Video Annotations Quality Descriptors

Mashkoor Malik, Brian Kennedy
NOAA OAR, Office of Ocean Exploration and Research

Megan Cromwell
NOAA NESDIS, National Centers for Environmental Information

Jessica Robinson
Ocean Networks Canada

NOAA Environment Data Management Workshop
Silver Spring, MD
April 23, 2018
Outline

• Use of images and video in NOAA
• Need for video/Imagery metadata
  • Annotations provide description of video observations
• How good the annotations are?
  • Has every frame needing an annotation has an annotation?
  • Are annotations accurate?
• Implementation of image processing
• Recommendations
NOAA/Video/Imagery data challenge

• No estimates for total NOAA optical data volume
• NMFS alone estimates 902 TB of optical data
• Use and reuse of video/images challenging
• Extremely labor intensive analysis
  – Rough calculations:
  – 50-150 sea days / year
  – 8-10 hours/day
  – 10 – 20 platforms
  – 4000 – 30,000 hrs/year (Media time)
  – 500 – 3750 work days/year (Discovery/Analysis time)
Metadata is critically needed for imagery use

• Metadata
  – ‘a set of data that describes and gives information about other data’

• Two types of imagery metadata (Jain & Hampapur, 1994)
  – Operational metadata: equipment, software, date, spatial coordinates etc. *(ISO 19115-2:2009 Part 2: Extensions for imagery and gridded data)*
  – Subject metadata: Automated or human interpreted. Annotations to enable content based search *(No standard exists)*
Subject metadata generation challenges

• Tools to effectively collect subject metadata

• Assess quality of metadata
  – Quantity of annotations
  – Correctness of annotations

• Standardized vocabulary
OER video workflow
Annotations play a key role!!

- ROV Video
- Sensor Data Environmental Data (Depth, Temp, etc.)
- Classification Standards
- Citizen Science Portal Video Annotations
- Live Streaming
- Real-time Annotations and Comments (SeaScribe + Event log)
- Post Cruise Annotations (SeaTube)
- Standardized Annotations
- Habitat Characterization
- Video Portal (High Resolution Archive)
- Video/Annotation Browsing and Discovery (SeaTube)
- Animal Guide
- Oceanographic Data Archive
OER video workflow
Annotations play a key role!!

- ROV Video
- Sensor Data
  Environmental Data
  (Depth, Temp, etc.)
- Classification
  Standards
- Citizen Science Portal
  Video Annotations
- Habitat Characterization
- Post Cruise Annotations
  (SeaTube)
- Real-time Annotations
  and Comments
  (SeaScribe + Event log)
- Standardized
  Annotations
- Video Portal (High
  Resolution Archive)
- Oceanographic Data
  Archive
- Video/Annotation
  Browsing and Discovery
  (SeaTube)
- Animal Guide
OER Video portal ...
Great application - But only as good as meta data

Video Data Management Modernization Initiative schematic
(Image from VDMMI report 2016)
Assessing quality of annotations

• Do annotations cover all the observations?
  • Time and distance interval between annotations
  • How many observations were missed?

• Are annotations of sufficient quality?
  • Are annotations correct?
  • Are annotations correctly time stamped?
  • Do annotations follow standardized vocabulary
Subject metadata generation challenges

• Tools to effectively collect subject metadata

• Assess quality of metadata
  – Quantity of annotations
  – Correctness of annotations

• Standardized vocabulary
Video annotation tools challenge
ONC Seatube / Seascribe implemented in 2017

http://dmas.uvic.ca/SeaTube
Video annotation tools challenge
ONC Seatube / Seascribe implemented in 2017

Onshore and on ship annotations

Video and annotations linked

Search annotations, discover videos

post cruise annotations addition and edits

Map browsing display

2017 video statistics
~1600 Hours of video
97 Dives
~30 k Annotations

http://dmas.uvic.ca/SeaTube
Subject metadata generation challenges

- **Tools to effectively collect subject metadata**

- **Assess quality of metadata**
  - Quantity of annotations
  - Correctness of annotations

- **Standardized vocabulary**
Image processing

- Automated organism identification
  - An active field of research
  - Lots of potential
  - Still not deployment ready

- Can image processing aid with annotation process?
  - Which frames are different enough from adjacent frames to warrant an annotation
  - Vast majority of frames have no new information
  - **Goal:** Present researchers the only frames that require an annotation
Methodology

● Extract images from video
  ○ Used FFMPEG python subprocess call
  ○ Extracted frames for every 1 s
  ○ EX1706 Dive 02

● Identify frames that differ from adjacent frames
  ○ Histogram comparison in open CV
    ● Hellinger, Chi-square, Intersection, correlation
  ○ Hellinger Distance found to perform best

● Compare these results with human annotations
Hellinger distance

\[ d(H_1, H_2) = \sqrt{1 - \frac{1}{\sqrt{H_1 H_2 N^2}} \sum_I \sqrt{H_1(I) \cdot H_2(I)}} \]

H1, H2 are the histogram of two frames
N = No. of bins
I = Bins interval
Testing algorithm performance

RBG histograms of all images

Hellinger distance between frame pairs
Testing algorithm performance

RBG histograms of all images

Hellinger distance between frame pairs

User picked threshold
Histogram comparison over 5 minutes video

Other algorithms tested but Hellinger distance was found to perform better
Threshold 0.5

+1.49996e9

SeaScribe annotations

Histogram difference threshold crossed
Threshold 0.5 (zoomed)
Threshold 0.2

SeaScribe annotations

Histogram difference threshold crossed
Threshold 0.2 (zoomed)

SeaScribe annotations

Histogram difference threshold crossed
## Comparison of histogram vs. Seascribe annotations

### EX1706 Dive 02

<table>
<thead>
<tr>
<th>Method</th>
<th># of annotations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seascribe</td>
<td>215</td>
</tr>
<tr>
<td>Hellinger (Threshold 0.2)</td>
<td>563</td>
</tr>
<tr>
<td>Hellinger (Threshold 0.3)</td>
<td>251</td>
</tr>
<tr>
<td>Hellinger (Threshold 0.5)</td>
<td>59</td>
</tr>
</tbody>
</table>
Comparison of histogram vs. Seascribe annotations

Cnidaria Hydzoa Siphonophorae (Siphonophore): multiple individuals

Seascribe annotations were overlaid on frame as text
Frames with histogram difference > threshold

Frame with histogram difference > threshold were marked with a line
0.5 Threshold - Change in scene while zooming in
0.2 Threshold while zooming in
0.1 Threshold while zooming in
0.2 Threshold
0.1 Threshold
Histogram differences results

- 0.1 Detection threshold detected all the frames with seascribe annotations and then some more
- Frame contents changes:
  - Inclusion of a small organism does not change the histogram of the entire frame
- As ROV is operated while zooming in various objects - lower threshold detects changes in frames
- False positives results
- Is this better than viewing the entire video?
Is there any advantage of non-observations?
Is there any advantage of non-observations?

- **Computer Detection**
- **Human Annotation**

No need for human review?
Subject metadata generation challenges

• Tools to effectively collect subject metadata

• Assess quality of metadata
  – Quantity of annotations
  – Correctness of annotations

• Standardized vocabulary
How accurate are the annotations?

• Still not ready to address this question
• Need to develop validation strategies
  • Finding gold standard annotations (Schoening et al., 2016)
  • Comparison of real time and post cruise annotations
    • Use of Dr. Chris Kelley post cruise re-annotations?
    • Deep sea coral research and Technology program (DSCRTP)

Recommendations

• Improving the change detection algorithm
  • Pixel comparison for small changes
  • Sub-histograms of frames (frame divided into 4 images)

• Possibility of test deploying it in real time stream to alert users at times when an annotation is needed
  • Linking it with SeaTube real time streaming?

• DSCRP detailed annotations are now available for cross referencing

• Subject metadata standards
  • Based on time interval?
“Metadata liberates us, liberates knowledge”
Dr. David Weinberg- Technology commentator

Thank You
Questions ?

OceanExplorer.NOAA.gov
Bibliography


Abstract

Several of NOAA offices including NOAA Office of Ocean Exploration (OER) rely on the underwater videos and associated annotations for studying oceanographic processes and distribution of biota and fauna. Video annotations describe in text what is observed in the videos and provide necessary metadata to browse, search and analyze the videos. Vast amounts of NOAA video are currently un-annotated or have been annotated for a specific application hindering re-use of video data for multiple purposes. The key metadata requirements for video annotations include precise time of observation, adherence to standardized taxonomic vocabulary, completeness and correctness. Manual assessment of these quality metric(s) is not realistic. Automated and semi-automated methods that can define the quality of video annotations, by assessing the use of standardized vocabulary and completeness, can help prioritize the video annotation work. During 2017 OER implemented a ‘distributed’ annotation logging system onboard NOAA Ship *Okeanos Explorer* where participating scientists entered annotations online through an online web browser. A total of 993 hours of video was collected over 97 remotely operated vehicle (ROV) dives along with 29,315 annotations. Using these videos and annotations as test data set, this paper proposes a methodology to describe the video annotation quality. The entered annotations are compared against the standard agreed upon vocabulary to identify the annotations that have deviated from the standardized vocabulary. Using image analysis technique, the video frames are identified where a change in visual scene is detected. This provides a method to extract key frames of videos where an annotation is needed but missing. In order to provide scientists access to videos, associated audio, and annotation post cruise, OER has made all of 2017 ROV videos available online in partnership with Ocean Network Canada (ONC) at dmas.uvic.ca/SeaTubeV2. SeaTube provides the ability to browse the videos and enter/edit the annotations through an online interface. Quality descriptors implemented in conjunction with an online video annotation entry system can decrease the annotation workload by guiding scientists to video frames that are missing annotations as well as assessing if a video needs further annotations.
Adopting a vocabulary standard

Implementation of Worms
What is upcoming?

- New version of seatube with live streaming
- Standardized vocabulary
- Auto-complete function
- Possibility of alerting users when a new annotation is needed
Every other frame vs. 5 s, 10 s and 30 s interval

Selection of frame interval
OER video workflow
Incorporation of ONC tools during 2017 operations

Operational metadata (Ship and sensor information)
Raw Video
Human generated metadata (Observations: Annotation, Audio)
Ancillary data (CTD, samples, Bathymetry, Backscatter etc.)

Metadata includes
- Sensor information
- Time / spatial extent
- Few key words

2017 statistics
- ~ 1600 Hours of video
- 97 Dives
- ~30 k Annotations

How confident OER is about the annotations?