5/25/2015

**ZOOMIE v 1.0 (Zooplankton Multiple Image Exclusion)**

Moritz S. Schmid\*, Cyril Aubry, Jordan Grigor, Louis Fortier

\* Moritz.Schmid@takuvik.ulaval.ca

Takuvik Joint International Laboratory, Laval University (Canada) – CNRS (France), UMI3376, Département de biologie et Québec-Océan, Université Laval, Québec, Québec G1V 0A6, Canada

1. Introduction

ZOOMIE is an image treatment tool developed to ensure optimal quality for images collected with the Lightframe On-sight Keyspecies Investigation (LOKI) System, an underwater zooplankton camera system. ZOOMIE does that by identifying cases where multiple pictures of the same specimen have been taken (hereafter referred to as double images), a phenomenon that frequently occurs when imaging plankton in a constrained volume during vertical deployments. The process of identifying double pictures can be carried out manually but is very time consuming. By applying ZOOMIE, the time needed to identify double images is substantially reduced. It is essential to account for double images when representative distributions of images are wanted

ZOOMIE can automatically filter thousands of images based on previously extracted image parameters (e.g. area, mean grey pixel value, kurtosis; here extracted using the LOKI browser software (Isitec GmbH; <http://www.isitec.de/start.htm>)). The filtering is based on a set of rules that compares the image parameters of multiple images in order to detect double images and exclude them. The set of rules can be changed easily in the ZOOMIE scripts so that researchers can easily adapt the thresholds for finding double images necessary for their LOKI settings. After running the actual script to find double images, other scripts can be executed to automatically transfer images flagged for exclusion to a new folder.

Finally, the results can be visualized on an internal homepage, using the actual images which are linked to the database. Here we can validate the outcome of the processing and we can manually adapt the outcome through dragging and dropping of images to verify if any images were wrongly allocated to a double image group.

Although ZOOMIE was developed for LOKI images and the exclusion of double images, ZOOMIE could easily be adapted to handle other tasks requiring the handling and comparison of large numbers of images.

2. Problematic addressed by ZOOMIE

LOKI takes continuously images of marine plankton (mostly mesozooplankton from 200 μm up to 30 mm) on a vertical haul from close to the seabed to the surface. The channel where plankton is imaged is very thin, so the flow of particles through that channel is very sensitive to changes in pressure and the general movement of the system in the water column. During ship-based sampling, due to inherent environmental variability in the ocean, the heaving movement of the ship is transferred through the winch cable into the system and causes the system to move in a heaving fashion.

In addition, increasing amounts of plankton biomass in the net in front of the camera and in the camera channel can also change the pressure system and lead to changing flow through the camera channel.

In a perfect scenario the system would move straight up through the water at a constant speed and without pressure fluctuations in the camera channel, leading to only single images of animals being taken.

However, when particles go through the imaging channel slower than the calculated speed due to pressure changes, this results in multiple images per zooplankton being taken.

Visual verification cannot be completely abandoned, but with ZOOMIE the process is substantially faster than detecting all double images manually.

We want to encourage researchers using other plankton imaging systems to consider the problem described here and to take action. Our experience shows that double images occur rather frequently. If not accounted for, this problem will lead to biased distribution profiles and overestimates of plankton abundances.

3. How does ZOOMIE detect double images

The following processes are triggered (brief description): The underlying idea is to find images that are a) very similar (basic image parameters such as mean grey pixel and kurtosis are compared), b) taken only milliseconds apart and c) which are positioned together very closely as described by x and y coordinates, and are therefore likely the same animal (LOKI does not save the complete frame of view as a picture when deployed, instead it extracts *in-situ* only the regions of the overall frame where it detects particles. Therefore, every image that is saved on the solid state drive (SSD) during deployment has x and y coordinates associated with it which tell us where the particle was in the overall frame). We can use this to reconstruct particle movement across the imaging view. If this is the case we have identified one or several double images that should be flagged and later on excluded.

4. Basic setup instructions

Below are a list of instructions for setting up ZOOMIE on your PC.

a) First, download the WAMP server package at <http://www.wampserver.com/en/> , install it and configure it so that it can communicate with phpMyAdmin or similar software.

b) Now, add the loki folder from the ZOOMIEv1.0 package into the WAMP directory at: C:\wamp\www\loki

c) Next, load the oceanographie.sql file from the ZOOMIEv1.0 package into phpMyAdmin.

d) Next, you should import your image data (for example as exported from LOKI browser) from a CSV file or similar into the pictures table of the oceanography database (phpMyAdmin). Use the file headertemplate.csv from the ZOOMIEv1.0 package to make sure your column names are correct. The pictures table in the oceanography database consists of columns for each image parameter that was extracted as for the time when the image was taken etc. This data will be used by the script to filter for double images.

e) If a visualization of the images on the internal validation webpage is wanted, the raw/underlying images have to be placed into the image folder at: C:\wamp\www\loki\public\pictures

f) Now you can start with processing the data. Follow the next chapter for that.

5. Workflow and commands

After turning on all WAMP services your internal website should be at:

<http://loki.local/>

The phpMyAdmin site with the oceanography database (where the csv table with the image data has to go) is here:

<http://localhost/phpmyadmin/sql.php>

Now we are ready to run the first commands:

**1)**

First you have to generate new IDs for your data by running this command in the sql tab of the pictures table:

update pictures set id = MD5(Image);

This command will update the column ID and you should see unique values there from now on.

**2)**

Now we have to set the timestamp to the value NULL. Do this by writing again in the sql tab:

update pictures set timestamp = null;

**3)**

Now, start navigating in the command prompt to: cd C:\wamp\www\loki

**4)**

Now run the first ZOOMIE script to parse the file names. Parsing the file names means here that the script will break the unique LOKI image names into its components and put them into separate columns. This way we get the necessary columns for time and image position (x and y).

Execute the following in the command prompt:

php artisan parse:filename

Debugging parse:filename:

Check the following

1. Make sure there is no second header row. Sometimes mistakes during import can lead to multiple header rows. Might need some refreshes (reload database by clicking on pictures table)
2. Make sure that the anticipated values are in the posx and posy timestamp columns and that the algorithm gives back 0 rows left when running parse:filename

**5)**

Run the double filter algorithm by copy pasting the following command in the command prompt:

php artisan processing:data

Debugging the processing:data command:

If for some reason the values for the column “state” are not “real” before running the processing:data command, run the reset:images command. You also have to run the reset command if you want to reprocess an already processed data table. The script will only process rows where the “processed” column value is “0”. You also have to run the reset command if you want the consecutive numbering in the column “group\_id” to start at 1 again.

To run the reset command, type the following into the prompt:

php artisan reset:images

**6)**

**Verifying double images**

Go to: <http://loki.local/>

If you processed some data and the underlying raw images are in the image folder then you will see the results of the double filter algorithm here. Images will be grouped together by column group\_id. If multiple images have the same group\_id, this means that the first image is the original one and all others after are double images. You can also drag and drop images if you want to manually adjust the outcome of the double filtering which will change the values in the underlying database as well. In addition you can also manually exclude a picture by clicking delete, or put an image into a new group (if it was wrongly assigned as a double image) by clicking new group.

**7)**

**Reallocate command**

After the doubles are verified, run the reallocate command to make sure that only one image in every group ID has the value “real” as “state”.

Execute:

php artisan reallocate:real

**8)**

**Move bad images**

This command will move all images in the image folder which have as a state “double” in the column “state” into a new folder that is created automatically and is called “bad”.

Execute:

php artisan move:badimages

**Now you separated the wanted unique images from the unwanted double images and are ready to further analyze the images.**

---------------------

The ZOOMIE package consists of:

The zipped folder ZOOMIEv1.0 consisting of:

1. ZOOMIE – Introduction v 1.0.docx
2. CSV header template (headertemplate.csv)
3. The loki folder which has to be placed into the WAMP folder on your local machine.
4. The database file oceanographie\_2015-03-30.sql for phpMyAdmin which you can then fill with your own data. I left it populated with sample data.

ZOOMIE was built on Windows 7, 64 bit.

Thank you to Nicolas Garneau for initial set up (paid basis, Université Laval).

Please let us know if your project benefits from ZOOMIE and if you make any adjustments/improvements to the ZOOMIE package. It would be great to hear from you.

Please contact the authors for more information and collaboration.

Moritz S. Schmid

<http://www.schmidscience.com>

Moritz.Schmid@takuvik.ulaval.ca

MoritzSSchmid@gmail.com