Midwater Classification Data and Results

Filename: Poeobius Database, Test Images and Results.doc
\\Tornado\ProjectLibrary\900260.AVED\Project.Documents\Project.UserManual

Created: Wednesday, July 20, 2005, 1:02:44 PM
Last Modified: Friday, January 06, 2006, 1:24:18 PM

The directories and subdirectories containing the training and test images as well as the generated matlab files and results for poeobius classification can be found on **nanomia** in **/RAID**. The **nanomia** backups are organized in such a way that the original movies or master ppms are separated from results. The original movies and avis for the training set are found in **/RAID/video/midwater/Poeobius/2613_Poeobius_training_movies**. The original test master ppms are found in **/RAID/video/midwater/Poeobius/Poeobius_testing** which contains four subdirectories representing each of the tests**.**

The images in the training databases can be found in **/RAID/classifier/results/ Poeobius_Classification/training_sets**. Images from the sets used for testing can be found in /**RAID/ classifier/results/Poeobius_Classification/testing_sets** and MATLAB files generated from image collection, training and classification can be found in **/RAID/ classifier/results/Poeobius_Classification/matlab_files**. Within these directories the subdirectory structure outlined below is followed.

Poeobius Database

The poeobius images database used for training purposes is located in the subdirectory, **/midwater-DB**. This directory is organized as follows:

The directory **background-all** contains all the poeobius images annotated in dive 2613 as well as a large number of unique images to represent an Other class contained in **Otherbackground**. These images have the background removed.

The directory **background-best** contains a smaller sampling of poeobius annotated in dive 2613. Some poeobius images from **background-all** had been removed based on a run that trained and tested the whole sample. If the image failed to be recognized as a poeobius it was then evaluated as to whether it should be part of the sample set. The Other class, **Otherbackground**, is the same as the one in **background-all** as the images had already been evaluated for their quality. These images also were generated with the background removed.

The directory **background-best-add16** contains 16 poeobius that had been added to the training set from the dive 1583 test sets. From one of the 10 minute segments of this dive, all the poeobius were added into the training set. As for the other 10 minute

segment, all the images from this set were evaluated against the **background-best** database explained above. The poeobius that had been partially or wholly recognized by the classifier were added into the training set and thus removed from a newly created test set. Partial recognition is interpreted to mean a poeobius being correctly classified in at least one of the events assigned to the same animal. If one or more events for a given poeobius were recognized, all the corresponding events for this poeobius were added to the **background-best-add16** directory. The Other class, **Otherbackground**, is the same as the one in **background-all**. All the images in this database were generated with the background removed.

The directory **background-best-add38** contains 38 poeobius – the previous 16 already added and 22 new poeobius that had been partially or wholly recognized. These poeobius derived from the same test set created from Dive 1583 where 16 poeobius were removed and evaluated against the **background-best-add16** database. The Other class, **Otherbackground**, is the same as the one in **background-all**. All the images in this database were generated with the background removed.

The square directories - **square-all, square-best, square-best-add16, square-best-add38** - were created in a similar vein but use images that were made square using the natural scenery.

The directory **PBSremoved-bad_images16** contains images from the 16 poeobius added into the database but which were of poor quality and thus not added into the **background-best** training set. These images were kept in case one wished to add them into the **background-all** training set. This directory contains both images made square using the natural scenery and square images with background removed.

The directory **PBSremoved-bad_images38** contains images from the additional 22 poeobius that had been added into the database but which were of poor quality and thus not added into the **background-best** training set. These images were kept in case one wished to add them into the **background-all** training set. This directory contains both images made square using the natural scenery and square images with background removed.

AVED Results and Original Movies used for Poeobius database

The subdirectory **2613_Poeobius_training_movies** located on nanomia in **/RAID/video/midwater** contains all the preprocessed quicktime movies from dive 2613, the dive used to collect all the poeobius images contained in the original database (no additional poeobius added from test sets).
back
The subdirectory **2613_AVED_results** located on nanomia in **/RAID/classifier/results/Poeobius_Classification** contains all the processed results of each movie clip in their corresponding folders. Each of these directories contains an events subfolder which has the collection of poeobius images and other images gathered from a specific movie clip. This same data is also collectively gathered in the **other**

directory for other images and the **poeobius** directory for poeobius images under the same main folder, in organized categories.

Test Data

The master ppms are in the subdirectory are located on nanomia in **RAID/video/midwater/Poeobius/Poeobius_testing.** Each of the subfolders represents one of the ten minute midwater segments used for testing purposes.
The images generated for the tests were gathered from two dives – dive 1583 and dive 1381. Two ten minute segments of each dive using square and background images were tested against the corresponding database.

The images used for testing purposes can be found on **/RAID/classifier/results/Poeobius_Classification/test_sets**.
Directory **1583_01_52_49_00** contains the data generated from a 10 minute AVED run. The subfolder **\events** contains three subfolders **background, square** and **off-transect**. The off-transect directory contain event images that were out of the ten minute transect range and thus not included in the test set. The background contain all the event images of the ten minute test segment with background removed and square contain all the event images that were created using the natural scenery. On **nanomia** in **/RAID/classifier/results/Poeobius_Classification_test_sets/1583_01_52_49_00** there is also a subdirectory with the name **1583_01_52_49_00-PBSremoved-events** which contain all the events but with poeobius images removed from the event set. This was used for testing purposes. The images were added into the **background-best-add16**and **square-best-add16** mentioned above and thus also exist in **background-best-add38** and **square-best-add38.**

The subdirectories - **1583_02_53_36_00, 1381_01_44_30_00** and **1381_03_28_00_00** - are organized in a similar manner and correspond to the remaining thirty minutes of the test set. They all contain an events folder used for testing purposes. However, on **nanomia** the **1583_02_53_36_00** contains two other subdirectories **1583_02_53_36_00-rem16-events** and **1583_02_53_36_00-rem38-events.** These two subfolders have the same events as the events folder but with 16 or 38 poeobius removed from the events set. These were used for testing purposes. These poeobius were either classified correctly or some of their events were classified as poeobius after an initial run. The entire set of images corresponding to these poeobius were removed from the testing set and added to training sets to see if there was an improvement in the performance of the classifier. The training sets containing the additional images can be found in the **background-best-add16** and **square-best-add16** mentioned above. The initial 16 removed can be found in all of the training sets with additional images whereas the ones removed from the **1583_02_53_36_00-rem38-events** directory set can only be found in **background-best-add38** and **square-best-add38.**

MATLAB Data Collection and Results

The MATLAB files generated from classification can be found in the subdirectory **/matlab_files** . The subdirectory **DB** contains the files generated from collecting the database images and the resultant training data based on them as well as collecting the test images. The training files are in the **training** subfolder and testing files are likewise in the **testing** subfolder.

The training folder contains the subfolders: **background-all, square-all, square-best-add16, background-best, square-best** and **square-best-add38.** These directories contain the matlab files produced from collecting and training the images derived from the image directory that bears the same name.

The testing subfolder contains the subfolders: **1381_01_44_30_00-square, 1381_03_28_00_00-square, 1583_01_52_49_00-background, 1583_01_52_49_00-square, 1583_02_53_36_00-background, 1583_02_53_36_00-square, 1583_02_53_36_00-square-rem16, 1583_02_53_36_00-square-rem38** as well as some tests conducted using the images in the training database ie. **PBSbackground-best** or **OTHERsquare**. All of these subdirectories contain the matlab files produced from collecting the test images contained in the image directory bearing the same name.

The results from the MATLAB classifier are stored in **/matlab_files/** in the subdirectory **RESULTS**. The naming convention is again applied. Thus the **RESULTS** directory contains the following subdirectories: **1381_01_44_30_00-square-best, 1583_02_53_36_00-square-all, 1381_03_28_00_00-square-best, 1583_02_53_36_00-square-best, 1583_01_52_49_00-background-all, 1583_02_53_36_00-square-best-16, 1583_01_52_49_00-background-best, 1583_02_53_36_00-square-best-38, 1583_01_52_49_00-square-all, 1583_01_52_49_00-square-best, 1583_02_53_36_00-background-all, 1583_02_53_36_00-background-best** as well as results from other tests. Each of these subdirectories have matlab files that detail the classification results. The names of the files in these directories specify the method used to determine the results. For example the directory **1583_02_53_36_00-square-best** contains the following results:

1583_02_53_36_00-squareandbest-majority.mat – Derived from using an **and** switch in the file **test_ljmNL3.m**. Uses the **square-best** training set on the **1583_02_53_36_00 square** test set. Classifies each **event** in the test set using the **majority** method.

1583_02_53_36_00-squareandbest.mat – Generated MATLAB file containing the classification of each individual **file** in the test set using the same training and test set as well as the same parameters.
1583_02_53_36_00-squareandbest-probability.mat - MATLAB file containing the results of **event** classification based on the **probability** method using the same training set, testing set and parameters.
1583_02_53_36_00-squareorbest-majority.mat – The same as above but using an **or** switch in the file **test_ljmNL3.m.**

1583_02_53_36_00-squareorbest.mat - The same as above but using an **or** switch in the file **test_ljmNL3.m.**
1583_02_53_36_00-squareorbest-probability.mat - The same as above but using an **or** switch in the file **test_ljmNL3.m.**