A multi-modal event detection system for river and coastal marine monitoring applications

Edel O'Connor, Alan F. Smeaton, Noel E. O'Connor
CLARITY: Centre for Sensor Web Technologies
Dublin City University, Glasnevin, Dublin 9
Email: edel.oconnor@computing.dcu.ie. alan.smeaton@dcu.ie, noel.oconnor@dcu.ie

***Abstract*—This work is investigating the use of a multi-modal sensor network where visual sensors such as cameras and satellite imagers, along with context information can be used to complement and enhance the usefulness of a traditional in-situ sensor network in measuring and tracking some feature of a river or coastal location. This paper focuses on our work in relation to the use of an off the shelf camera as part of a multi-modal sensor network for monitoring a river environment. It outlines our results in relation to the estimation of water level using a visual sensor. It also outlines the benefits of a multi-modal sensor network for marine environmental monitoring and how this can lead to a smarter, more efficient sensing network.**

I. Introduction

Water quality monitoring is an important part of monitoring our natural environment and includes monitoring the quality of both coastal and inland marine locations. It covers the detection of pollution and monitoring the development of harmful algal blooms as well as tracking coastal features and wave patterns. For many years water managers relied on field measurements for coastal monitoring and water quality evaluation. This involves costly, time and labour-intensive on-site sampling and data collection, transportation to laboratories for analysis, and then subsequent evaluation. This type of sampling is also too limited on temporal and spatial scales to adequately monitor the quality of water bodies on a long-term basis or to address the development of events such as harmful algal blooms and fish kills.

New technologies are helping to streamline the water quality monitoring process. In recent years, the use of in-situ wireless sensor networks (WSNs) for marine environmental monitoring has been investigated to allow continuous real-time remote monitoring of the marine environment at greater temporal and spatial scales. This provides an opportunity for long-term data collection at scales and resolutions that are difficult or impossible to obtain otherwise. Important indicators of water quality can be continuously monitored with the possibility of real-time alert notifications of harmful marine events. Greater temporal and spatial sampling also allows environmental processes and the well-being of our waterways to be monitored and characterised from previously unobtainable perspectives.

The last ten years have seen the emergence of a variety of initiatives from very simple WSN deployments to highly complex coastal observation systems which make physical, chemical and biological measurements. However the current state of the art in this technology still poses many challenges for environmental monitoring applications [1]. In particular the more sophisticated of these sensing devices - chemo- and bio-sensors - which are really of utmost importance in environmental monitoring applications, are not particularly suited to the large-scale long-term deployments that are required by environmental monitoring applications. Many of these devices have a limited deployment lifetime (e.g. samples) before they begin to experience signal drift and require maintenance. They are also high in cost and require significantly more energy than their less sophisticated counterparts [1]. What in essence is required is adaptive sensing environment whereby these sensors can sample more efficiently and be placed more effectively in pollution hotspots.

There is also an issue with data reliability. Maintaining continuous reliable signals from sensors deployed in the marine environment can prove very challenging and failure of sensing devices is not uncommon. It is also the case that in the future large-scale deployments of chemo-bio sensor platforms will more than likely consist of low-cost unreliable devices that will be used to modify the operating characteristics of the more sophisticated platforms which are less densely distributed [2]. Therefore future challenges now lie in understanding how sensors can be used collaboratively in a hierarchical network ranging from relatively low-cost dumb devices to the more expensive sophisticated platforms.

Other issues also arise in relation to the deployment of sensor networks for marine environmental monitoring. In-situ sensors can improve the scale of sensing but only up to a point. They have limited spatial resolution as they are in effect single point sensors and often the region of interest in a marine environment may be quite vast. Furthermore, due to the expense and logistical difficulties often associated with the deployment of an in situ sensor network in marine environments, it may be difficult to monitor a wide area over long periods of time. Certain environments may not even be suited to monitoring by an in-situ WSN e.g the turbulent nature of the surf zone often makes it difficult to successfully maintain in-situ instrumentation [3]. Finally certain events may occur that may not necessarily be immediately detected by in-situ instrumentation. For example if there is pollution floating on our water, water managers may not be automatically alerted by readings from the in-situ observations. However it may be vital that this is attended to immediately.

Our work is investigating the use of a multi-modal sensor

network where visual sensors such as cameras and satellite imagers, along with context information can be used to complement and enhance the usefulness of a traditional in-situ sensor network in measuring and tracking some feature of a river or coastal location. Multiple sensing modalities provide increased information, the characterisation of an environment from multiple different perspectives, greater spatial resolution and the otherwise unobtainable detection and early warning of certain marine events. It also leads to a smarter network whereby different modalities can be used to control the operating characteristics of the more sophisticated nodes in the network which improves efficiency.

This paper focuses on our work in relation to the use of an off the shelf camera as part of a multi-modal sensor network for monitoring a river environment. It outlines our results in relation to the estimation of water level using a visual sensor. It also outlines the benefits of a multi-modal sensor network for marine environmental monitoring and how this can lead to smarter, more efficient sensing network.

II. Related Work

The work carried out in this paper follows on from work carried out in [4] which also focuses on estimating river water level using visual sensing. The approach described in this paper again uses the land water boundary in order to determine water level. However it adopts a more sophistiacted classification approach and a different camera angle that contains more distinct features. It also takes into account difficulties that were encountered in the work reported in [4] where it was difficult to model pixels as water or land due to the varying lighting conditions. It attempts to overcome these difficulties and to take a more adaptable approach that can easy be applied to a variety of images from different camera angles.

In other studies, coastal video systems have been identified as effective tools for coastal monitoring. A prime example of this is a major European research project entitled CoastView [5]. This focused on the development of video systems in support of coastal zone management utilizing Argus technology. Argus stations consist of optical systems developed for nearshore sampling [6]. The CoastView project demonstrates the use of fixed video remote sensing systems to partially ameliorate some of the problems associated with in-situ measurements of waves, currents, and morphological change. Davidson et al. [5] refers to some of the research carried out investigating algorithms for the quantitative extraction of geophysical signals from image data including morphology [7], flows [8] and wave parameters [9] and refers to the various scientific literature that has tested and reviewed the reliability, accuracy and versatility of coastal video systems. Alexander and Holman [3] used time exposure images from the Argus Network in the quantification of nearshore morphology and Chickadel and Holman [10] investigated the use of video techniques for measuring longshore currents. The Argus system allowed large amounts of data to be gathered over larger spatial areas for longer time periods than could be achieved with the in-situ instrumentation.

In [11], Goddijn-Murphy et al explore the possibilities of employing a conventional digital camera, as an alternative low-cost technique to satellite imagers or multi-spectral radiometers, to estimate water composition from optical properties of the water surface. This paper presents the method that was used to acquire digital images, derive RGB values and relate measurements to water quality parameters. Measurements were taken in Galway Bay and in the North Atlantic. Both yellow substance and chlorophyll concentrations were successfully assessed using this method. In [12], Iwahashi et al also investigate detecting water level using a land water boundary. However their work is using a a video signal as opposed to still images. Also they aim to classify pixels as land or water, where they are assuming that the land region contains solid objects with fine texture full of high frequency components. This is not always the case with the land region in the images in this study, which under certain conditions can partially appear visually similar to water.

Other studies have investigated the use of imagers not only in the context of monitoring a marine environment but also in other forms of environmental monitoring applications. In [13], a scalable end-to-end system for vision-based monitoring of avian behaviour during a nesting cycle is presented. The manual collection of phenological data can prove to be labour intensive and thus requires the use of innovative new methodologies such as the use of digital cameras e.g [14], [15]. In [14], Graham et al. investigate the use of visible light digital cameras in determining the dynamics of expanding leaf area for Rhododendron occidentale, a deciduous understory shrub.

III. Visual Sensing System

A visual sensing system has been developed for two marine locations. The first system is based upon a river environment situated at the River Lee, Cork Ireland. An in-situ sensor network was deployed at five locations along the river as part of the Deploy project [1]. As part of this research we deployed a camera at one of these locations. We gathered contextual information such as rainfall radar images from Met Eireann [2], along with in-situ and visual data from the site and a visual data analysis and data modelling toolkit was developed around these data streams. Secondly a satellite image analysis system was developed with a view to analysing and searching satellite information for the Galway Bay area on the west coast of Ireland. This site was chosen due to the initialisation of a national test and demonstration research infrastructure supporting a range of sensing technologies known as SmartBay. An outline of this system is provided in [16]. However the focus of this paper will be on the visual based sensing system developed for the site at the River Lee.

This paper demonstrates our results in relation to the estimation of water level at the site. It outlines the benefits of incorporating multiple sensing modalities into an environmental event detection system, the parameters that can be identified

[1] $http://www.deploy.ie$
[2] $http://www.met.ie/latest/rainfall_radar.asp$

and classified, the difficulties encountered and the benefits of investigating and modelling the relationships between various sensor streams.

IV. Deploy

Deploy was an eighteen month demonstration project which was co-funded by the Irish Marine Institute and the Irish Environmental Protection Agency, running from April 2009 until October 2010. It demonstrates long-term deployment of a multi-sensor sensor monitoring system at five different sites on the River Lee. This system provided data which allows the temporal and spatial variations in water quality to be examined, and the investigation of issues associated with deployment of such a system. The demonstration sites of the Deploy project were chosen at different points along the River Lee catchment which were representative of varying conditions along the river. One station was deployed near the source at Gougane Barra, two stations were deployed in the Inniscarra reservoir, another station was deployed in the main channel of the river (Lee Road) and the fifth station was deployed in Cork City (Lee Maltings) where the river has entered the estuary, which is tidal and partially saline. This was the site chosen for this study and for the rest of this paper will be referred to as Lee Maltings.

V. Site

There were two main reasons for choosing the Lee Maltings site as the location of the study. Firstly this site is located at the Tyndall National Institute which forms part of the CLARITY research collaboration. The Tyndall National Institute provides resources such as network and power which allows the easy instrumentation of the site with a camera network.

Secondly it must be noted also that the Lee Maltings site represents a very interesting site to monitor along the River Lee but also very difficult. The site is positioned on the north channel of the River Lee at the Tyndall National Institute. It is located near the upper end of the estuary on a left hand bend of approximately 70 degrees. Water levels at the site are influenced by spillage from the Iniscarra dam and the site is also tidal with a tidal range of approximately 4m.

A. Camera

An AXIS PTZ Network camera was deployed overlooking the banks of the River Lee at the Tyndall Research Institute, Cork, Ireland [17]. It is controlled remotely from a desktop PC at Dublin City University (DCU). A software application was developed to interface with the camera every minute. Each minute it moves the camera to four different positions in order to save images from the camera at four different angles. These angles are as follows: pan-right-zoom ($ca - wall$), pan-left-zoom ($ca - trees$), pan-up-zoom ($ca - sky$) and finally a full zoom-in ($ca - wall$) on the water . The images from the camera are stored to a desktop PC at DCU for further analysis. Initially the application was interfacing with the camera at ten minute intervals. Following re-evaluation, it became apparent that a greater sampling frequency was required. Occasionally large changes in environmental conditions were evident from one sampled image to the next. Hence the sampling rate was changed to one minute intervals in order to capture a greater amount of interesting events at the site. The camera was fully deployed and linked up to the network at the Tyndall Research Institute from May 14 2008. Due to initial problems with camera positioning and stability, data is available for analysis from July 2008.

B. In-situ and visual sensing parameters

The in-situ parameters measured at the Lee Maltings site by the multi-sensor in-situ sensing system include conductivity, chlorophyll, temperature, dissolved oxygen and depth. The images from the camera were then analysed in order to highlight image features that could be used to complement the information being retrieved from the in-situ sensor network. Analysing outdoor data is challenging due to the wide-range of environmental conditions and their rapid changes. Varying river and weather conditions, extreme changes on lighting and reflections on the water are representative of some of the challenges presented. The assessment of a feature such as water colour which can be used to estimate water quality parameters such chlorophyll and turbidity is difficult if not impossible under these circumstances. Previous studies where cameras were used to estimate water colour were carried out using much in more specialised circumstances and not in the case where a camera is just placed in a building overlooking the water e.g. [11]. However the purpose of this work is to examine how we can use relatively low cost off-the-shelf webcam type devices for complementing in-situ sensor networks.

As previously outlined, the Lee Maltings site is tidal. Depth can give us an indication of a variety of conditions at the Lee Maltings site such as temperature, dissolved oxygen and conductivity readings. Figure 1 demonstrates the influence of depth on a variety of the in-situ sensor readings. The conductivity data also illustrates dilution in the River Lee due to dam releases from the Iniscarra reservoir.

Therefore the estimation of depth from the camera images is a really important indicator of conditions at the site. Our visual sensing system also undertook the detection of other image features such as objects floating on the water, boats, water turbulence etc. However it is really only the extraction of depth that can be linked up with the in-situ sensor readings which is the prime focus of this research. The detection of objects such as boats or floating objects is also an example of how a visual sensor can complement and enhance the use of an in-situ sensor network, as these are objects that cannot be detected or immediately detected by an in-situ sensor network. Thus the incorporation of a multi-modal sensor network for monitoring the marine environment leads to the detection of an increased number of events. Figure 2 demonstrates images where a boat and some material can be seen floating on the top of the water. However the focus of this paper will be on estimating water depth using a visual sensor at the Lee Maltings site. As well as providing an indication of current

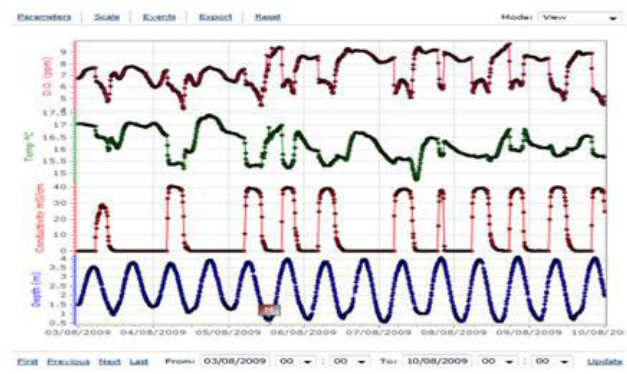

Fig. 1. The relationship between depth and various in-situ parameters at the Lee Maltings site. The conductivity data also illustrates dilution in the River Lee due to dam releases from the Iniscarra reservoir .

Fig. 2. A boat on the river and scum floating on the top of the water. Our visual sensing system can detect objects such as these floating on the top of the water

conditions at the site, continuous monitoring of water level is important for flood warnings and also for navigational and recreational safety.

VI. Estimation of water depth using a visual sensor

Changing depth is a feature of almost any inland waterway. If the water-land boundary is visible, visual imaging is a practical means for determining the water-level. As previously outlined the camera deployed overlooking the river at the Lee Maltings site, pans to four different angles every minute - pan-right-zoom ($ca-wall$), pan-left-zoom ($ca-trees$), pan-up-zoom ($ca-sky$) and finally a full zoom-in on the water ($ca-centre$). These four camera angles can be seen in Figure 3

In order to analyse the relationship between the sensor readings and features in the images, a software tool was developed to enable the visualisation of the images and the nearest in-situ sensor reading that corresponds to the time that image was captured (See Figure 4). From analysis of the images from $ca-wall$ along with the in-situ depth readings, it is apparent that certain features in the images become visible in a certain order as water levels are decreasing and can thus provide an estimation of water level at the site. This can be clearly seen in Figure 5. As the depth of the water decreases, the first feature to becomes visible is the appearance of rocks beneath the trees in the far left of the image (*feature 1 - rocks at trees*). The second feature to become visible is the appearance of rocks in the far right of the image (*feature 2 - rocks at far wall*). The third feature to become visible is the appearance of rocks in the near right of the image ((*feature 3 - rocks at near wall*)) and finally the final feature to become apparent that will indicate depth is the appearance of a small island in the middle of the water (*feature 4 - island*). Thus if each of these features can be accurately detected, then this can provide a very good indication of water levels at the site. Each of these features are used to delineate a certain type of water level e.g. the appearance of *feature 1* denotes water level 1, the subsequent appearance of *feature 2*, denotes water level 2, and so forth. Previous work [4] demonstrated the difficulty with accurately detecting the boundary of such a feature due to colour changes in the image. Thus an approach which allows the detection of such a feature or a series of features appears to be more robust to these conditions. The following section outlines the approach and results for the detection of each of these features.

VII. Methodology - Detection of Depth Features

The detection of these features in the image is far from trivial due to the huge lighting changes of the water. Sometimes the water can appear almost black mid-day due to reflections on the water and the appearance of rocks at each of the individual features is unclear even to the human eye. This therefore renders it difficult for an image processing algorithm to accurately detect the appearance of each of the features.

A. Data

As previously outlined an image is captured approximately every minute by the visual sensor network. This leads to over $25,000$ per week being captured for one camera angle alone. Thus for 4 camera angles, there are over $100,000$ images captured per week. For this study, one week of images from May 1-7 2009 was chosen due to the display of a number of events associated with changes in depth. This week of data was manually annotated four times over in order to have a set of ground-truth images for each of the features outlined above. For each feature, the images were annotated as follows - 1. feature present 2. no feature present 3. slight feature present. The third annotation is used where the annotation of the image is slightly ambiguous. The feature is at an intermediate stage of appearance whereby it is slightly present but not as apparently as when annotated as 'present'. This is due to intermediate changes in the water level. These images were given their own class so that it could be decided when carrying out classification whether it is better to have three-way classification for these features or two way-classification

whereby images where the feature is annotated as being slightly present are just classified simply as 'present' or 'not present'.

B. Image Analysis - Feature Sets

The Matlab image processing toolbox (Version R2009A) was used for processing images and extracting relevant image features. For each of the depth features to be detected (i.e. depth features 1-4) a set of image features were extracted from the images at the relevant points using the four sets of ground truth data. A variety of features were extracted including colour features such as *average hue, normalised hue histogram, average saturation, normalised saturation histogram, average value, normalised value histogram*, texture features such as *average entropy, normalised entropy histogram*, edge features such as *normalised edge histogram, number of pixels marked as an edge in feature area, percentage of pixels marks as an edge in mask area, number of pixels marked as an edge in mask area after a fill operation, percentage of pixels marked as an edge in mask area after a fill operation* and other features such as *average brightness, normalised brightness histogram, average luminance, normalised luminance histogram*. Various feature sets were examined for the detection of each depth feature.

These feature sets were then extracted for approximately 1200 (400 positive instances, 400 negative instances, 400 slightly present instances) random images from the ground truth dataset for each of the four depth features. Therefore approximately 4800 instances are analysed in this study. However it should be noted in the case of *feature 4- island*, there were not 400 hundred instances available in the dataset for the class *slightly present*. Therefore in this case any available instances were used.

These feature sets were then input into a Support Vector Machine (SVM) classifier which is detailed more below. After initial testing the feature sets outlined in Table I were the most successful for detecting each of the depth features outlined above. Thus these feature sets are used in the remainder of the study. The features in the first row of the table are the features used in all the features sets, with the features in the following rows outlining those that are specific to a particular feature set.

C. Support Vector Machine (SVM) - Classification

Support Vector Machines(SVMs) work well with large feature sets and after training they are very quick to classify new observations. With the correct parameters, they are known to work as well or better than most classification methods [18]. Initially a simple thresholding approach was attempted for classifying the presence of the various depth features. However this involved manually testing and setting thresholds for image features for each depth feature under investigation. With an SVM, features can be extracted for the location of interest in the image, formatted and input into an SVM for training or classification. This is a much more efficient and successful approach to classification of each of the depth features.

All Feature Sets	average hue, normalised hue histogram, average saturation, normalised saturation histogram, average value, normalised value histogram, average entropy, normalised entropy histogram
Feature Set 1	+ normalised edge histogram
Feature Set 2	+ normalised edge histogram, percentage of pixels marked as edge in feature area after a fill operation
Feature Set 3	+ percentage of pixels marked as edge in feature area after a fill operation
Feature Set 4	+ normalised edge histogram, average brightness, normalised brightness histogram
Feature Set 5	+ normalised edge histogram, normalised luminance histogram

TABLE I
VARIOUS FEATURE SETS EXAMINED IN THE STUDY. ALL FEATURE SETS CONTAIN THE FEATURES OUTLINED IN ROW 1 IN ADDITION TO THE FEATURES OUTLINED IN THEIR SPECIFIC ROWS.

Fig. 3. The angle of the images captured by the camera - labelled as follows - trees, wall, sky, centre

D. SVM Parameters

LibSVM [3], an integrated software for Support Vector Classification, is used in the Weka [4] data analysis environment for classification of the presence of the various depth features. Normalisation of features is carried out in the Weka environment and classification is carried out using two different kernel parameters - an RBF kernel and a linear kernel. The

[3] $http://www.csie.ntu.edu.tw/\ cjlin/libsvm/$
[4] $http://www.cs.waikato.ac.nz/ml/weka/$

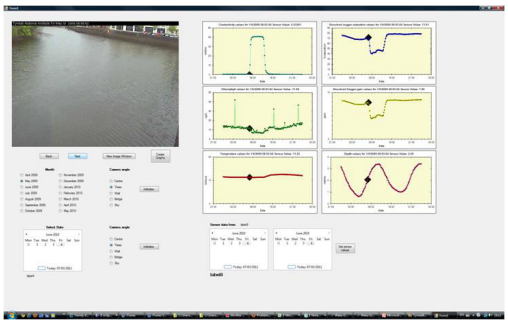

Fig. 4. Visual sensor analysis tool - enables the analysis of visual data alongside in-situ sensor readings in order to examine features and relationships between features and in-situ sensor data.

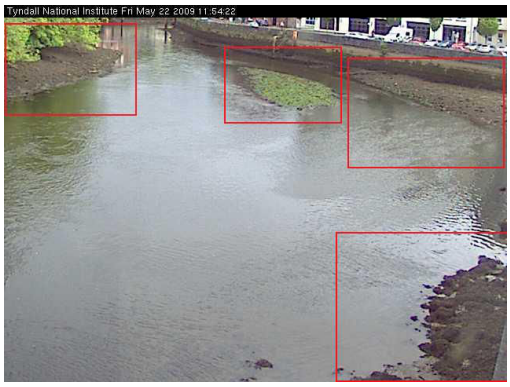

Fig. 5. The features highlighted in the image become visible in order with changing depth.

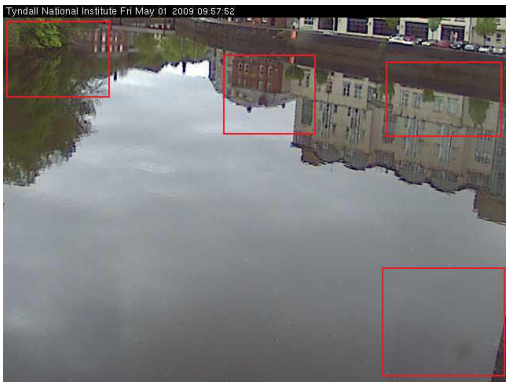

Fig. 6. Image where none of the depth features are present. It also demonstrates huge image processing challenges with reflections on water and changes in lighting.

RBF kernel can handle a nonlinear relationship between class labels and attributes, however it also may be the case that if the number of features is large, performance may not be improved from mapping data to a higher dimensional space. Default parameters in the Weka environment were used for both kernel parameters. However future work may involve optimising these. However satisfactory results were found without optimisation, thus this process was not carried out in the context of this study. Ideally optimisation of the (C, γ) space would be carried out for SVMs using an RBF kernel (*SVM-rbf*) and of the C space for SVMs with a linear kernel (*SVM-linear*). In this study a C values of 1 and a γ value of 0 was used for *SVM-rbf* and a C value of 1 is used for *SVM-linear*. Ten-fold cross validation is used for evaluation of the model. It is found that *SVM-linear* performs better and it is these results that are reported here. However optimisation of parameters for SVM-rbf could result in an increase in performance.

VIII. Results

As previously outlined, certain features in the images become visible in a certain order as water levels are decreasing and can thus provide an estimation of water level at the site. Thus if each of these features can be accurately detected, then this can provide a very good indication of conditions at the site. The following outlines our results in relation to detection of each of these depth features in the images. For each of the depth features, three types of evaluation were performed. As previously outlined, in certain images there is ambiguity in whether an image would be classified as having the feature present or not present as the visibility of the feature is at an intermediate stage due to changing water levels. Thus it is examined whether to have 3 way classification - a) feature present b) feature not present c) feature slightly present or a 2-way classification whereby these images where the feature is slightly present are classified as a) feature present b) feature not present. Therefore for the detection of each depth feature three sets of results are presented where the classifier is evaluated for three different classification scenarios:

- $3class$ - 3 way classification - rocks, no rocks, slight rocks.
- $2classN$ - 2 way classification negative - rocks, no rocks - where the images where rocks are slightly detected is regarded as a non detection.
- $2classP$ - 2 way classification positive - rocks, no rocks - where the images where the rocks are slightly detected is regarded as a detection.

Figure 6 shows an image where none of the features are present. It also demonstrates the challenges in terms of processing these images with huge reflections on the water such as that of the buildings seen in this image. There is huge variability among images in terms of reflections and lighting. In the following sections, for each classification approach the feature set that produced the most accurate results, the accuracy, the F-Measure and the area under an ROC Curve are reported. The F-Measure represents the harmonic mean of precision and recall and is calculated using the following formula 1:

$$F = \frac{2 * precision * recall}{precision + recall} \quad (1)$$

The area under an ROC curve represents the discrimination ability of the classifier, with a perfect result being 1. The F-Measures and ROC Areas reported here are the weighted average of the values for each of the individual classes in question for each of the classification scenarios. It should also be noted that for the *2classN* or *2classP* classification scenarios the datasets are merged between *slight rocks* and *no rocks* or *slight rocks* and *rocks*. This means there will be more instances of the merged class in the dataset presented to the classifer. However this represents a real-world scenario whereby the two datasets are not evenly distributed. Although the accuracy of the classifier may be working off a higher baseline, the ROC area presents the classifiers ability to distinguish between the two classes despite this biased distribution of positive and negative instances..

Fig. 7. a) Rocks at trees b) Slight rocks at trees

Class	FSet	Accuracy	F-Measure	ROC Area
$3class$	F2	81.7124	0.819	0.863
$2classN$	F1-F2	92.3525	0.922	0.901
$2classP$	F5	88.3624	0.882	0.858

TABLE II
RESULTS FOR EACH OF THE CLASSIFICATION SCENARIOS FOR THE DETECTION OF ROCKS AT THE TREES

Fig. 8. a) Rocks at far wall b) Slight rocks at far wall

Class	FSet	Accuracy	F-Measure	ROC Area
$3class$	F4	86.0349	0.862	0.895
$2classN$	F5	90.6899	0.906	0.888
$2classP$	F5	93.35	0.932	0.906

TABLE IV
RESULTS FOR EACH OF THE CLASSIFICATION SCENARIOS FOR THE DETECTION OF ROCKS AT THE FAR WALL

A. Depth Feature 1 - rocks at trees

Figure 7 (a) shows a sample image where rocks are appearing under the trees and Figure 7 (b) shows an image where rocks are only slightly appearing under the trees. From the results shown in Table II, it is clear that the most successful classification scenario is *2classN* where 2-way classification is used and images where rocks are only slightly appearing at the trees are classified as a negative detection. This has a high classification accuracy of 92.35% and and ROC Area of 0.901. Overall the classification accuracies are extremely satisfactory considering the difficult data that is being dealt with as can be seen in previous sample images. The results of the three way classification *3class* are very promising considering it is often difficult even for the human eye to distinguish between all three classes. This has a classification accuracy of 81.71% and an ROC Area of 0.863. From the confusion matrix outlined in Table III, it is clear that many of the images denoted as having no rocks are being classified as a slight detection of rocks and vice versa. This is not surprising since it can be difficult even for the human eye to distinguish between these classes. It is also consistent with the fact that doing two way classification whereby both these sets of images are considered to be the same class, produces the highest accuracy.

B. Depth Feature 2 - rocks at far wall

Figure 8 (a) shows a sample image where rocks are appearing at the far wall and Figure 8 (b) shows an image where rocks are only slightly appearing at that point in the image. From the results shown in Table IV, it is clear that the most successful classification scenario is *2classP* where 2-way classification is used and images where rocks are only slightly appearing at the trees are classified as a positive detection. The accuracy produced here is very high at 93.35%, with an ROC Area of 0.906. This is in contrast to the results for {emphdepth feature 1 - rocks at trees - where *2classN* produced more accurate results. Again the classification accuracies are satisfactory. There is a better accuracy for the three way classification *3class* than was achived for *depth feature 1*, with a classification accuracy of 86.03% and an ROC area of 0.895 which is extremely promising. From visual analysis of the images, it is clear that the visual distinction of the classes is more apparent than for *depth feature 1*. Therefore these result are consistent with this observation. From the confusion matrix in Table V it can also be seen that there is almost a consistent number of images of the class *slight rocks* that are incorrectly classified as *rocks* or *norocks*. Similarly there is also not much difference in the number of images of class *rocks* that are classified as *slight rocks* and the number of images of class *no rocks* that are classified as *slight rocks*.

C. Depth Feature 3 - rocks at near wall

Figure 9 (a) shows a sample image where rocks are appearing at the far wall and Figure 9 (b) shows an image where rocks are only slightly appearing at that point in the image. From analysis of the images, this feature is more visually more distinguishable than the other two depth features previously considered and this is apparent from the results presented in

	rocks	no rocks	slight rocks
rocks	340	14	47
no rocks	5	319	77
slight rocks	25	52	324

TABLE III
CONFUSION MATRIX FOR THE RESULTS OF THE *3class* CLASSIFICATION SCENARIO FOR DEPTH FEATURE 1 - ROCKS AT TREES

	rocks	no rocks	slight rocks
rocks	350	2	49
no rocks	10	339	52
slight rocks	28	27	346

TABLE V
CONFUSION MATRIX FOR THE RESULTS OF THE *3class* CLASSIFICATION SCENARIO FOR DEPTH FEATURE 2 - ROCKS AT FAR WALL

Fig. 9. a) Rocks ar near wall b) Slight rocks ar near wall

Fig. 10. a) island feature present b) island feature only slightly present

Class	FSet	Accuracy	F-Measure	ROC Area
$3class$	F5	92.8512	0.928	0.946
$2classN$	F2	98.3375	0.983	0.982
$2classP$	F5	93.5162	0.934	0.918

TABLE VI
RESULTS FOR EACH OF THE CLASSIFICATION SCENARIOS FOR THE DETECTION OF ROCKS AT THE NEAR WALL

Class	FSet	Accuracy	F-Measure	ROC Area
$3class$	F5	89.7155	0.896	0.917
$2classN$	F1-F2	92.4508	0.924	0.923
$2classP$	F5	94.3107	0.943	0.94

TABLE VIII
RESULTS FOR EACH OF THE CLASSIFICATION SCENARIOS FOR THE DETECTION OF THE ISLAND FEATURE

Table VI. The accuracy of the 3-way classification *3class* is extremely impressive at 92.85% and with an ROC Area of 0.946 which is higher than that achieved for the results of *3class* of any of the previous depth features considered. It is also interesting to note that from the confusion matrix, it can be seen that the main problem is the distinction between the *no rocks* and *slight rocks* classes. This is consistent with the results from *2classN* which demonstrates that when these two classes are merged that a classification accuracy of 98.34% and an ROC Area of 0.982 can be achieved which is an excellent result. When the classes *rocks* and *slightrocks* are merged the classification accuracy is lower at 93.52% and a ROC area of 0.918.

D. Depth Feature 4 - island

Figure 10 (a) shows a sample image where an island like feature can be seen in the middle of the water and Figure 10 (b) shows an image where this feature can only be slightly seen. Table VIII shows that the accuracy of the 3-way classification is quite good at approximately 89.76% and an ROC Area of 0.917. Similar to the accuracies produced for *depth feature 2 - rocks at far wall*, the *2classP* classification scenario produces the highest accuracy of 94.31% and a ROC area of 0.94. When the classes *no island* and *slightisland* are merged in *2classN* the accuracy is lower at 92.45% and a ROC area of 0.923 . From the confusion matrix produced by the *3class* classification scenario outlined in Table IX it is clear that the class *slight island* was as equally inclined to be classified inaccurately as class *island* as class *no island*. Also class *island* was classified incorrectly as *no island* more frequently than as class *slight island*, with a similar scenario for class *no island* where incorrect classifications were more as class *island* than *slight island*. It should be noted here that a limited number of instances of the class *slight island* were available to the dataset, which renders the 3-way classification results more impressive.

E. Discussion

From these it results it is clear that each of the four depth featured can be detected to a very high accuracy and the classifier has a high ability to distinguish between the classes. The most successful results were for *depth feature 3 - near wall rocks*. This is consistent with visual analysis of the images where it can be seen that this feature is far more distinguishable throughout an array of lighting conditions than each of the other features. Three different classification types were examined - 3-way classification - *3class* - where the ambiguous images in which it was difficult to decide if a feature was present or not due to intermediate changes in water level were given a class of their own for each of the features, 2-way classification - *2classN* - where these ambiguous images where the feature was slightly present were regarded as a non-detection of the feature and merged with the no rocks or no island class, and finally two-way classification - *2classP* where these images where the feature was slightly present was regarded as a positive detection and merged with the rocks or island class. For two of the features - *Feature 1 - rocks at trees* and *Feature 3 - rocks at near wall* the accuracies were

	rocks	no rocks	slight rocks
rocks	396	3	2
no rocks	5	351	45
slight rocks	7	24	370

TABLE VII
CONFUSION MATRIX FOR THE RESULTS OF THE *3class* CLASSIFICATION SCENARIO FOR DEPTH FEATURE 3 - ROCKS AT NEAR WALL

	island	no island	slight island
island	380	12	8
no island	22	363	16
slight island	18	18	77

TABLE IX
CONFUSION MATRIX FOR THE RESULTS OF THE *3class* CLASSIFICATION SCENARIO FOR DEPTH FEATURE 3 - ISLAND

better for *2classN* where these ambiguous images were merged with the *no rocks* class. In these scenarios it was also clear from the confusion matrices of the 3-way classification that the classifier found it most difficult to distinguish between the *rocks* and *no rocks* classes. For the other two depth features, *2classP* produced the highest accuracies. Overall accuracies were very satisfactory producing over 90% accuracy in many cases for the best 2-way classification and over 85% in the 3-way classification except for *depth feature 1- rocks at trees*. Other metrics such as ROC Area demonstrated the classifier's ability to distinguish between the classes with values of over 0.9 for the best 2-way classification and reaching 0.9 for three out of the four 3-way classifications. Although in many cases results between feature sets were very similar, it is apparent that feature sets 2 and 5 seem to produce the highest accuracies overall. From our analysis feature set 7 appeared to be the most unstable. Overall these results are very promising considering it can be difficult at times even to distinguish each of the features visually with reflection and rapid changes in lighting conditions at the site.

IX. Linking Parameters - Visual and In-Situ

Now that it has been determined that each of these depth features can be successfully detected, it is clear that our visual sensor can be used in order to provide an estimation of water level at the site. However as previously outlined our camera takes not just one image but four images every minute. It pans right - $ca - wall$(these are the images used in the preceding study), pans left - $ca - trees$, zooms in on the water - $ca - centre$ and pans up to the sky - $ca - sky$. Therefore each of these images can represent an individual sensing stream in our environmental monitoring network. However these streams need to be aligned in some way. Figure 11 shows images from $ca - trees$ and $ca - centre$. Clay or rocks begin appearing under the trees in *ca-trees* at the same time as the first *depth feature - rocks at trees* appears in $ca - wall$. In $ca - centre$ the rocks feature begins to appear at the same time as *depth feature - rocks at near wall* in $ca - wall$.

When delineating classes in the images from $ca - trees$ and $ca - centre$ we don't have the same situation as we had with images from $ca - wall$ where the appearance or disappearance of four different depth features denotes different water levels. However we can use this information in order to delineate classes in images from the two other camera angles. Each of the images are taken within 20 seconds of each other. Therefore using the image timestamps, we can annotate images from $ca - trees$ and $ca - centre$ according to our annotations for $ca - wall$ where the various water levels are more clearly distinguishable. We can then extract features from the images grouped into each class and attempt to use a classifier to delineate between the various classes. Thus we would have each of the camera angles representing a sensor stream measuring different classes of water level and this therefore provides redundancy in the network. This work forms part of our ongoing research efforts in combining multiple sensing modalities for marine environmental monitoring applications.

Fig. 11. a) camera angle panned towards trees - $ca - trees$ b) camera angle zoomed in on water - $ca - centre$

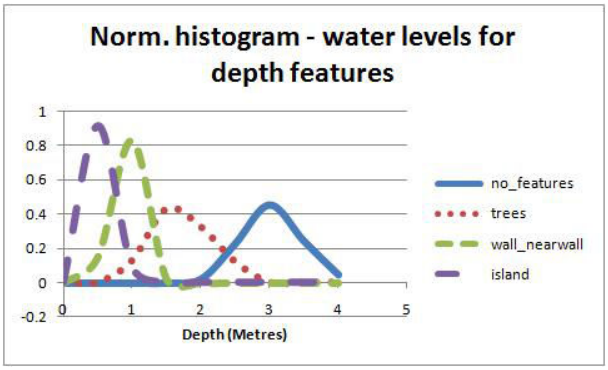

Fig. 12. Normalised histogram showing the distribution of water levels for the various depth features

It also needs to be examined how information from our visual sensors can be linked to the in-situ depth data. Figure 12 shows a normalised histogram showing the relationship between the appearance of the various depth features and water level readings from the Deploy water depth sensor. The curve representing *no-features* shows the normalised distribution of water depth values when there are no depth features present in the images from the training set. The curve entitled *trees* shows the normalised distribution of values when *depth feature 1 - rocks at trees* is present or slightly present **and** none of the other three depth features are present (i.e *depth feature 2 - rocks at far wall*, *rocks at near wall* and *island*). The curve *wall-nearwall* shows the normalised distribution of values when *depth feature 2 - rocks at far wall* **or** *depth feature 3 - rocks at near wall* is present **and** *depth feature 4 - island* is not present. Finally the curve *island*shows the normalised distribution of depth values when *depth feature 4 - island* is present. It is clear from this histogram that there is a clear distinction between the distribution of water depth values output by the depth sensor for the varying appearance of the four depth features.

X. Conclusion

In conclusion, it clear that a visual sensor can be used to complement the use of an in-situ sensor network in a river environment. In times of extreme events where an in-situ sensor may go offline or during times of failure due to stresses in the environment or failure of components, a visual sensor can provide an estimation of conditions at the site. The four depth features utilised in this study can also be extended to

include other types of features such as detecting the amount of wall that can be seen so a risk of flooding can be assessed. As previously outlined estimation of depth is extremely useful at the site since it can provide an indication of a variety of conditions. Other parameters that can be picked up by the visual sensing system include the detection objects floating on the water, weather conditions etc.

The aim of this study is to examine the effectiveness of incorporating a low cost off the shelf non specialised camera in an environmental sensing network. Thus there are also limitations to the use of the visual sensor in the manner that it is being examined in this study. For example it is difficult to pick up water colour through extreme changes in lighting conditions and reflections on the water. However it is clear that there are huge benefits to incorporating such a device in a network. It can provide low cost long term sensing that requires little or no maintenance. It can provide an estimation of conditions at the site, it has a wider spatial resolution than a single point sensing device and it can detect events that cannot always be detect by an in-situ sensor network. Such a device could be used hierarchically in a network whereby if it senses change in the site it can send a message to the more sophisticated nodes in the network to take a measurement and provide a precise measurement. This the efficiency and effectiveness of the network can be improved whereby there may be sensors that only have a limited number of samples before requiring maintenance.

Finally another objective of this work is to examine how such a multi-modal sensor network can be used to tackle data reliability. It may not be viable to deploy a multitude of sensors monitoring the same parameter and sensors are can be inherently unreliable when deployed in the marine environment. In the literature trust and reputation models have been used for monitoring the trust of in-situ sensor nodes where there is a multitude of homogenous sensor nodes monitoring the same parameter (e.g. temperature). The aim of this work is to adapt such a model [19] to be used in a scenario where nodes in the network are heterogenous and represent a variety of sensing modalities. This means that a visual sensing stream may be able to help determine the trust of an in-situ node in the network and determine if reported events are real. It may determine whether their are abnormalities associated with the readings from the sensor . If abnormalities are detected then an alert may be sent to the site manager to carry out maintenance and the data can be flagged as unreliable so that analysis is not affected.

ACKNOWLEDGMENT

Based on research funded by the Dept. of Communications, Marine & Natural Resources under the Strategy for Science, Technology and Innovation (2006-2013) and by Science Foundation Ireland under grant 07/CE/I1147.

REFERENCES

[1] D. Diamond, S. Coyle, S. Scarmagnani, and J. Hayes, "Wireless sensor networks and chemo-/biosensing," *Chemical Reviews*, vol. 108, no. 2, pp. 652–679, 2008.

[2] D. Diamond, K. T. Lau, S. Brady, and J. Cleary, "Integration of analytical measurements and wireless communications: Current issues and future strategies,," *Talanta*, vol. 75, pp. 606–612, 2008.

[3] P. Alexander and R. Holman, "Quantitative analysis of nearshore morphological variability based on video imaging," *Marine Geology*, vol. 208, no. 1, pp. 101–111, 2004.

[4] E. O'Connor, C. . Conaire, A. Smeaton, N. O'Connor, and D. Diamond, "River water-level estimation using visual sensing," in *EuroSSC 09 4th European Conference on Smart Sensing and Context*, 2009.

[5] M. Davidson, M. V. Koningsveld, A. de Kruif, J. Rawson, R. Holman, A. Lamberti, R. Medina, A. Kroon, and S. Aarninkhof, "The coastview project: Developing video-derived coastal state indicators in support of coastal zone management," *Coastal Engineering*, vol. 54, no. 6-7, pp. 463 – 475, 2007.

[6] R. Holman and J. Stanley, "The history and technical capabilities of argus," *Coastal Engineering*, vol. 54, no. 6-7, pp. 477 – 491, 2007.

[7] S. G. J. Aarninkhof, I. L. Turner, T. D. T. Dronkers, M. Caljouw, and L. Nipius, "Nearshore subtidal bathymetry from time-exposure video images," *Geophysical Research*, vol. 110 (C6), 2005, art. No. C06011 JUN 23 2005.

[8] C. C. Chickadel, R. A. Holman, and M. H. Freilich, "An optical technique for the measurement of longshore currents," *Geophysical Research*, vol. 108 (C11), p. 3364, 2003.

[9] R. Holman and C. Chickadel, "Optical remote sensing estimates of the incident wave angle field during ncex," in *International Conference on Coastal Engineering*, 19-24 Sept 2004.

[10] C. Chickadel and R. Holman, "Measuring longshore current with video techniques," *AGU Fall Meeting Abstracts*, pp. C9+, Dec. 2001.

[11] L. Goddijn-Murphy, D. Dailloux, M. White, and D. Bowers, "Fundamentals of in situ digital camera methodology for water quality monitoring of coast and ocean," *Sensors*, vol. 9, no. 7, pp. 5825–5843, 2009. [Online]. Available: http://www.mdpi.com/1424-8220/9/7/5825

[12] M. Iwahashi, S. Udomsiri, Y. Mai, and S. Fukuma, "Water level detection for river surveillance utilizing jp2k wavelet transform," in *IEEE Asia Pacific Conference on Circuits and Systems, 2006. APCCAS 2006.*, 2006.

[13] S. Ahmadian, T. Ko, S. Coe, M. Rahimi, S. Soatto, M. Hamilton, and D. Estrin, "A vision system to infer avian nesting behavior," University of California, Los Angeles, Tech. Rep., 2007.

[14] E. A. Graham, E. M. Yuen, G. F. Robertson, W. J. Kaiser, M. P. Hamilton, and P. W. Rundel, "Budburst and leaf area expansion measured with a novel mobile camera system and simple color thresholding," *Environmental and Experimental Botany*, vol. 65, no. 2-3, pp. 238 – 244, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/B6T66-4TPPF2G-1/2/bd08f949e37262d8ca2265021f86042b

[15] A. Richardson, J. Jenkins, B. Braswell, D. Hollinger, S. Ollinger, , and M. Smith, "Use of digital webcam images to track spring green-up in a deciduous broadleaf forest," *Ecosystem Ecology*, 2007.

[16] E. O'Connor, J. Hayes, C. O'Conaire, A. Smeaton, N. O'Connor, and D. Diamond, "Image processing for smart browsing of oceancolor data products and subsequent incorporation into a multi-modal sensing framework," in *In RSPSoc Remote Sensing and Photogrammetry Society Annual Conference with Irish Earth Observation Symposium 2010 : 1-3 September 2010, Cork, Ireland.*, 2010.

[17] E. O'Connor, A. F. Smeaton, N. E. O'Connor, and D. Diamond, "Integrating multiple sensor modalities for environmental monitoring of marine locations," in *SenSys '08: Proceedings of the 6th ACM conference on Embedded network sensor systems*. New York, NY, USA: ACM, 2008, pp. 405–406.

[18] T. Segaran, *Programming collective intelligence*. O'Reilly, 2007.

[19] S. Ganeriwal, L. K. Balzano, and M. B. Srivastava, "Reputation-based framework for high integrity sensor networks," *ACM Transactions on Sensor Networks*, vol. 4, 2008.