WP1: Video Data Analysis

Leading : UNICT
Participant: UEDIN

Department of Electrical, Electronics and Computer Engineering
University of Catania (Italy)

Fish4Knowledge Review Meeting - December 14, 2011 - Catania - Italy

Teams

- UNICT Team
  - Di Salvo Roberto
  - Giordano Daniela
  - Kavasidis Isaak
  - Palazzo Simone
  - Spampinato Concetto
- UEDIN Team
  - Boom Bas
  - Fisher Bob
  - Huang Xuan

Overview

- Underwater ecological observation system
- Description and Objectives of the WP
  - Fish Detection (UNICT)
  - Fish Tracking (UNICT)
  - Fish Description (UNICT)
  - Fish Recognition (UEDIN)
  - Fish Clustering (UEDIN)

- 9 cameras continuously recording during daylight
- Video stream is divided into 10 minute long videos:
  - Multiple resolutions (320x240 and 640x480)
  - Multiple formats, such as MPEG-1/2/4, WMV, FLV
  - Different frame rates ranging from 5 fps to 30 fps.
- 4000 hours of video now recorded available at *http://gad240.nchc.org.tw/*

- Underwater scenes
  - multimodal background
  - sudden and gradual light changes
  - bad weather conditions
  - murky water
  - algae on camera lens
  - periodic movements
- Targets
  - Erratic motion in three dimensions
  - Sudden changes in appearance
  - Non-rigid movement
  - Fish occlusion

Features of Underwater Video Data

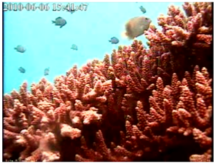
Sample Images

NPP3-1

NPP3-2

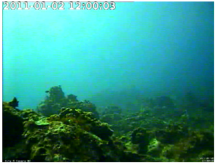
NPP3-3

NPP3-4



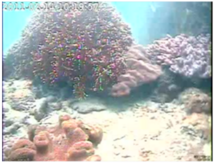
LanYu-1

LanYu-2



HoBiHu-1

HoBiHu-2



HoBiHu-3

Workpackage 1

- **Objectives**
  - O1.1 Fish/marine animal detection, tracking;
  - O1.2 Extract a set of properties to describe fish;
  - O1.3 Fish Recognition and identify clusters of unrecognised fish;
- **Tasks**
  - T1.1 **Fish Detection**: Background/foreground modeling algorithms able to deal with marine domain
  - T1.2 **Fish Tracking**: Covariance model to handle occlusions and temporary loss of fish
  - T1.3 **Fish Description**: Affine invariant fish descriptors
  - T1.4 **Fish Recognition and Clustering**: Recognition using a combination of colour, texture, active appearance models and special purposes features such as head, tail, fin size estimates.

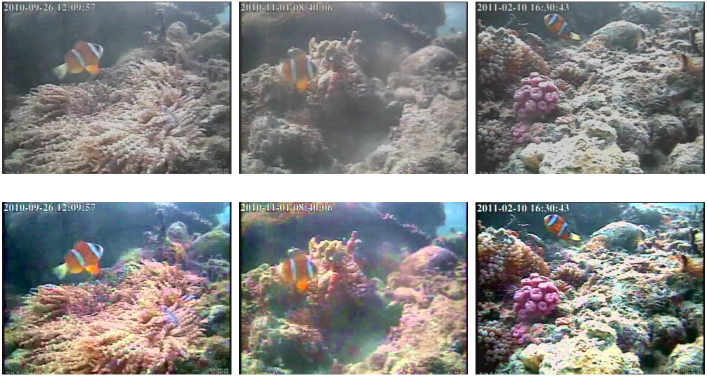
T1.1: Fish Detection
Description and Motivations

- **Objective**
  - Detection algorithms should be able to handle both the effects occurring in underwater scenes and frequently changes in size and appearance of fish
- **Methods**:
  - Mixture of *pdf*s (Gaussian and Poisson)
  - Intrinsic Model
  - Wave-back
  - Adaptive Multi-distribution Model

- **Pre-processing**
  - Frame Enhancement
- **Post-processing**
  - Blob Level: Quality Score
  - Pixel Level: Contours Improvement

T1.1: Fish Detection
Pre-processing: Frame Enhancement

- Total independence of the image formation process, and no a priori knowledge of the environment
- Contrast stretching both in RGB and in HSI space

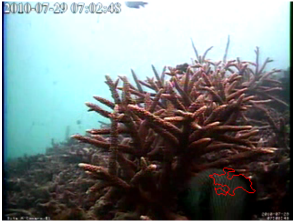


Quality Score: score describing how sure we are that a detected blob be a fish:

- Difference of color at object boundary
- Difference of motion vectors at object boundary
- Internal color homogeneity
- Internal motion homogeneity

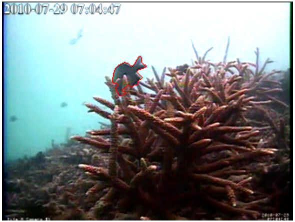
T1.1: Fish Detection

Post-processing: Examples of quality scores



Quality Score: 0.39

Quality Score: 0.61



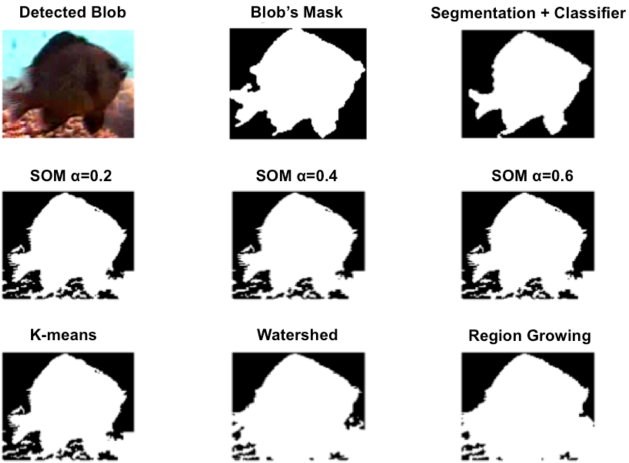
Quality Score: 0.75

Quality Score: 0.89

- **Segmentation Methods**
  - Self-Organizing Maps (SOM)
  - Watershed
  - Region Growing
  - K-Means
- **Segmentation wrapped inside a classifier**
  - Correct segmentation not based on some low-level image homogeneity of the object, i.e., color, grayscale, or texture, but on the probability of correct classification of a proposed segmentation for a given class

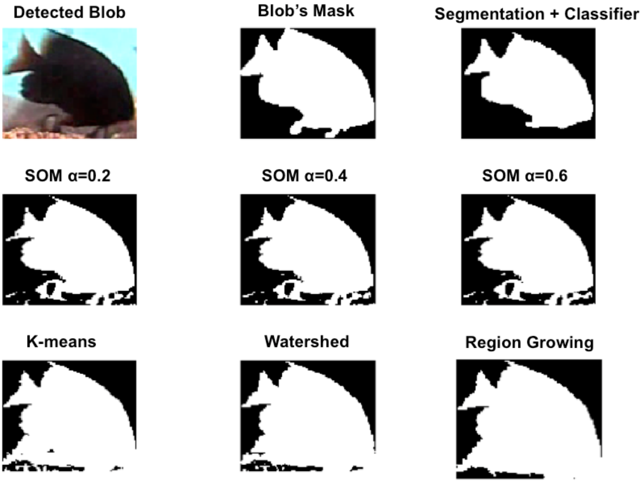
T1.1: Fish Detection

Post-processing: Contour Improvement



Fish Detection

Post-processing: Contour Improvement



- Aspects to deal with:
  - The appearance of a fish changes continuously because of lighting, orientation, non-rigidity
  - Occlusions might temporarily hide an object
  - Searching region limited to a neighbourhood of the object
- Solution:
  - To represent in a compact way both spatial and appearance information and the correlation between them.

T1.2: Fish tracking
Covariance based tracking algorithm

- Feature vector: RGB values, hue, local histogram moments
- Covariance matrix
- Förstner's distance used to compute the similarity between covariance matrices
- Adaptive search area to handle the temporary loss of tracked objects

T1.2: Fish Tracking
Qualiy score

- Quality score computed for each tracking decision as the average of:
  - *Shape ratio variation*
  - *Histogram difference*
  - *Direction smoothness*
  - *Speed smoothness*
  - *Texture difference*
  - *Temporal persistence*

T1.2: Fish Tracking
Qualiy score

Quality Score: 0.91

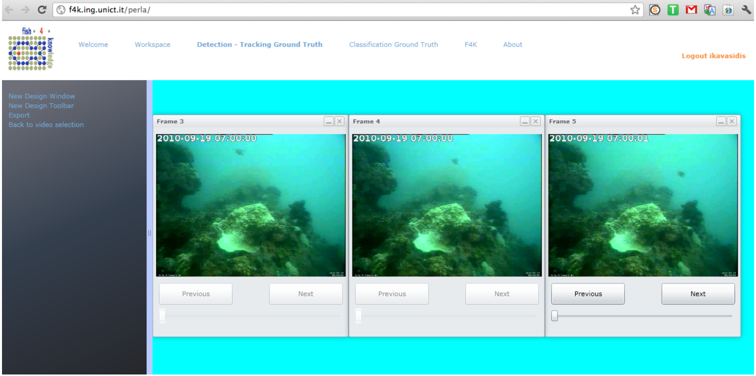
Quality Score: 0.81

Quality Score: 0.63

Quality Score: 0.71

Repair tracking failures: tracking as an optimization problem where the global maximum score has to be found in consecutive tracking decisions for each trajectory

PERLA: Performance Evaluation and gRound truth LAbeling
http://f4k.ing.unict.it/perla



Performance Evaluation

Ground Truth

- Ground truth quality (between 0 and 1) assessed by using PASCAL Score and Euclidean Distance Score with a very accurate ground truth carried out on a subset of objects
- 5 videos with the highest ground truth qualities: resolution of 320×240 with a 24-bit color depth at a frame rate of 5 fps

Video	Description	N_F
1	Dynamic Background Striped Fish Texture	156
2	Highly Dynamic Background	1373
3	Typhoon Frequent illumination variations Very low contrast	1790
4	Typhoon Plants movements	34
5	High illumination Striped Fish Texture	840

Fish detection rate (DR) and false alarm rate (FAR)

	No pre/post-proc.		Image. enhanc.		Blob post-processing	
	DR	FAR	DR	FAR	DR	FAR
AGMM	70%	18%	79%	16%	86%	11%
APMM	67%	20%	76%	17%	84%	8%
IM	70%	16%	74%	14%	87%	7%
WB	58%	20%	66%	13%	75%	5%
AMDM	73%	17%	79%	12%	89%	9%

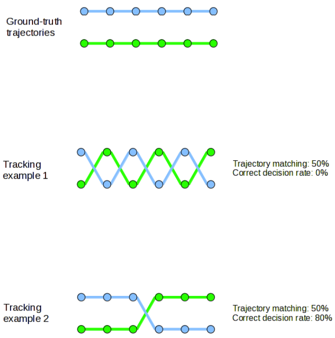
Pixel detection rate (PDR) and pixel false alarm rate (PFAR)

	No pre/post-proc.		Image. enhanc.	
	PDR	PFAR	PDR	PFAR
AGMM	92.6%	18.1%	92.4%	16.2%
APMM	92.7%	21.4%	89.4%	23.0%
IM	87.4%	25.1%	89.0%	23.6%
WB	94.6%	28.2%	93.2%	27.2%
AMDM	93.8%	21.6%	92.7%	17.0%

- Using segmentation, the PFAR drops by about 5-10%.

Performance Evaluation
Fish Tracking

- Correct Counting Rate (CCR).
- Average Trajectory Matching (ATM)
- Correct Decision Rate (CDR)



Comparison between the results obtained by the proposed algorithm and CAMSHIFT on the ground-truth data

	Covariance tracker	CAMSHIFT
CCR	91.3%	83.0%
ATM	95.0%	88.2%
CDR	96.7%	91.7%

Computation time per frame by algorithm and pre/post-processing levels

	No pre/post-proc.	Image. enhanc.	Blob post-processing
AGMM	25 ms	60 ms	75 ms
APMM	30 ms	70 ms	85 ms
IM	120 ms	160 ms	190 ms
WB	85 ms	120 ms	140 ms
AMDM	60 ms	90 ms	115 ms

Performance Evaluation
Database Content Overview

Total number of processed videos, detections and fish

Number of processed videos	2825
Number of detections	3869473
Number of fish	456622

Number of processed videos, detections and fish by algorithm

	AGMM	*APMM*	*IM*	*WB*
Number of processed videos	2825	2825	2825	2825
Number of detections	731049	708292	1326058	1104074
Number of fish	97267	91925	177609	89821

Number of processed videos, detections and fish by location

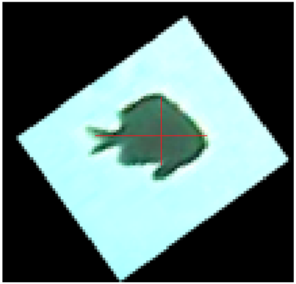
	NPP-3	*HoBiHu*	*Lanyu*
Number of processed videos	2367	545	138
Number of detections	1007794	43926	3572
Number of fish	123528	7753	603

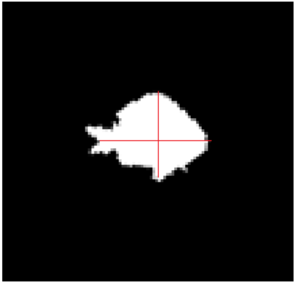
Color		Texture		Motion		Contour	
Name	**Resp.**	**Name**	**Resp.**	**Name**	**Resp.**	**Name**	**Resp.**
Background Scoring *	UC	Gabor Filter	UC/UE	Motion Vector	UC	Rigid Points *	UC
RGB, nor RGB	UE	SIFT	UC/UE	FTLE	UC	CSS	UC
HSV, HSL	UE	GC-SIFT	UE	Periodic Motion Analysis *	UC	Curvature Points	UC
L_{ab}	UC/UE	PCA-SIFT	UE			Fourier Descriptors	UC
Joint Histogram	UC	Covariance	UC			TPS	UE
Transf. Color *	UC	Co-occurences	UC			ASM/AAM	UE
Color Moments	UC	Spots/Stripes	UE			MDL	UE
HSV SIFT *	UC	Symmetry Hierarchies *	UC			Shock Graph	UE
RGB SIFT *	UC					Mellin Transform	UE
						Wavelet	UC
						Implicit Polynomials *	UC

Preliminary List of Fish Descriptors that will be used in detection, tracking and recognition processes. Most of these descriptors have been already implemented except the ones indicated with *.

T1.4: Fish Recognition

Fish Descriptors

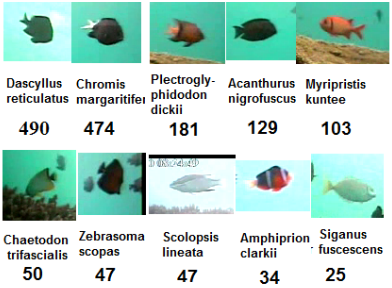




- 30 color features
  - 5 parts: head, tail, top, bottom, whole fish
  - 2 attributes: mean and variance
  - 3 descriptors: normalized Red & Green, H in HSV
- 1 boundary feature
  - radio of fish tail's variance and whole fish's.

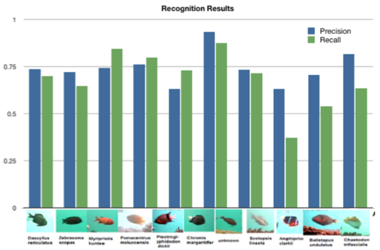
T1.4: Fish Recognition

Preliminary Results



- Classifier: Linear PEGASOS SVM
- 4 fold cross validation

Preliminary results in terms of precision and recall

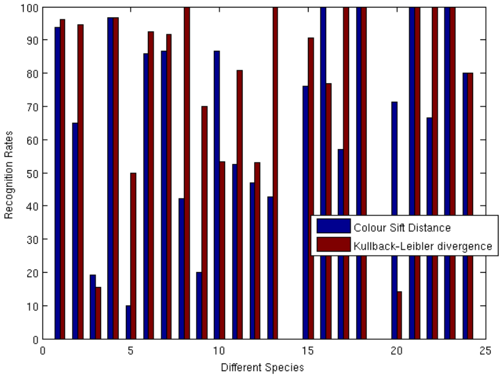


Average precision and recall, respectively, 0.736, 0.701

- Purpose:
  - Supporting Ground-Truth Annotation
  - Supporting recognition, recognising cluster of fish instead of single fish
- Two methods:
  - **Bag of Features:**
    Sift Features with Color Information
  - **Kullback-Liebler Divergence:**
    Color, Texture and Shape features

T1.4: Fish Clustering
Preliminary Results

- **Ground Truth Data**: 3424 Fish images, 25 Fish Species, Unevenly distributed
- **Colour Sift**
  Total Recognition Rate: 87.4%
  Mean-Class Rec Rate: 68.2%
- **KL Divergence**
  Total Recognition Rate: 92.6%
  Mean-Class Rec Rate: 79.8%



Conclusions

- Satisfactory performance of fish detection and tracking
- Expectation of improved performance when processing higher-resolution videos
- Implementation of affine invariant descriptors for colour, texture, motion and shape/contour
- Preliminary results on fish recognition are encouraging
- Effective image clustering methods

Publications

- Accepted:
  1. C. Spampinato, S. Palazzo, A. Faro - Event Detection in Crowds of People by Integrating Chaos and Lagrangian Particle Dynamics, Proceedings of the 2011 3rd International Conference on Information and Multimedia Technology (ICIMT 2011) Dubai, UAE, December 28-30, 2011
  2. S. Palazzo - Object Tracking: State of the Art and Online Performance Evaluation. Proceedings of the IEEE International Conference on Computer and Multimedia (CAMAN 2012), March 9-11, 2012 Wuhan, China.
- Under revision:
  1. C. Spampinato, S. Palazzo, B. Boom, J. van Ossenbruggen, I. Kavasidis, R. Di Salvo, F-P. Lin, D. Giordano, L. Hardman, B. Fisher, "Understanding Fish Behavior during Typhoon Events in Real-Life Underwater Environments", Special Issue on Real-life Events in Multimedia: Detection, Representation, Retrieval, and Applications, Multimedia Tools and Applications (MTAP) Journal, Springer.
- Submitted:
  1. C. Spampinato, S. Palazzo, D. Giordano, F. Lin, Y. Lin, Covariance based Fish Tracking in Real-Life Underwater Environment, VISAPP 2012.