific automated annotation procedures in real time as accuracy and processing speeds increase (Schoening et al., 2015b).

Research on automated fish species classification is still a growing theme, with current settings requiring a static video recording platform. Traditionally such tasks take place in controlled environments, such as dams or aquaculture facilities, where underwater video provides a useful alternative to other methods (that require fish to swim through constrained passages), such as acoustics, resistivity counters or infrared beams (Morais et al., 2005). Tracking objects in video can be implemented by using different image filtering protocols, such as those for pixel size, grey-levels or red/green/blue enhancement (Costa et al., 2009). There are successful examples using global appearance shape descriptors (e.g. Mokhtarian and Abbasi, 2002) and combining texture features with shape descriptors (Spampinato et al., 2010), although these are often designed for only a small number of species. Shortis et al. (2013) presents a synthesis of the current approaches to the recognition and measurement of fish in video sequences and the analysis of the methodologies.

2.11. Quantification of image features

Annotation may include the quantification of features, with image/object area and length the most frequently calculated. Some MIAS facilitate measurements through on-screen measurement tools, data recording, and conversion of pixel measurements to 'real' units (e.g. mm). Size measurements require knowledge of the scale (e.g. size of a known object, a scaling laser system or altimetry and position data). Quantification requires area knowledge and is easier with images taken perpendicular to an assumed flat seafloor; the camera footprint can be estimated using camera lens specifications together with pan, tilt, zoom and altitude, that can be exported to the annotation log. Solutions to deal with distortion and parallax have been discussed by Morris et al. (2014). Area estimates can also be obtained in oblique views, by applying a perspective grid (Wakefield and Genin, 1987), through the use of five laser diodes and trigonometry solving software such as Benthic Imager (Pilgrim et al., 2000) or VARS (Schlining and Jacobsen Stout, 2006), or from parallel laser lines (Dias et al., 2015). MIAS developed for stereo cameras can also allow measuring lengths and other 3D information (e.g. Event Measure; see also Dunlop et al., 2015). Automated computation of the spatial pixel resolution for each image, based on the laser point positions (Pilgrim et al., 2000; Pinkard et al., 2005) and the footprint sizes, has been successfully developed in BIIGLE software (Schoening et al., 2015a).

Among the different techniques for area analysis, point count analysis of photo quadrats has been used for several decades (Pante and Dustan, 2012) as an efficient way to estimate percent cover and monitor large areas. Available tools include fixed-point counts (e.g. FISH_ROCK, TransectMeasure), random point counts (or both, e.g. COVER), and planar area analysis (e.g. BIIGLE, CPCe, DIAS, photoQuad, TransectMeasure). Segmentation tools based on hierarchical segmentation algorithms and extraction of advanced 2D morphometric descriptors (e.g. areal cover and perimeter estimation) have also been developed to assist manual annotation (e.g. Seascape, photoQuad). A review of the methods available for estimating abundance and percent cover from underwater still images and video has been provided by several authors (Jamieson et al., 2013; Jones et al., 2009; Tran, 2013).

2.12. Output files

Annotation outputs are normally recorded in comma separated value (CSV) files, text files or, as database-specific files. Other common outputs are web or web-GIS enabled formats (e.g. HTML, CSV, KMZ; Table 2). Using the time code or image name as the key indexing fields, the output files are mostly differentiated by (i) having several descriptors in the same output line (referring to the same time or image name; e.g. CATAMI, COVER, OFOP, VARS), or (ii) one line per descriptor, with image name/time code repeated where necessary. Some MIAS integrate hierarchical classification, for example for taxonomic purposes (e.g. FISH_ROCK, VARS, EventMeasure, TransectMeasure).

2.13. Additional tools

2.13.1. Capture of still images from video

Still images (frame grabs) can be captured during video annotation. Such images captured in real time can be used to rapidly illustrate occurrences of interest during the dive, for illustrating the dive report or preparing subsequent surveys. Only a small number of software packages have this feature available (e.g. Frame-Grabber, OFOP, VirtualVan, IRLS). Images captured in posterior annotation may be used to illustrate annotated objects on video, and in object catalogues or databases.

A small number of MIAS include a tool to automate the capture of still images from video (e.g. Frame-Grabber, Cover, Jason Virtual Van, OFOP, TransectMeasure). These images may be captured in real time at specific time intervals, for example to illustrate platform activities. Still images can also be recorded in real time by manual activation, or with annotation (e.g. Frame-Grabber and VirtualVan). In posterior annotation, images may be extracted at a fixed time interval or by distance intervals (accounting for platform position; e.g., OFOP). This feature facilitates the subsampling of large video surveys. Many of these tools offer a choice of image file type generated.

2.13.1.1. Randomization of imagery for annotation. Some MIAS allow the randomization of imagery to be annotated (e.g. VARS), using a random number generator to select the following image or video clip to be annotated. This randomization aids in reducing possible bias in the analysis (Howell et al., 2014).

2.13.1.2. Combine image annotations with other data. Scientists commonly combine annotation data with data from other sensors during data analysis (e.g. ROV position, self logging CTDs or other physical/chemical sensors). Some MIAS include tools to aggregate all data in one output file (e.g. Cover, OFOP). OFOP can further process position data and spline this position data to the second, pairing it with the video data with a similar time code.

2.13.1.3. Preliminary annotation data analysis. Some MIAS provide tools for simple data analysis (Hagedorn et al., 2008). A few examples include basic statistic outputs by EventMeasure, and basic calculations for BRUVS (e.g. MaxN, time of arrival, and 3D measurements). CPCe provides basic statistics and Shannon–Weaver diversity index (Magurran, 2013). For example, tools in BIIGLE plot annotation data as histograms or scatterplots. Standardized and customizable database queries can also provide a preliminary analysis of datasets (e.g. VARS). Advanced analysis of output files is generally not included in MIAS, with the exception of Observation, an ArcGIS extension of the ADELIE suite, that dynamically calculates the video swath and generates population density matrices.

2.13.1.4. Scale information and size measurements. Some tools support absolute or relative measurement of sizes in cm by enabling computing the pixel/cm relation based on reference information like laser points. This is helpful for distance and biomass estimation or the assessment of other masses (condition to the availability of an appropriate size/mass formula (Ontrup et al., 2009; Schoening et al., 2015a).

3. Databases

In every area of science, much investigation now depends not on new experiments, but on databases in which observational evidence has been stored (Buneman et al., 2005). Imagery annotation data may be viewed as a temporal phenomenon: a body of related descriptors or other text that may change with time, as new annotations are added and existing ones are updated.

A number of MIAS interact with a database (see Table 2), either as an integral part of the MIAS (e.g. BIIGLE, VARS) or with some customization freedom (e.g. OFOP). Databases often include a reference to the source imagery file and some interact with other meta-databases such as the Ocean Biogeographic Information System (Grassle, 2000) or Global Biodiversity Information Facility (Edwards et al., 2000). Examples of databases, systems architecture and data workflows are presented in Supplementary Material A (A.2, A.6, A.18, A.20, A.26 and A.29).

3.1. Browsing and editing

The capability of browsing and editing annotation data with the GUI can prove very useful (Table 3). Some MIAS allow browsing through imagery stored on a personal computer or an online database system (e.g. BIIGLE, DIAS, CATAMI, Squidle, VARS), including browsing and query of previous annotations (as retrieved by keywords for instance). Browsing allows annotation data to be validated by trained experts and for quality control purposes. However, in a large number of MIAS, browsing is limited to video replay or viewing still images.

Some MIAS, like EventMeasure, allow the annotation file to be queried and the associated imagery is then shown, where the video file is treated as constituent frames that can be browsed. Similarly, once a list of annotations is uploaded to OFOP, the video pauses automatically at each annotation. Browsing in MIAS that incorporate a database is done through query of the database (e.g. VARS) and in some cases through a different application (e.g. ADELIE, BIIGLE, see Supplementary Material A). In DIAS a “single patch classification view” provides a display of existing annotations in a rapid serial representation to the annotators. Posterior annotation in several packages is developed by uploading a dataset and either overwriting or appending the previous annotations, or by creating a brand new file. Only certain systems allow reviewing and editing annotations (e.g. BIIGLE, DIAS, CPCe, ImageJ, VARS, photoQuad). The capability of browsing and edit annotation data allows the user to consult/validate previous annotations which can be a key issue on revisiting databases. In general, current MIAS lack support for efficient annotation workflow, provide poor representation of data on a timeline (Hagedorn et al., 2008).

4. Selecting a suitable MIAS

No single MIAS meets all needs, but one may be more suited for the specific survey and data product needs to fulfill the scientific aims of a project. In choosing a particular MIAS, it is important to consider the research questions and deployment systems, and timeline for data use. In certain situations an image annotation software might be unnecessary, such as for simple annotations over very small datasets, poor access to computing, or where there is no need for long-term databasing.

The main aspects to consider are the operation mode (whether in real time or posterior to the acquisition), the suitability to the annotation protocol to be applied (Schoening et al., 2016) and the costs and effort for implementation. Additional tools might be considered, as they relate to specific annotation requirements. In the case of implementing a MIAS in a research center, the capacity of interaction with a database (e.g. integration with existing database, or to enable management of data from multiple annotators or projects), and maintenance/updating of the MIAS, should be particularly considered. Different annotation scenarios and requirements should be established based on the short, mid and long-term objectives. If the scientific questions change, the data workflow may change, and the capacity of the MIAS to adapt or be customized may be valuable.

4.1. Operation mode and platform type

The platform type(s) determine the annotation environment (Figs. 2 and 3). An important question is if real time annotation is needed? In real time scenarios, is the aim solely to annotate or to have a data management solution (including data inputs, data display), or even to provide feedback to assist in platform deployment? In posterior annotation, will video and/or image be annotated?

The use of video or still images, depends on (i) the research platform, (ii) the imagery collected (camera angle, definition), (iii) the type of analysis for (iv) addressing a specific question, a decision process to be evaluated in an inverted order if a scientific approach is to be adopted. As the questions change, the data workflow is expected to change, and the capacity of particular software to manage different equipment sets and workflows, as well as post hoc analysis, relies on the variety of tools and/or capacity for customization.

4.2. Costs and effort

MIAS license costs (where applicable), often relate to the complexity of the package (e.g. functionality or tools), flexibility for customization, implementation effort, and customer support, training and user documentation. The need for functionalities such as data display, re-annotation of imagery via browsing and editing annotations, and additional tools that can aid in minimizing analysis time, should also be taken in account. While a practical and simple solution can be the most adequate, more complete MIAS reduce the need for third party software and additional data analysis techniques a video annotator needs to be acquainted with.

4.3. Database interaction

The MIAS interaction with a database can be very useful for long-term establishment of an imagery-based research group or facility, and substitution or interaction with the existing system components should be considered. Additionally, large integrative projects (e.g. Joint Programming Initiative, Healthy and Productive Seas and Oceans; Rasenberg, 2011) and international directives (INSPIRE, 2007) have operational procedures that foresee the sharing of data products by the different institutions.

4.4. The ideal MIAS

From an image annotator/scientist perspective, the ideal MIAS is one that performs the functions needed (on the specific operation mode, dataset and research question), as simply as possible, with customizable functions as necessary, free of charge, with easy set-up and maintenance. For implementation in research centers, the ideal MIAS would be a more complete package, including all features of real-time and posterior annotation packages, and be easily configured to interact with any database format. It would be versatile, expandable, stable, easily and universally deployable software.

The most versatile MIAS would allow for real-time input of vessel data, deployed sensor data, video, still images and annotations. The annotation interfaces would be customizable, and allow for multiple annotations via intranet and internet. Capture of still images from video could be done manually, and automatic detection and recording of events or objects of interest in still images/video would be integrated. It could continuously estimate the field of view (based on sensor data) and export all data to a database in real time. It would also, ideally, produce customized activity reports and enable preparation of surveys. To enable customization, it would consist of an application program interface and a scripting language for additional programming. Open-source MIAS with tools developed by the wider community would offer opportunities for troubleshooting, support and future improvement. Considering the tools available such software could be capable of assisting with camera platform deployment and platform navigation.

In posterior annotation, such ultra-versatile MIAS would allow previously annotated files to be browsed and edited, and relevant sensor data be uploaded. It would include video and still image analysis tools (e.g. frame extraction by distance or time), tools to quantify features (e.g. area analysis, segmentation tools, etc.), and tools for basic analysis of the resulting image annotation data (frequency, abundance, ecological indexes, etc.). Ultimately, it would have computer-assisted and automated annotation tools to process imagery and highlight events of interest. Such MIAS would provide a means for browsing and querying databases of annotations and metadata. Lastly, an equivalent web-interface should be available, to allow multiple users to access contents and perform annotation remotely.

5. Future trends

Underwater image annotation is a 21st century science, with an ever-increasing number of MIAS being developed. Dozens of institutes are using or developing MIAS, while MIAS are being used commercially to meet the monitoring obligations of offshore oil and gas industries and future deep-sea mining activities. Future developments will improve MIAS and the quality and volume of annotation data. Some anticipated developments are listed below:

An increasing number of MIAS will become available outside developing institutions, or be released as commercial products.
More MIAS will become freely available and open-source, allowing for community-driven development.
A greater variety of system set-ups and data workflows will be available.
The capacity for receiving and displaying platform and sensors data in real time will be improved. Tools are expected to develop to assist in survey preparation and deployments, functioning as versatile data and mission management systems.
Improvements to the functionality for visual, graphical and contextual review of annotations are expected.
Improved tools for in-depth processing and data analysis will accelerate the publication of data products (e.g. spatial analysis, ecological indices, etc.).
Cross-validation techniques are expected to provide improved assessments of intra and inter-observer agreement, allowing attributing bias in relation to particular datasets.
Specialized MIAS tools will address specific research needs and methods (e.g. image filtering for photomosaic).
Integration of computer-assisted tools for event detection and species identification will become increasingly available.
As catalogues (e.g. image-based taxonomic identification keys) are increasingly available, they will be integrated into MIAS to assist both human annotation and the development of machine-learning algorithms.
Standardized classification schemes in online databases (e.g. WORMS) will be integrated into MIAS.
In-house and online database interaction will improve, allowing increased data sharing.
Web-based technologies will become a major driving force for the collection of content (e.g. Serpent Project; Hudson et al., 2005), interacting with MIAS to produce user-generated content.
Live web-streams and real-time, multiple annotation will become more frequent and economically feasible, and involve the public and international community.

Acknowledgments

Information in this manuscript was presented by the first author at the Marine Imaging Workshop, in 7–10 April 2014, at the National Oceanographic Center Southampton, United Kingdom. It was further developed following some of the conference presentations and results. JNGP was supported by the doctoral grant from the Regional Directorate for Education, Science and Culture, of the Regional Government of the Azores (M3.1.2/F/062/2011). GD was supported by the doctoral grant from the Portuguese Science Foundation (SFRH/BD/73205/2010). MB is supported by the Helmholtz Alliance ROBEX. TS is supported by the German BMWi grant 03SX344A TIMM-IS2U. LARSyS-Associated Laboratory no. 9 is funded by the Portuguese Foundation for Science and Technology (FCT) through pluriannual and programmatic funding schemes (OE, FEDER, POCI2001, FSE) and by the Azores Directorate for Science and Technology (DRCT). This is publication number 35856 of The Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.pocean.2016.07.005.

References

Aguzzi, J., Company, J.B., Costa, C., Matabos, M., Azzurro, E., Mànuel, A., Menesatti, P., Sardà, F., Canals, M., Delory, E., 2012. Challenges to the assessment of benthic populations and biodiversity as a result of rhythmic behaviour: video solutions from cabled observatories. Oceanogr. Mar. Biol. – Annu. Rev. 50, 235.

Aguzzi, J., Costa, C., Robert, K., Matabos, M., Antonucci, F., Juniper, S.K., Menesatti, P., 2011. Automated image analysis for the detection of benthic crustaceans and bacterial mat coverage using the VENUS undersea cabled network. Sensors 11, 10534–10556.

Althaus, F., Hill, N., Edwards, L., Ferrari, R., 2013. CATAMI Classification Scheme for Scoring Marine Biota and Substrata in Underwater Imagery – A Pictorial Guide to the Collaborative and Automation Tools for Analysis of Marine Imagery and Video (CATAMI) Classification Scheme. CATAMI.

Bell, K.L.C., Brennan, M.L., Flanders, J., Raineault, N.A., Wagner, K. (Eds.), 2016. New Frontiers in Ocean Exploration: The E/V Nautilus and NOAA Ship Okeanos Explorer, 2015 Field Season.

Benoit-Bird, K.J., Au, W.W., 2006. Extreme diel horizontal migrations by a tropical nearshore resident micronekton community. Mar. Ecol. Prog. Ser. 319, 1.

Beuck, L., Vertino, A., Guillaumont, B., Savini, A., Laffargue, P., Tempera, F., Van den Beld, I., Kutti, T., Freiwald, A., Grehan, A., et al., in press. Protocol for standardised underwater video analysis of European cold-water coral habitats with special emphasis on coral-fish interactions. Deep-Sea Res. II (in press).

Buneman, P., Bose, R., Ecklund, D., 2005. Annotation in Scientific Data: A Scoping Report. Technical Report. University of Edinburgh.

Burr, B., 2006. VACA: a tool for qualitative video analysis. In: Proceedings of the CHI '06, Extended Abstracts on Human Factors in Computing Systems, 22–27 April 2006, Montreal, Canada, pp. 622–627.

Cline, D.E., Edgington, D.R., Mariette, J., 2007. An automated visual event detection system for cabled observatory video. In: MTS/IEEE Oceans 2007 Conference Proceedings, Vancouver, Canada September 2007. IEEE Press.

Costa, C., Pallottino, F., Angelini, C., Proietti, M., Capoccioni, F., Aguzzi, J., Antonucci, F., Menesatti, P., 2009. Colour calibration for quantitative biological analysis: a novel automated multivariate approach. Instrument. Viewpoint, 70–71.

Costello, M.J., Bouchet, P., Boxshall, G., Fauchald, K., Gordon, D., Hoeksema, B.W., Poore, G.C., van Soest, R.W., Stöhr, S., Walter, T.C., Vanhoorne, B., 2013. Global coordination and standardisation in marine biodiversity through the World Register of Marine Species (WoRMS) and related databases. PLoS ONE 8 (1), e51629.

Cox, R.V., Haskell, B.G., Lecun, Y., Shahraray, B., Rabiner, L., 1998. On the applications of multimedia processing to communications. Proc. IEEE 86, 755–824.

Datta, R., Li, J., Wang, J.Z., 2005. Content-based image retrieval: approaches and trends of the new age. In: Proceedings of the 7th ACM SIGMM International Workshop on Multimedia Information Retrieval. ACM, pp. 253–262.

Davies, J.S., Guillaumont, B., Tempera, F., Vertino, A., Beuck, L., Ólafsdóttir, S.H., Smith, C., Fossa, J.H., van den Beld, I., Savini, A., Rengstorf, A., Bayle, B., Bourillet, J.-F., Arnaud-Haond, S., Grehan, A., in press. A cold-water coral biota classification scheme for ecosystem-based management of the deep sea. Deep-Sea Res. II (in press).

Davies, C.E., Moss, D., Hill, M.O., 2004. EUNIS Habitat Classification Revised 2004. Report to: European Environment Agency-European Topic Centre on Nature Protection and Biodiversity, pp. 127–143.

Dias, F., Gomes-Pereira, J., Tojeira, I., Souto, M., Afonso, A., Calado, A., Madureira, P., Campos, A.S., 2015. Area estimation of deep-sea surfaces from oblique still images. PLoS ONE 10 (7).

Dunlop, K.M., Kuhnz, L.A., Ruhl, H.A., Huffard, C.L., Caress, D.W., Henthorn, R.G., Hobson, B.W., McGill, P., Smith, K.L., 2015. An evaluation of deep-sea benthic megafauna length measurements obtained with laser and stereo camera methods. Deep Sea Res. Part I 96, 38–48.

Edgington, D.R., Cline, D.E., Davis, D., Kerkez, I., Mariette, J., 2006. Detecting, tracking and classifying animals in underwater video. In: Oceans 2006. IEEE, pp. 1–5.

Edwards, J.L., Lane, M.A., Nielsen, E.S., 2000. Interoperability of biodiversity databases: biodiversity information on every desktop. Science 289, 2312–2314.

Federal Geographic Data Committee (FGDC). 2012. FGDC-STD-018-2012. Coastal and marine ecological classification standard. Reston (VA): Federal Geographic Data Committee.

Gebali, A., Branzan-Albu, A., Hoeberechts, M., 2014. Detection of salient events in large datasets of underwater video. In: IEEE Oceans. IEEE. 2012, Oct 14, pp. 1–10.

Grassle, J., 2000. The Ocean Biogeographic Information System (OBIS): an on-line, worldwide atlas for accessing, modeling and mapping marine biological data in a multidimensional geographic context. Oceanography 13, 5–7.

Hagedorn, J., Hailpern, J., Karahalios, K.G., 2008. VCode and VData: illustrating a new framework for supporting the video annotation workflow. In: Proceedings of the Working Conference on Advanced Visual Interfaces. ACM Press, Napoli, Italy, pp. 317–321.

Hauptmann, A.G., 2005. Lessons for the future from a decade of informedia video analysis research. In: Proceedings of the International Conference on Image and Video Retrieval. Springer, Berlin, Heidelberg.

Howell, K.L., Bullimore, R.D., Foster, N.L., 2014. Quality assurance in the identification of deep-sea taxa from video and image analysis: response to Henry and Roberts. ICES J. Mar. Sci.: J. Conseil 71, 899–906.

Huffard, C.L., Kuhnz, L.A., Lemon, L., Sherman, A.D., Smith Jr., K.L., 2016. Demographic indicators of change in a deposit- feeding abyssal holothurian community (Station M, 4000m). Deep-Sea Res. I: Oceanogr. Res. Papers 109, 27–39.

Hudson, I., Jones, D., Wigham, D., 2005. A review of the uses of work-class ROVs for the benefits of science: lessons learned from the SERPENT project. Underwater Technol. 26, 83–88.

INSPIRE, E., 2007. Directive: Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 Establishing an Infrastructure for Spatial Information in the European Community (INSPIRE). Off. J. Eur. Union 108, 50.

Itti, L., Koch, C., 2000. In: NATO (Ed.), Target Detection Using Saliency-Based Attention. Research and Technology Organization. Defense Technical Information Center Document, pp. 1–3.

Itti, L., Koch, C., Niebur, E., 1998. A model of saliency-based visual attention for rapid scene analysis. IEEE Trans. Pattern Anal. Mach. Intell. 20, 1254–1259.

Jaffe, J.S., Moore, K.D., McLean, J., Strand, M.P., 2001. Underwater optical imaging: status and prospects. Oceanography 14, 66–76.

Jamieson, A., Boorman, B., Jones, D.O.B., 2013. Deep-sea benthic sampling. In: Eleftheriou, A. (Ed.), Methods for the Study of Marine Benthos. John Wiley & Sons, Ltd., Chichester, pp. 285–348.

Jones, D., Bett, B., Wynn, R., Masson, D., 2009. The use of towed camera platforms in deep-water science. Underwater Technol. 28, 41–50.

Juniper, S.K., Garrett, J., Shepherd, K., Tamburri, K., Wallace, K., 2000. IRL: an interactive real-time, logging system for ROVs. In: OCEANS 2000 MTS/IEEE Conference and Exhibition. IEEE, pp. 465–473.

Kämpfe, T., Nattkemper, T.W., Ritter, H., 2004. AQUISAR: image retrieval in underwater webcam images. In: Proceedings of 5th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Lisboa, Portugal.

Kipp, M., 2001. ANVIL – a generic annotation tool for multimodal dialogue. In: 7th European Conference on Speech Communication and Technology (Eurospeech 2001), Aalborg, Denmark.

Kitayama, T., Tanaka, K., Shimabukuro, R., Hase, H., Ogido, M., Nakamura, M., Saito, H., Hanafusa, Y., Sonoda, A., 2012. Efficient management and promotion of utilization of the video information acquired by observation. In: AGU 2012 Fall Meeting Abstract San Francisco, California, 3–7 December 2012.

Lerner, S., Maffei, A., 2002. In: Institution, W.H.O. (Ed.), The Jason II Virtual Control Van System, Data Acquisition System, Web-Based Event Logger, and SeaNet. WHOI Technical Report.

Leslie, M., Scott, N., Guillemot, E., Auger, V., 2010. Video acquisition, archiving, annotation and analysis: NEPTUNE Canada's real-time georeferenced library of deep sea video. In: OCEANS 2010. IEEE, pp. 1–9.

Magurran, A.E., 2013. Measuring Biological Diversity. John Wiley & Sons, p. 179 pp.

Mokhtarian, F., Abbasi, S., 2002. Shape similarity retrieval under affine transforms. Pattern Recogn. 35, 31–41.

Morais, E.F., Campos, M.F.M., Padua, F.L., Carceroni, R.L., 2005. Particle filter-based predictive tracking for robust fish counting. In: Proceedings of the 18th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI'05). IEEE, pp. 367–374.

Morris, K.J., Bett, B.J., Durden, J.M., Huvenne, V.A., Milligan, R., Jones, D.O., McPhail, S., Robert, K., Bailey, D.M., Ruhl, H.A., 2014. A new method for ecological surveying of the abyss using autonomous underwater vehicle photography. Limnol. Oceanogr.: Methods 12, 795–809.

Nattkemper, T., 2012. Are we ready for Science 2.0? In: Proceedings of 4th International Conference on Knowledge Management and Information Sharing (KMIS), 2012, Barcelona, Spain.

Ontrup, J., Ehnert, N., Bergmann, M., Nattkemper, T.W., 2009. BIIGLE-Web 2.0 enabled labelling and exploring of images from the Arctic deep-sea observatory HAUSGARTEN. In: Proceedings of the IEEE OCEANS Conference. IEEE, Bremen, Germany, pp. 1–7.

Osterloff, J., Nilssen, I., Nattkemper, T.W., 2016. A computer vision approach for monitoring the spatial and temporal shrimp distribution at the LoVe observatory. Methods Oceanogr. 15–16, 114–128.

Pante, E., Dustan, P., 2012. Getting to the point: accuracy of point count in monitoring ecosystem change. J. Mar. Biol. 2012,. 7 pages Article ID 802875.

Pilgrim, D., Parry, D., Jones, M., Kendall, M., 2000. ROV image scaling with laser spot patterns. Underwater Technol.: Int. J. Soc. Underwater 24, 93–103.

Pinkard, D.R., Kocak, D.M., Butler, J.L., 2005. Use of a video and laser system to quantify transect area for remotely operated vehicle (ROV) rockfish and abalone surveys. In: OCEANS, 2005. Proceedings of MTS/IEEE. IEEE, pp. 2824–2829.

Porteiro, F.M., Gomes-Pereira, J.N., Pham, C.K., Tempera, F., Santos, R.S., 2013. Distribution and habitat association of benthic fish on the Condor seamount (NE Atlantic, Azores) from in situ observations. Deep Sea Res. Part II 98A, 114–128.

Purser, A., Bergmann, M., Lundalv, T., Ontrup, J., Nattkemper, T.W., 2009. Use of machine-learning algorithms for the automated detection of cold-water coral habitats: a pilot study. Mar. Ecol. Prog. Ser. 397, 241–251.

Rasenberg, M.M.M., 2011. Joint Program Initiative Healthy and Productive Seas and Oceans. Institute for Marine Resources & Ecosystem Studies. Report C107/11. October 2011.

Reynolds, J.R., Greene, H.G. (Eds.), 2008. Marine Habitat Mapping Technology for Alaska. Alaska Sea Grant for North Pacific Research Board. 282 pp.

Rosenblum, K.L., Zeanah, C., McDonough, S., Muzik, M., 2004. Video-taped coding of working model of the child interviews: a viable and useful alternative to verbatim transcripts? Infant Behav. Dev. 27, 544–549.

Sameoto, J.A., Lawton, P., Strong, M.B., 2008. An Approach to the Development of a Relational Database and GIS Applicable Scheme for the Analysis of Video-Based Surveys of Benthic Habitats. Canadian Technical Report of Fisheries and Aquatic Sciences 2818: iv, 34 p.

Schlining, B., Jacobsen Stout, N., 2006. MBARI's video annotation and reference system. In: Oceans Conference, Boston, Massachusetts. 1–5 p.

Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nat. Methods 9, 671–675.

Schoening, T., Ehnert, N., Ontrup, J.R., Nattkemper, T.W., 2009. BIIGLE Tools – a Web 2.0 approach for visual Bioimage database mining. In: Information Visualisation, 2009 13th International Conference. IEEE, pp. 51–56.

Schoening, T., Bergmann, M., Ontrup, J.r., Taylor, J., Dannheim, J., Gutt, J., Purser, A., Nattkemper, T.W., 2012. Semi-automated image analysis for the assessment of megafaunal densities at the Arctic deep-sea observatory HAUSGARTEN. PLoS ONE 7, e38179.

Schoening, T., Kuhn, T., Bergmann, M., Nattkemper, T.W., 2015a. DELPHI – fast and adaptive computational laser point detection and visual footprint quantification for arbitrary underwater image collections. Front. Mar. Sci. http://dx.doi.org/10.3389/fmars.2015.00020 2015.

Schoening, T., Langenkämper, D., Steinbrink, B., Brün, D., Nattkemper, T.W., 2015b. Rapid image processing and classification in underwater exploration using advanced high performance computing. In: IEEE Oceans, Washington, DC, USA.

Schoening, T., Osterloff, J., Nattkemper, T.W., 2016. RecoMIA-Recommendations for Marine Image Annotation: lessons learned and future directions. Front. Mar. Sci. 3, 59.

Shneiderman, B., 2008. Science 2.0. Science 319, 1349.

Shortis, M.R., Ravanbakskh, M., Shaifat, F., Harvey, E.S., Mian, A., Seager, J.W., Culverhouse, P.F., Cline, D.E., Edgington, D.R., 2013. A review of techniques for the identification and measurement of fish in underwater stereo-video image sequences, SPIE Optical Metrology 2013. Int. Soc. Opt. Photon., 87910G

Smeaton, A.F., 2005. Large Scale Evaluations of Multimedia Information Retrieval: The TRECVid Experience, Image and Video Retrieval. Springer, pp. 11–17.

Solan, M., Germano, J.D., Rhoads, D.C., Smith, C., Michaud, E., Parry, D., Wenzhöfer, F., Kennedy, B., Henriques, C., Battle, E., 2003. Towards a greater understanding of pattern, scale and process in marine benthic systems: a picture is worth a thousand worms. J. Exp. Mar. Biol. Ecol. 285, 313–338.

Spampinato, C., Chen-Burger, Y.-H., Nadarajan, G., Fisher, R.B., 2008. Detecting, tracking and counting fish in low quality unconstrained underwater videos. Visapp 2008 (2), 514–519.

Spampinato, C., Giordano, D., Di Salvo, R., Chen-Burger, Y.-H.J., Fisher, R.B., Nadarajan, G., 2010. Automatic fish classification for underwater species behavior understanding. In: Proceedings of the First ACM International Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams. ACM, pp. 45–50.

Strachan, N.J.C., 1993. Recognition of fish species by colour and shape. Image Vis. Comput. 11, 2–10.

Tran, M., 2013. Mapping and Predicting Benthic Habitats in Estuaries Using Towed Underwater Video. University of Technology. Sydney School of the Environment, Sidney, p. 156.

Vevers, H.G., 1951. Photography of the sea floor. J. Mar. Biol. Assoc. United Kingdom 30. 101:111.7.

Vondrick, C., Patterson, D., Ramanan, D., 2013. Efficiently Scaling up Crowdsourced Video Annotation. Int. J. Comput. Vis. 101, 184–204.

Waldrop, M., 2008. Science 2.0: Great New Tool, or Great Risk? Wikis, blogs and other collaborative web technologies could usher in a new era of science Or not. Sci. Am.

Wakefield, W.W., Genin, A., 1987. The use of a Canadian (perspective) grid in deep-sea photography. Deep Sea Res. A: Oceanogr. Res. Papers 34 (3), 469–478.

Zhang, D., Islam, M.M., Lu, G., 2012. A review on automatic image annotation techniques. Pattern Recogn. 45 (1), 346–362.