Automatic monitoring of birds in marine video content

T'jampens Roeland[1], Hernandez Francisco[1], Vandecasteele Florian[2] and Verstockt Steven[2]

[1] Flanders Marine Institute (VLIZ), InnovOcean site, Wandelaarkaai 7, 8400 Oostende, Belgium
E-mail: roeland.tjampens@vliz.be

[2] Data Science Lab, Ghent University - iMinds Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium

Technology makes it possible for scientists to study animals in their natural environment, without being present at that location. For example a camera at a nesting area allows scientists to observe from a distance. This reduces time spent on visiting the targeted area and limits the number of nest disturbances. The resulting video content gives scientists a unique insight in the life and behavior of the animals filmed. However, this leads to many hours of raw video content, which still have to be processed.

In the framework of LifeWatch, the Flanders Marine Institute (VLIZ) installed several cameras in a marine environment. Next to behavioral aspects, scientists are interested in the number of birds present during a certain time frame. This could be done by continuously watching and counting birds throughout the video content. But this is very time consuming, so an automatic processing is researched. The biggest problem encountered is the typical background in marine video content, namely the sea. Water is very dynamic, and next to waves and water ripples, also light reflections and shadows are visible on the surface, and can easily be interpreted as birds.

Most research done on the image processing of marine video content is for maritime purposes (harbor security and control). In this research, the techniques for maritime processing are tested for the monitoring of birds. An automatic detection of ships is deemed less difficult, due to the larger, and above all, static form of ships.

To identify the foreground items (i.e. birds) in video content, the background is subtracted first. A very dynamic background (i.e. water) is hard to distinguish from the foreground and a common problem in dynamic background subtraction (DBGS). In this research it was found impossible to remove all waves from marine video content without also removing some birds.

To counter this problem a two-step detection of birds was created. At first, without removing any birds, as many waves and reflections on the water as possible are removed. In a second step, the resulting foreground candidates are classified as water or bird by an offline trained classifier. Finally all valid detections are tracked throughout the video content (individual bird tracks).

- Step 1: For the removal of the dynamic background, several pixel based and block based techniques were tested. Pixel based techniques have a higher detail (per pixel) and are less complicated, and thus faster than block based techniques. But neighboring foreground pixels must be merged into a foreground candidate in each frame.
- Step 2: Birds can be distinguished from water by an image texture analysis. Namely the pattern of image gradients for birds differ sufficiently from water gradients to train a bird-water classifier on. A very rudimentary explanation is to envisage a bird as the Japanese flag (centered object), and water as mostly horizontal bands, like the flag of the Netherlands.
- Step 3: Tracking throughout the video is done by linking overlapping detections in subsequent frames. Once two detections are linked, the trajectory of the object is used. New detections are checked with the location predictions of all active objects. This works well for most cases, but is insufficient for birds that fly by near the camera (no overlapping).

The software created in this research allows for an automatic processing of marine video content. The output consists of a new video fragment in which detected birds are marked (identifier, bounding box and track) and the number of birds for each frame is indicated. Next to the new video fragment, a metadata (json) file is written. This file can be queried to know when (frame number) and where (position) birds were seen throughout the given video.

This research allows for scientists to focus on the statistical analysis of birds in a video without having to look at the entire video fragment.

Keywords: dynamic background subtraction; texture analysis; image classification; object detection; tracking; seabirds; marine environment