Automatic methodology for mapping of coastal zones in video sequences

Natalia V. Revollo [b,c,d,*,1], Claudio A. Delrieux [c], Gerardo M.E. Perillo [a,e]

[a] Instituto Argentino de Oceanografía, CONICET, Florida, 8000 Bahía Blanca, Argentina
[b] Facultad de Ingeniería, UNJu, Gorriti 237, Jujuy, Argentina
[c] Dpto. de Ingeniería Eléctrica y de Computadoras, UNS-CONICET, Av. Alem 1253, Bahía Blanca, Argentina
[d] Instituto de Investigaciones en Ingeniería Eléctrica, UNS-CONICET, Av. Alem 1253, Bahía Blanca, Argentina
[e] Departamento de Geología, UNS, San Juan 670, Bahía Blanca, Argentina

ARTICLE INFO

ABSTRACT

Gathering precise and detailed geomorphological and dynamic information of coastal processes is increasingly required for environmental studies and coastal management policies as well. Traditional methods for in situ measurements, or remote sensing monitoring by satellites or airbone imagery, impose limitations and tradeoffs between image quality, operational costs, availability, and negative environmental effects. These limitations and tradeoffs restrict the kind of environmental studies that can be undertaken, specifically when a high spatial and temporal resolution is required over wide geographical areas. In the last decades, video monitoring systems have demonstrated to be a cost-effective alternative for this and other similar purposes. Notwithstanding that, video processing is not fully mature in the context of environmental monitoring in general, and, thus, most of the past and current efforts have been developed in an ad hoc basis. This has the drawback that most available solutions are hardly useful in contexts different from their original setup. In this work we develop an autonomous application for geographic feature extraction and recognition in coastal videos. Specifically, we address the classification and feature measurement of multiple beach zones, a topic addressed to a lesser extent by other projects. The system is designed to be deployed in inexpensive, off-the-shelf hardware, and open source software development frameworks, in a way such that the results can be easily replicated by other research groups. The initial setup and calibration requires very simple supervision, thus allowing the system to be used in a variety of coastal environments.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Satellite and airborne imagery and, more recently, in situ camera placement are among the preferred non-invasive monitoring techniques for obtaining qualitative and quantitative data. The information provided by these techniques is mainly used for decision making and establishing policies in coastal environments. Specially, beaches are unstable environments, which have continuous changes due to different phenomena including waves, winds and anthropic activities. The assessment of changes in these environments should be studied permanently, requiring continuous monitoring of the coast in a wide spatio-temporal scale (range from meters to kilometers and hours to weeks) keeping both the spatial and temporal resolution as high as possible (Holman and Stanley, 2007).

Detailed measurements of waves, beach profiles, currents or sediment transport are traditionally performed through in situ field studies. In most cases, traditional methods involve expensive and cumbersome transportation and installation of equipment, and the direct intervention of specialists in many fields. In addition, to be meaningful, these studies must be repeated periodically to measure the highly varying environments. For this reason, these campaigns have complex logistics associated and high operative costs that vary according to the monitoring site.

On the other hand, satellite or airbone imagery allows a wide coverage size of the study area, with the advantage of being able to sense regions with difficult access conditions. However, the use of these techniques is limited by their high cost in relation to the required image quality (Yatabe and Fabbri, 1986; Schowengerdt, 1997; Girard and Girard, 1999; Lillesand and Kiefer, 2000; Jensen, 2000). Also, in satellite imagery, the revisit time may not be adequate for understanding short-time phenomena (i.e., tidal variations, changes in wave conditions, storms events, etc. require at least one sample per hour for a reasonable monitoring, far beyond the possibilities of current high spatial resolution satellite constellations).

In the last decade, in situ installation of autonomous monitoring systems become an alternative in terms of cost and performance. Installation of cameras in fixed positions offers an ever increasing spatial and temporal resolution, which results in constant improvements in cost-
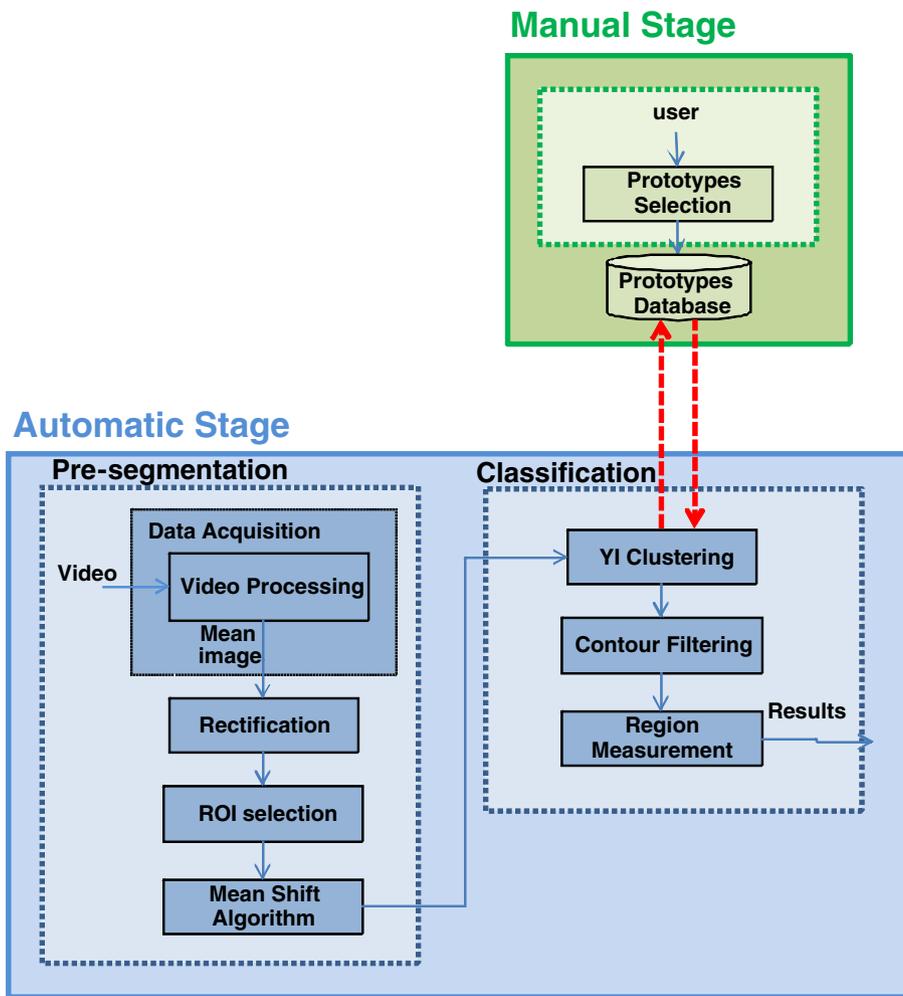
Manual Stage

Automatic Stage

Fig. 1. Manual and Automatic stages for mapping of coastal zones in video sequences.

effectiveness due to the technological breakthroughs associated to these devices. For this reason, it is already clear that autonomous video-based monitoring is a mainstream technology in environmental studies.

There are currently several research projects related with coastal video monitoring distributed worldwide. The Argus project is the pioneer in optical remote sensing of coastal management. The project was developed by the Coastal Imaging Lab (CIL) at the Oregon State University. During the 1980s, CIL began using time-lapse video measurements of wave run-up as a diagnostic method for sampling infragravity edge waves (Holman, 1981). Traditional measurement methods were replaced by optical sensors, allowing to locate many features such as submerged sand bars and rip current channels, which are not always visible by in situ techniques (Lippmann and Holman, 1989). The coastal information supplied by the video monitoring techniques enables the quantification of shoreline evolution and beach width, erosional and accretional sediment volumes at the intertidal beach, subtidal beach bathymetry, wave run-up and coastal state indicators with a high resolution in time through assimilation of computation models.

The project INDIA (Inlet Dynamics Initiative: Algarve) studies the Barra Nova Inlet in Portugal (Williams et al., 1998). INDIA involves a multi-Institute, multi-disciplinary study of the interacting hydrodynamics and morphodynamics occurring at tidal inlet entrances and along adjacent coastlines. Two main aims are followed by the project, namely the improvement in understanding complex interactive coastal processes, and the validation of numerical models. For this, the main activities are the monitoring and measurement of the surf zone

hydrodynamics using a video system and in situ observations of bedforms in the surf zone and offshore using a mobile instrument platform, which includes a video camera.

The Hazaki Oceanographical Research Station (HORS) developed a research facility for field measurements of various phenomena in the nearshore zone. HORS is located on the Hasaki coast of Japan, facing the Pacific Ocean. This study tries to show spatial and temporal variation of the surf zone hydrodynamics using a moored video system (Takewaka and Nakamura, 2001).

HORUS is a project for coastal monitoring using a video-based system created to contribute to the research and management of the environment. It is able to continuously measure changes in various natural areas and offers large spatio-temporal resolution (Osorio et al., 2012).

Cam-era project,[2] developed by National Institute of Water and Atmospheric Research, consists of a number of remote sites that monitor the coast of New Zealand. This has developed a set of tools for shoreline detection and beach width measurement to understand the beach changes for future shoreline management plans.

Another project for video monitoring is SIRENA,[3] a low cost and open source software. It was designed and developed by the Mediterranean Institute for Advanced Studies (IMEDEA, CSIC-UIB, Spain). SIRENA is composed of an in-field node and a central server that are remotely

---

connected. SIRENA provides the followings products: median image, variance image, snapshot, time stacks and pixels statistics.

Many of the above projects share a set of typical processing techniques, including timex or time exposure images, variance images, daily time exposure images and time-stack images. Timex images are a useful tool to detect shorelines (Kroon et al., 2007), submerged sand bar topography (Lippmann and Holman, 1989; Van Enckevort and Ruessink, 2001), and intertidal beach profile (Plant and Holman, 1997), among others. Variance images identify the location of the most variable zones of the beach compared with those that had less modifications.

Early projects began collecting and processing only gray scale images (Aarninkhof and Holman, 1999; Davidson et al., 2007). Simple information as shoreline location was first obtained using visual inspection of these images as an initial proxy, such as the 'Shore-Line

Intensity Maximum' (SLIM) model (Plant and Holman, 1997; Madsen and Plant, 2001). The SLIM model was highly robust for reflective beaches, but its features are often less clear for dissipative beaches.

Further improvement of shoreline detection used an analysis of the features in the correlogram of the cross-shore intensity and variance profile (Aarninkhof and Holman, 1999) or spatial gradients in intensity levels in rectified images (Davidson et al., 2007). However, the poor chromatic discrimination of gray-scale images was a limitation in more complex applications.

In order to increase the video monitoring capabilities, with the advent of better camera sensors, the subsequent developments started to include full color images. The technique described by Aarninkhof (2003) aims to segment a shoreline feature from a timex image, on the basis of distinctive image intensity features, sampled across subaqueous and sub-aerial beach. It uses Hue-Saturation-Value (HSV)
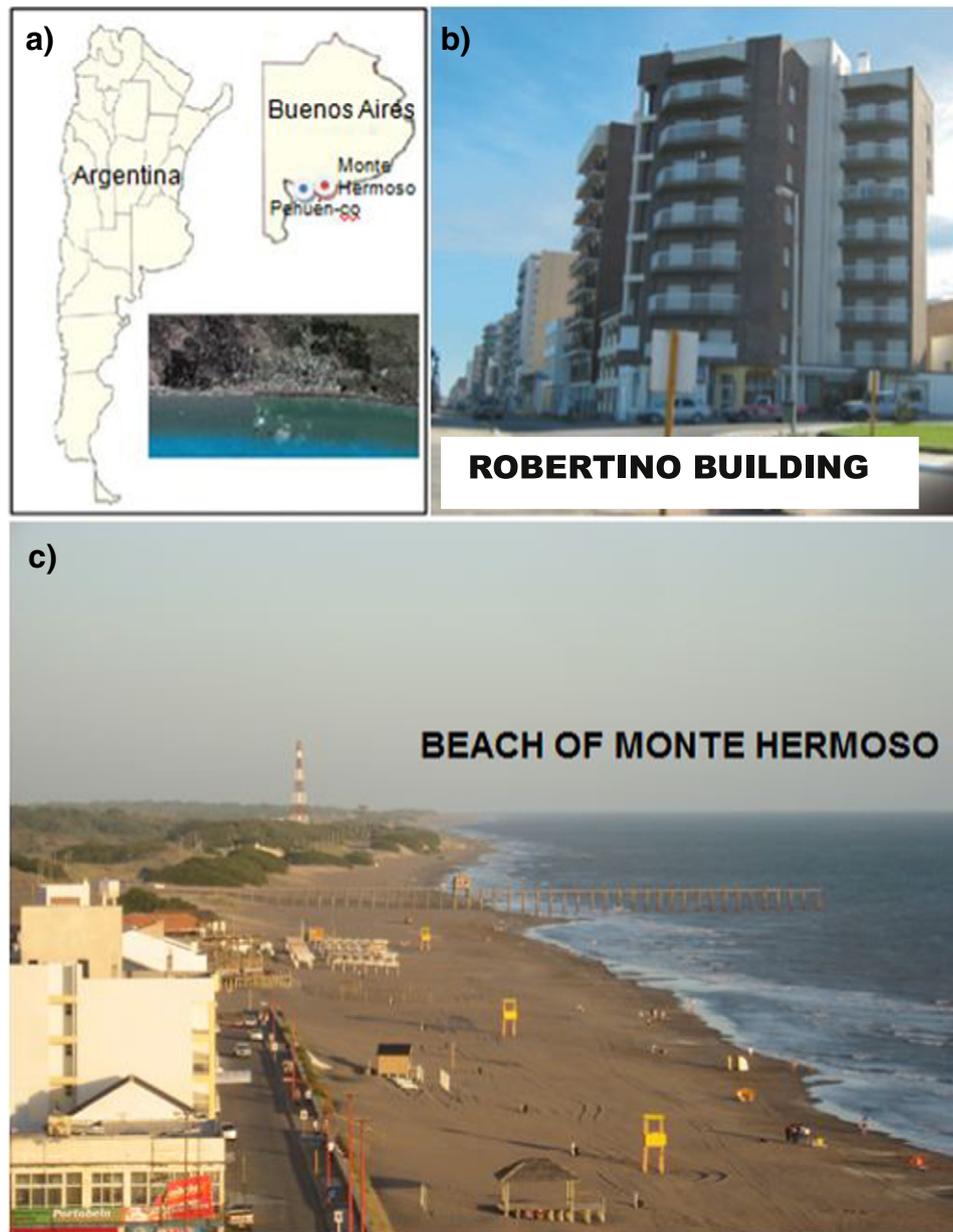
Fig. 2. Study area. a) Southern coast of the Buenos Aires Province, Argentina and location of Monte Hermoso Beach. b) Robertino Building. c) View of the beach from the top of the Robertino Building.

Fig. 3. Average image of a 5 min burst (4500 frames) taken at the Monte Hermoso Beach February 26, 2010.

color space, taking advantage of the split between chromaticity (HS) and value (V). In this way, the pixels that correspond to the different zones can be clustered together according to their chromaticity and luminance.

Quartel et al. (2006) presented a semi-automatic methodology to extract intertidal morphologic features in low tide from video images using object oriented image analysis. The classification algorithm uses the Maximum likelihood classification (MLC) and it considers the classes: dry sand, wet, sand and water of an ARGUS image.

Another approach consider an automatic algorithm for segmentation of coastal images (Hoonhout et al., 2015). The algorithm uses features with information of color, texture and visual appearance instead of color ones. With these features, a machine learning framework is used to train a model with ability to identify regions containing major classes such as water, sand, vegetation, sky and objects.

In this work, we present the development of an automatic, video-based geographic feature extraction and recognition system designed specifically for autonomous monitoring in beach environments. The system requires low-cost off-the-shelf hardware, was derived using open source software development frameworks, and it includes video

processing algorithms specifically developed for the determination of coastal features and their dynamics. The results provide a very precise and accurate monitoring of coastal zones (i.e., each of the possible beach segments that can be observed in a marine littoral). Features that makes our system different and more advanced than other previously presented.

Our initial aim is to perform accurate spatio-temporal determinations of the dry and wet beach, and breaker zones, since those are the most relevant for the geomorphological studies undertaken within our research Institute. Therefore, the present paper is focused in achieving simultaneous classification and measurement of multiple beach zones by means of an autonomous monitoring system. However, other similar applications can be tackled with no significant programming effort.

2. Methodology

2.1. Data acquisition

Digital videos are taken with a specially developed video monitoring station. The most simple station consists of a video camera and a conventional laptop that computes a video acquisition and a processing pipeline (Fig. 1). The testing of these algorithms was made in Monte Hermoso Beach on the southern coast of the Buenos Aires Province, Argentina (Fig. 2a). Monte Hermoso Beach is located along a stretch of the coast with a W-E general trend facing south. Therefore, the sun makes its path parallel and over the continent with respect to the camera. In the summer, sunrise appears at the horizon to the left of the area covered by the camera. The monitoring station was located in a fixed position at an elevation of 30 m on top of a building (Fig. 2b). The field of view covers all the different zones in a selected area, having a panoramic view of the beach (Fig. 2c). The beach is mainly dissipative with a relatively low slope (2° in average), (Huamantinco Cisneros, 2012). Therefore, in our particular situation, image rectification without an elevation model will incur in negligible errors. Further validation (see Section 3.1) confirms this. In situations with higher slope, a DEM would be required, and standard GIS procedures for orthorectification should be applied. The system captures bursts of 5 min of duration (15 Hz sampling rate), taken every 15 min, during daylight time. Videos are acquired in 640 × 480 pixels resolution, RGB color, in compressed MPEG format. A mean image which corresponds to the average of a 5 min burst (4500 images in total) is computed and stored (Fig. 3). For demostration of the method, we took several video captures during different months of the year representing varied meteorological, illumination and oceanographic conditions.
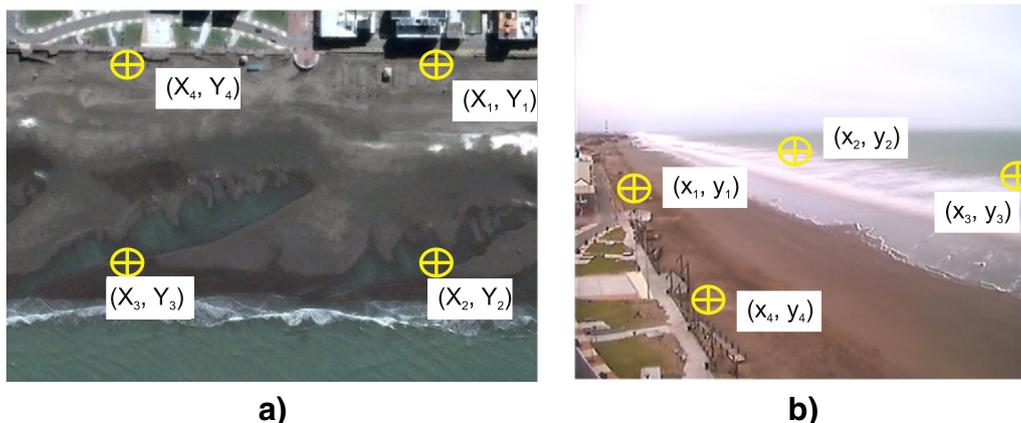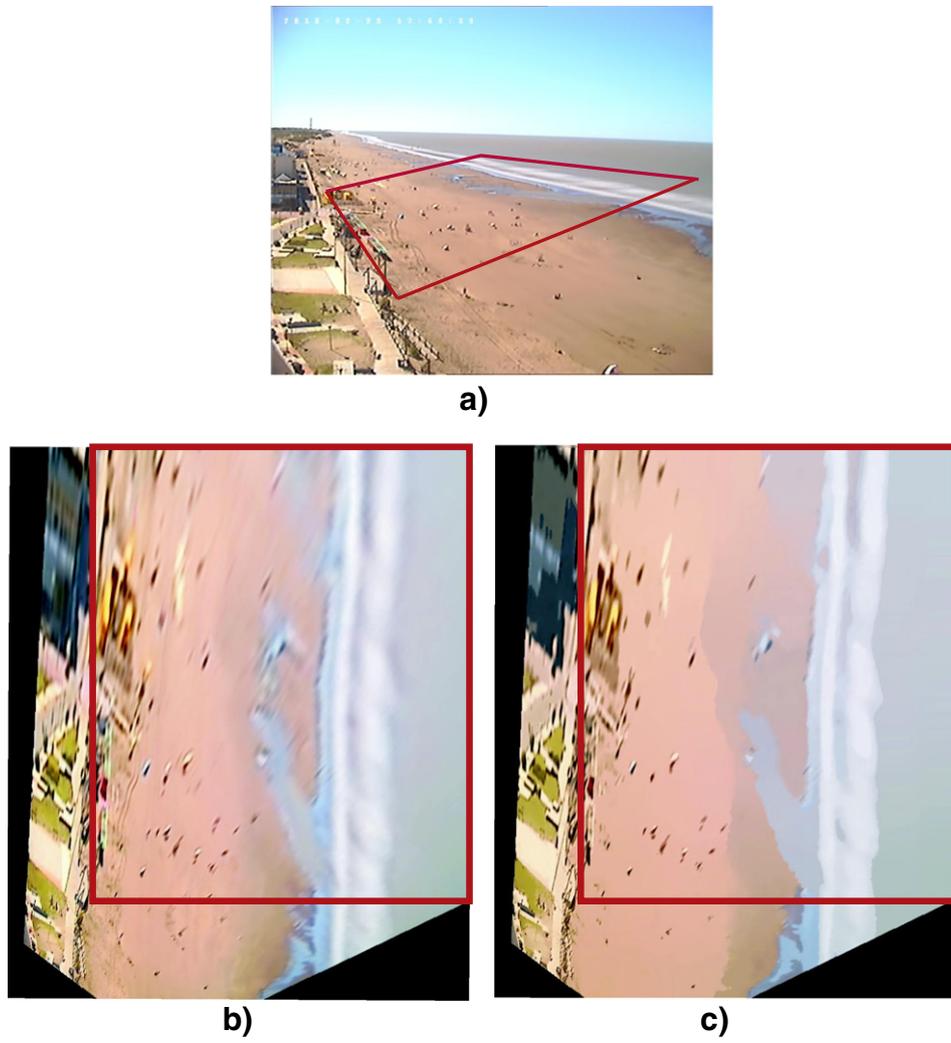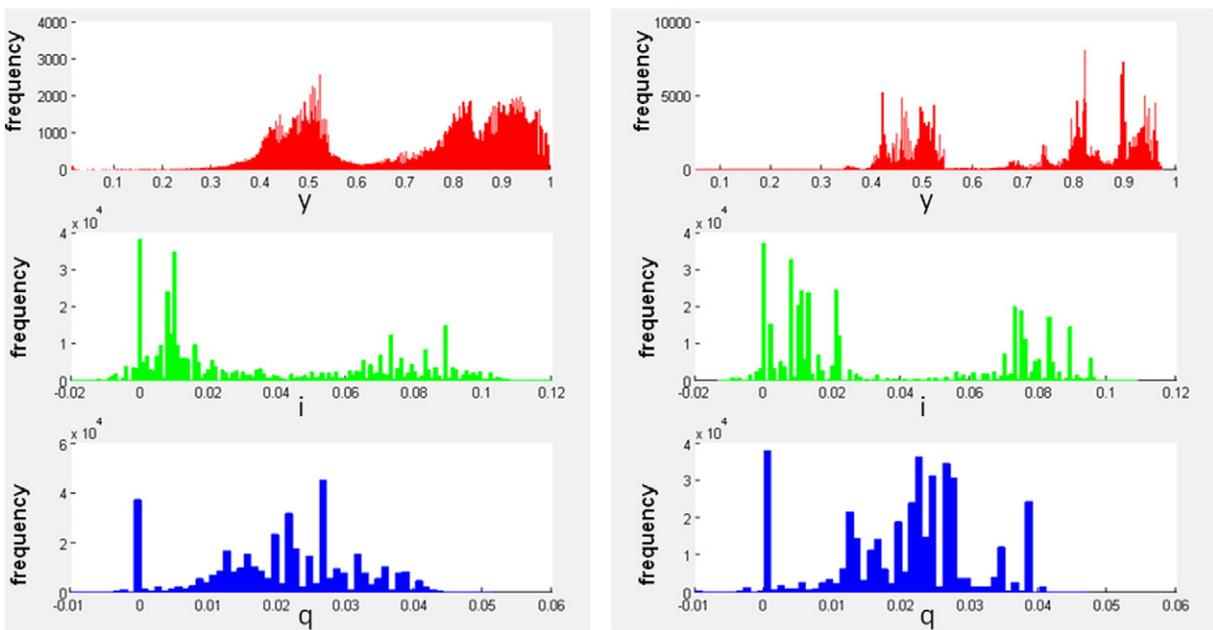

Fig. 4. Image rectification. a) GPS-established control points, and b) their respective pixel coordinates in the camera view.

Fig. 5. Pre-segmentation. a) ROI. b) A two-dimensional projective transformation was used to reproject the oblique image into a zenital one. c) ROI pre-segmented with the Mean Shift algorithm.

Fig. 6. Histograms corresponding to Fig. 5a and b (before and after applying mean shift, respectively). It can be appreciated that luminances (Y channel) and chromaticities (I and Q channels) clusterize in a few local modes, meaning that the intra-class color variances are reduced.
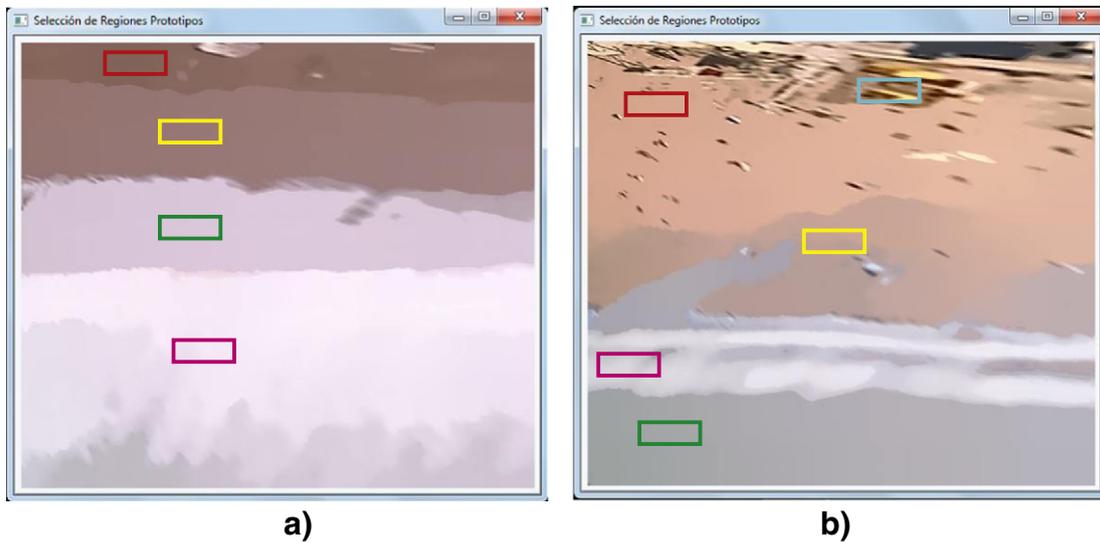
Fig. 7. Training areas: red (dry beach), yellow (wet beach), green (swash zone and water), magenta (breaker zone), and cyan (shadow). a) Monte Hermoso, February 26, 2010. b) Monte Hermoso, July 09, 2009.
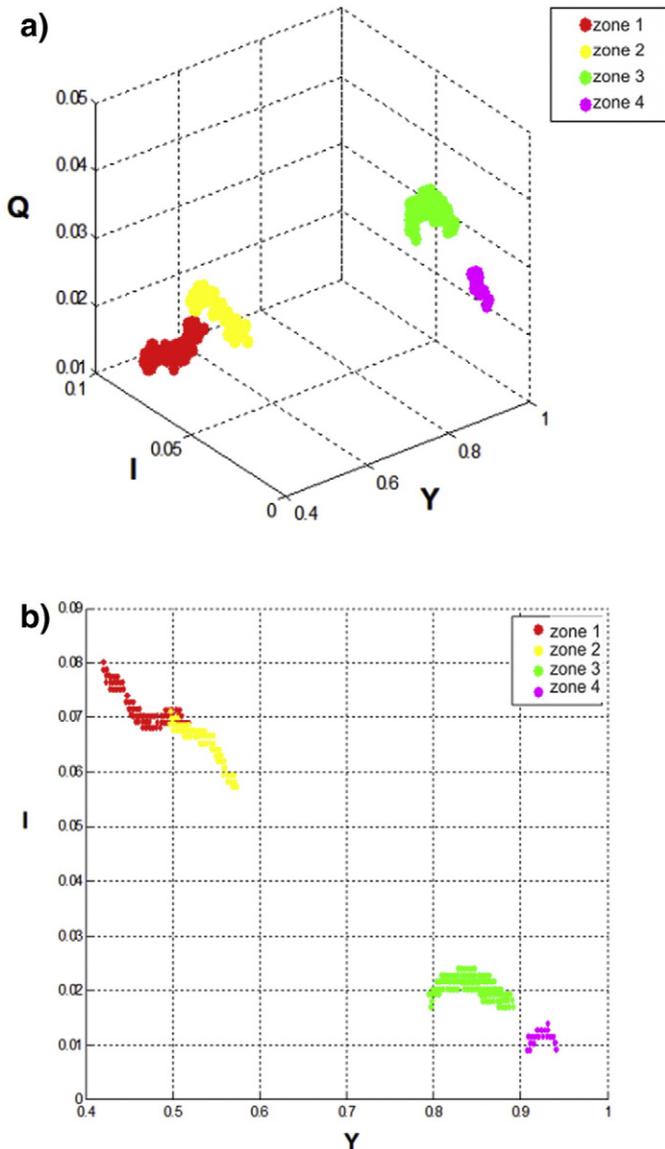

Fig. 8. Distribution of the pixel color in each training area a) in *YIQ* color space, and b) in *Y* vs. *I*.

2.2. Image rectification

To produce zenital views, a rectification procedure is applied on the oblique video frames (Perez-Muñoz et al., 2013). This has also the advantage of producing images where the spatial resolution is constant, therefore simplifying distance and area measurement (Szeliski, 2010). For this, a two-dimensional projective transformation was used to reproject the oblique mean images into zenital views. We briefly describe the method to estimate the projection matrix. For further details see (Faugeras, 1993; Hartley and Zisserman, 2004). Given a set of points with coordinates (x, y) in the reference system and their corresponding
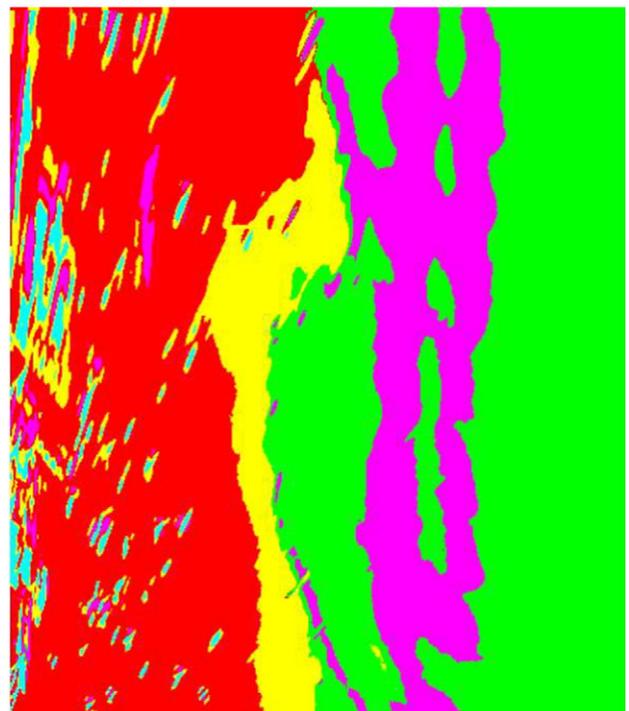

Fig. 9. Classification of the beach zones. Dry zone (red), wet zone (yellow), breaker zone (magenta), swash zone, water (green) and shadows (cyan).

coordinates (x', y') in an arbitrary system (in this case, the camera view), the projection among both systems can be computed using eight parameters (a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2) as follows:

$$x = \frac{a_1x' + b_1y' + c_1}{a_3x' + b_3y' + 1} \quad y = \frac{a_2x' + b_2y' + c_2}{a_3x' + b_3y' + 1} \tag{1}$$

Four noncollinear control points in both systems are required to estimate the projection parameters (the minimal solution). The accuracy of the reprojection depends on the precise location of the points. Initially, four coordinates (x_i, y_i) ($i = 1.0.4$) were obtained using a GPS RTK Sokkia Radian with an accuracy of 3 cm horizontal and 5 cm vertical. With the camera turned on, and pointing to the target area, the corresponding pixels (x_i, y_i) are marked on the screen. For the specific purposes of our study, the ground points form a rectangle of sides about 230 m parallel to the shoreline, and 160 m orthogonal to the shoreline (see Fig. 4). The procedure has to be undertaken during low tide. The rectification algorithm, using the four ground point and the four image pixels, is able to transform the oblique video frames into a zenital view in which there is no perspective distortion of the spatial resolution. These frames are the basis of the further processing steps (Fig. 5a, b).

2.3. Pre-segmentation

The classification algorithm will be applied only to the region of interest (ROI) (i.e, the region enclosed by the four control points) to
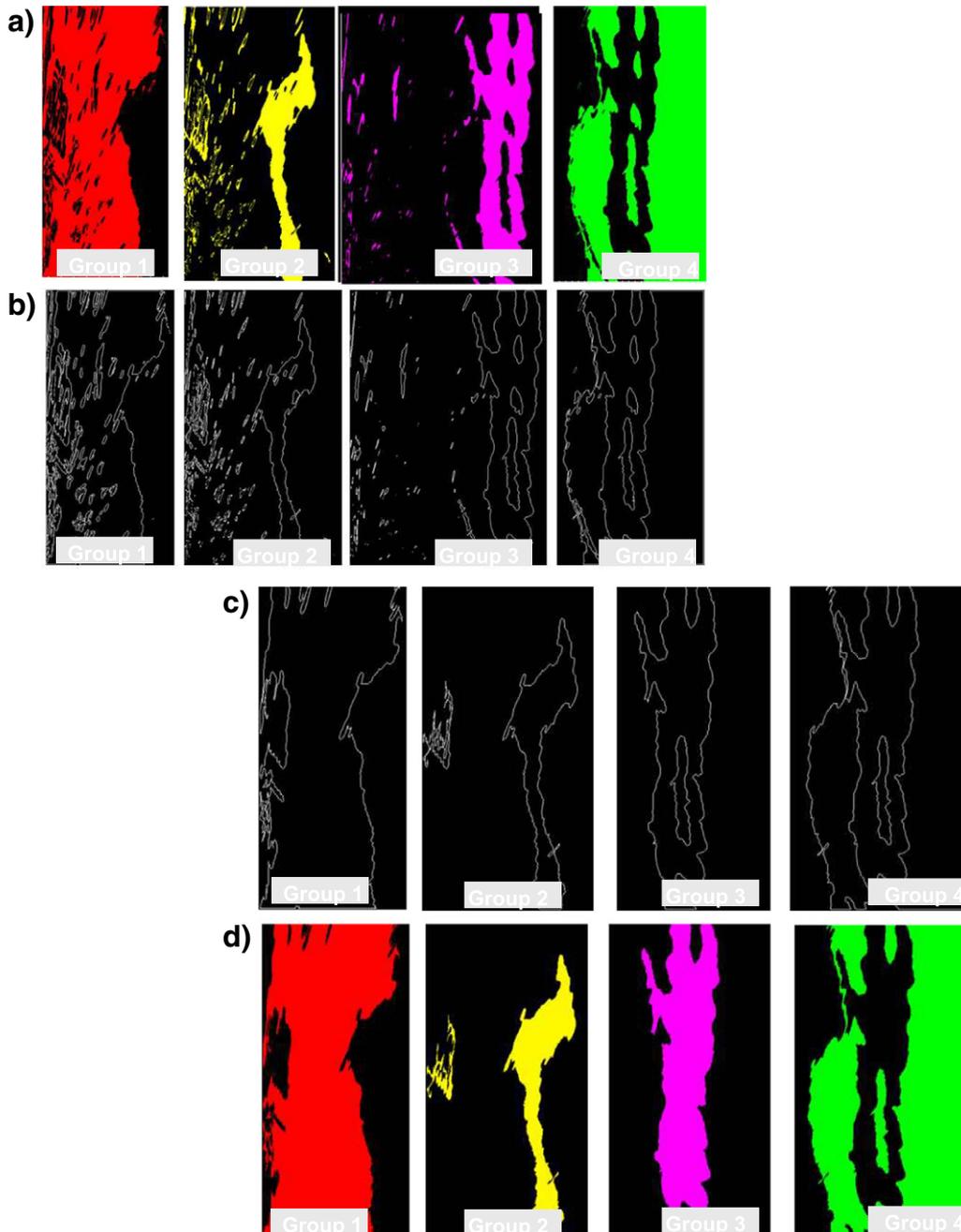

Fig. 10. Correction of the segmented image. a) Decomposition by groups of segmented image. b) Edge estimation. c) Filtering of the edges by size. d) Final segmentation.

avoid unnecessary processing time (Fig. 5b). Within the ROI, the color values in YIQ color space are clustered together using the Mean Shift algorithm (Fukunaga and Hostetler, 1975; Cheng, 1995; Comaniciu and Peter, 2002).

The Mean Shift algorithm is a widely used mode-seeking method, mostly used as a pre-segmentation method in computer vision. Basically it finds local maxima (modes) of an unknown density function (in our case, pixel distributions in color space), which reduces the color variance within classes and therefore enables more crisp region segmentation.

Given an initial population estimation x and a hypothesized kernel distribution K (usually Gaussian), the method iterates the determination of the weight of the neighbor distribution members to re-estimate x. The weighed mean of the density distribution within the window determined by K is given by:

$$m(x) = \frac{\sum_{x_i \in N(x)} K(x_i - x) x_i}{\sum_{x_i \in N(x)} K(x_i - x)}, \quad (2)$$

$$K(x_i - x) = e^{c\|x_i - x\|^2}, \quad (3)$$

where $N(x)$ denotes the neighbors of x for which the kernel function is nonzero. In every iteration, $m(x) - x$ is the hypothesized shift of the mean (and thus the name of the method), and therefore m is assigned to $m(x)$ and a new iteration is performed until the shift is below a convergence threshold.

The main reason for using YIQ instead of HSV color space is that, even though it splits luminance and chromaticity (as HSV does), the chromaticity space IQ is "metric" (vector space), while HS is not. In other words, distances in HS subspace has no clear meaning since H is an angle and S a distance. In this way, clustering-based segmentation algorithms in HSV space are unstable (with S close to zero, as is our case, tiny H variations may assign wrong classes to pixels). Also, since H is periodic, there is no exact formulation about neighborhood that may apply to the mean shift method. Other vectorial color spaces (Lab, Luv, L*a*b*, L*u*v*) were tested, yielding similar results as the simpler YIQ.

The result of this processing step is a pre-segmentation that can be appreciated in Fig. 5c. The color variance within the different regions is decreased. However, the Mean Shift algorithm only reduces the dispersion in the distribution of intensities, without generating the association of elements into classes. Specifically, in Fig. 6, the histogram of the mean image has peaks (modes) in the three channels of the YIQ space. After this pre-segmentation, the process of clustering pixels together according to their color properties is more robust.

2.4. Classification

Our classification scheme consists of a (supervised) training stage that must be performed only once, and a final (unsupervised) classification stage that is able to perform in the open. The purpose of these processing stages is to assign every pixel in the ROI to a specific class. The classes are described by a feature vector (or prototype), obtained only once during the calibration of the system. All the pixels in the same class determine a region in the image, which ideally should be coincident to the portion of the image where a specific beach zone is projected.

For this training, an expert selects small training rectangular areas in the different beach zones (dry, wet, breaker, swash zone and water), whose respective pixel colors will be used for establishing classification prototypes. In addition to the four mentioned zones, building shadows projected over the beach are considered as a fifth zone. The chosen training areas of the different zones are shown in Fig. 7, shadows of buildings in cyan, wet zone in yellow, dry zone in red, breaker zone in

magenta and swash zone and water in green. In Fig. 7b, the five classes can be seen, while in Fig. 7a, shadows are not present.

The distribution of the pixels' color within each training area in YIQ space (Fig. 8a) and its projection over the plane YI plane (Fig. 8b) exhibit a neat separation within areas. For this reason, the mean color of every zone can be taken as an appropriate prototype. Also, the distributions in the Y vs. I projections cluster together better than with the others projections (Fig. 8). For this reason, in the classification stage we only consider the Y vs. I of the pixels and prototypes.

Each image and the corresponding prototypes are stored in a database during the training stage (Fig. 1). In the automatic stage, to classify a new image, the database is consulted and the corresponding prototypes are selected automatically. Once the prototypes are defined for each class, the minimum distance method (Girard and Girard, 1999) is used for classification. The Euclid distance between every pixel and the prototypes is calculated for each pixel:

$$d_{k_p} = \sqrt{\left(Y_p - Y_k\right)^2 + \left(I_p - I_k\right)^2}, \quad (4)$$

where Y_k and I_k are the median values of the prototypes of each class k, and Y_p and I_p are the components of the pixel. The classification assigns the pixel p to the class for which the distance to its prototype is minimal. Fig. 9 shows the five classes resulting of the classification stage corresponding to Fig. 7b.
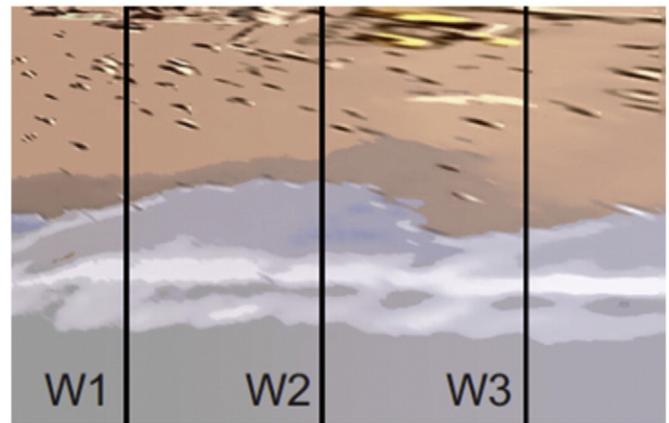
a)
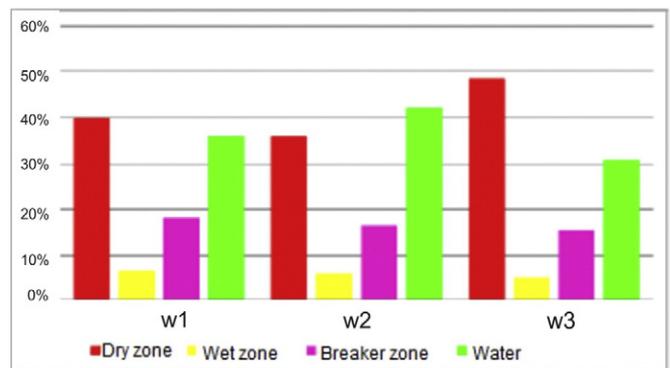
b)

Fig. 11. Beach Profiles of the Balneario Monte Hermoso corresponding to February 26, 2010. a) Location of the transects W1, W2, W3. b) Percentage variation of the transects according to the different areas.

Fig. 12. Mean images corresponding to different metereological conditions: sunset, sunrise, clouds or rain.
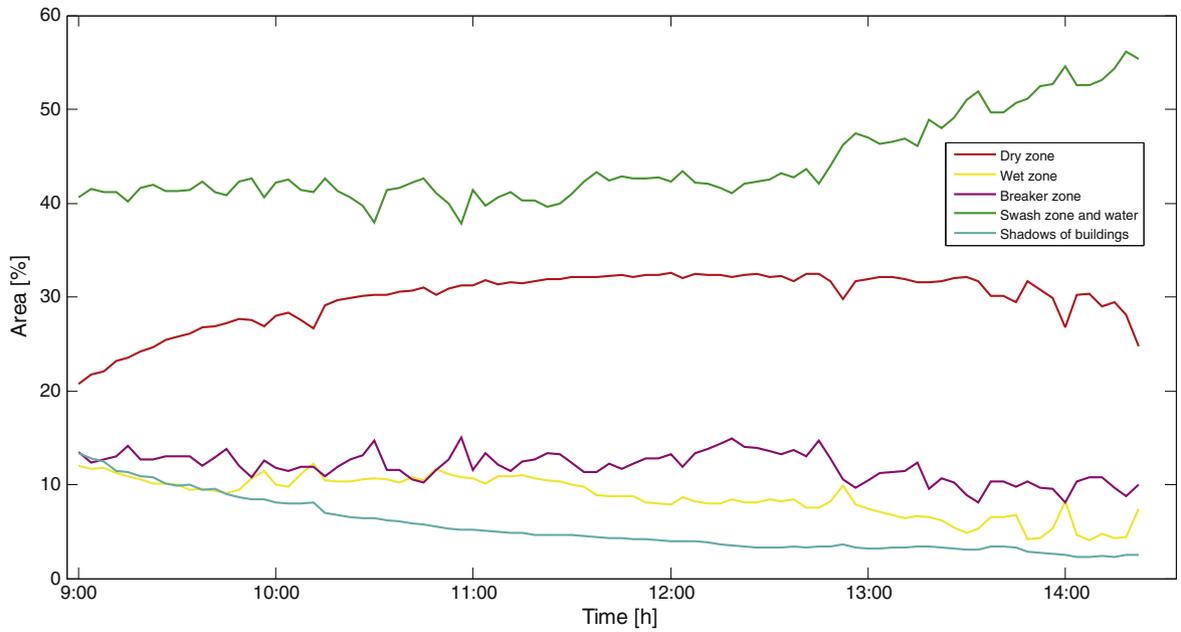

Fig. 13. Variation of areas in (m²) of the different beach zones along time.

2.5. Post-processing

The classification performed above is a per-pixel segmentation procedure that assigns each pixel in the ROI a class within the set of predetermined beach zones. The next purpose is to determine a clean representation of the area associated to each zone. However, as expected, the segmentation based on minimal distance produces some misclassified regions due to different uncontrolable conditions. This is illustrated in Fig. 10a, where the first zone segmented (dry beach) has 'holes' that do not belong to the class. In the wet beach, misclassified regions arise as small isolated groups of pixels, which also occurs in the third and fourth classes (breaker and water).

In order to obtain a clean representation for each zone, we compute the borders around the different image areas (Fig. 10b), retaining only the outermost contour associated to each zone (Fig. 10c and d). The final segmentation is produced by joining together the regions that were individually processed as explained above. In these figures, black regions represent the regions in the image that could not be classified. Within the dry zone, these regions mostly correspond to shadows, whereas in the wet zone they correspond to people in the beach.

2.6. Measurement

The planimetric area of each zone can be measured taking into account their correlative area in the picture, since after the zenital reprojection, every pixel in the reprojected image represents the same area unit. However, pixel counting has not enough precision if borders among areas are too convolved. For this reason, it is useful to represent the regions as enclosed in countours, represented using chain codes, from which the area calculation is performed with more precise numerical integration algorithms. Among the most used algorithms we can mention Chain Code (Freeman, 1970), Teh Chin Chain (Teh and Chin, 1989) and MSI (Cipolletti et al., 2012).

Another feature of interest in the beach is the width of the different regions. For this, we consider an straight transect orthogonal to the shoreline, which intersects the borders of the zones (Fig. 11a). The length of each region is determined by the distance between the intersection of the transect and the borders of the regions. As this information is registered in a video sequence, it is possible to measure the region width evolution along different transects (Fig. 11b).
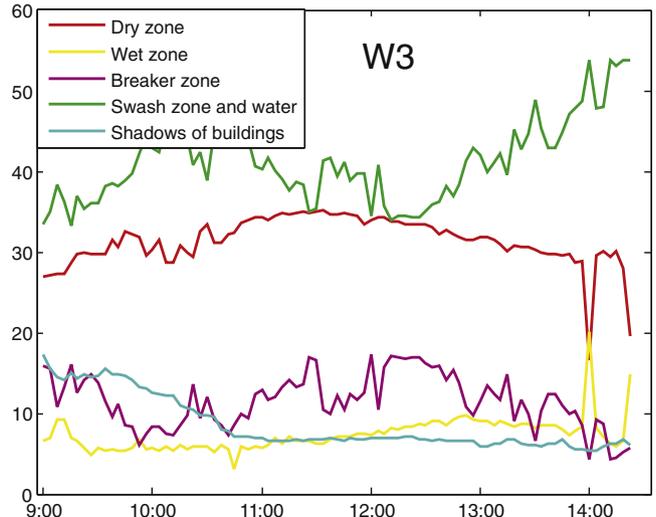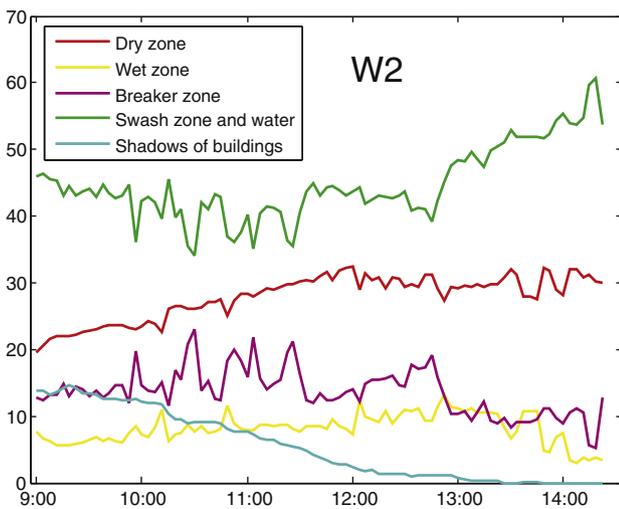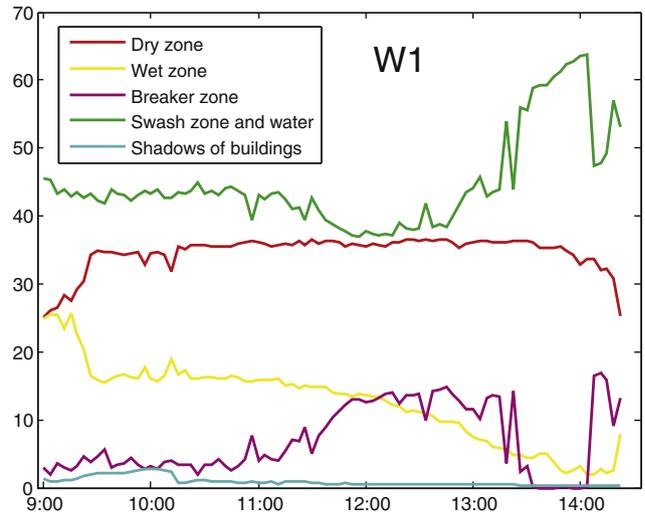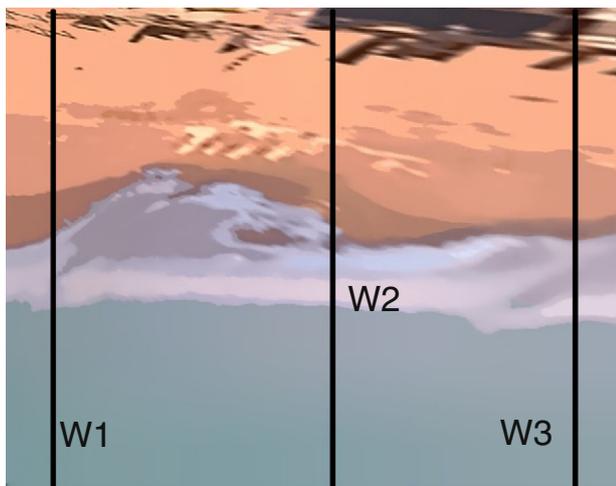

Fig. 14. Transects W1, W2, W3.

Fig. 15. Integration of automatically segmented beach zones in a Geographic Information System (original GIS is in Spanish for local managers).

3. Results and discussion

A set of videos were selected as representative examples of the results under different meteorological, oceanographic, and illumination conditions to study the robustness of the developed methodology (Fig. 12). After the rectification stage, we computed the mean image for 5 min video bursts. A preliminary analysis of these mean images allows the identification of the four zones of the beach (dry and wet beach, breaking zones, and swash zone and waters) but there are other elements that can be identified, for instance certain periodicities along the beach due to particular dynamic processes such as edge waves (Fig. 12a) or rhythmic bars (Fig. 12c). By themselves, changes in their area provide significant information regarding tidal stage, variations on the wave set up, changes in the breaking and offshore waters, which are all related to the modifications of the wave conditions offshore but also with variations in the underwater beach morphology.

Mean images are also very useful for quick qualitative identification of geomorphological features and, when rectified based on known coordinates, to estimate the areas of the different zones, along and across beach forms or particular sediment properties (i.e., areas of shell concentration or rock outcrops) and characteristics of the wave-induced conditions by following the form distribution. The presence of buildings or natural structures may cast shadows, and an unsupervised workflow must have a set of pixel prototypes to filter this kind of perturbation without losing information about the conditions at the beach.

The evolution of the area of each zone may be followed by plotting their changes in time along each mean image while enough illumination is available on the beach to adequately process the information. For example, we show the variation in the width and area of each zone during five hours in a specific spot of the beach of Monte Hermoso (Fig. 13). The selected video also includes, within the swash zone, the presence of a curved channel separating the swash bar from the dry beach. The wet zone (yellow) decreases when water covers the beach. The breaker zone stays with few oscillations like the dry zone (red). As the tide rises,

the water area (green) increases. Results of the variation of the beach zones are coincident with the qualitative information showed in images.

As mentioned above, another application of the segmentation of the different zones is the measurement of each zone lenght through a cross-shore profile. As an example, we take three transects W1, W2 and W3 to calculate the lenght of the beach (Fig. 14) and trace the variation of each zone. In transect W1 the areas of water offshore and dry beach have few variations. The wet zone decreases and the area of shadows is constant. It can be observed that when the breaker zone decreases the offshore increases correspondingly. The zone variation along transect W2 is similar to W1; however, in this case the area of the wet zone is constant and the curve of shadow decreases. In W3 the features are similar to W2.

As final result, we present the integration of the results in a Geographic Information System as a decision making tool.[4] In this case, the automatically segmented beach zones are georreferenced in a digital map of the Monte Hermoso beach (Fig. 15). This kind of information is increasingly required for coastal administration, which includes managment decisions and scientific studies. In addition to visual information, meteorological information was stored in the Geographic Information System. The meteorological conditions were taken and registered with the weather station designed and constructed by the Instituto Argentino de Oceanografía (IADO). The station is located 200 m to the W of the camera position along the coast and belongs to the coastal monitoring network facility of the Institute (EMAC, 2003). The meteorological parameters considered were air temperature, humidity, solar radiation, and wind direction and velocity.

3.1. Validation

The accuracy of the image rectification was tested using a set of distances between known points that were measured (Fig. 16a). These

---

[4] http://qgiscloud.com/Nrevollo/BeachZonesGIS.
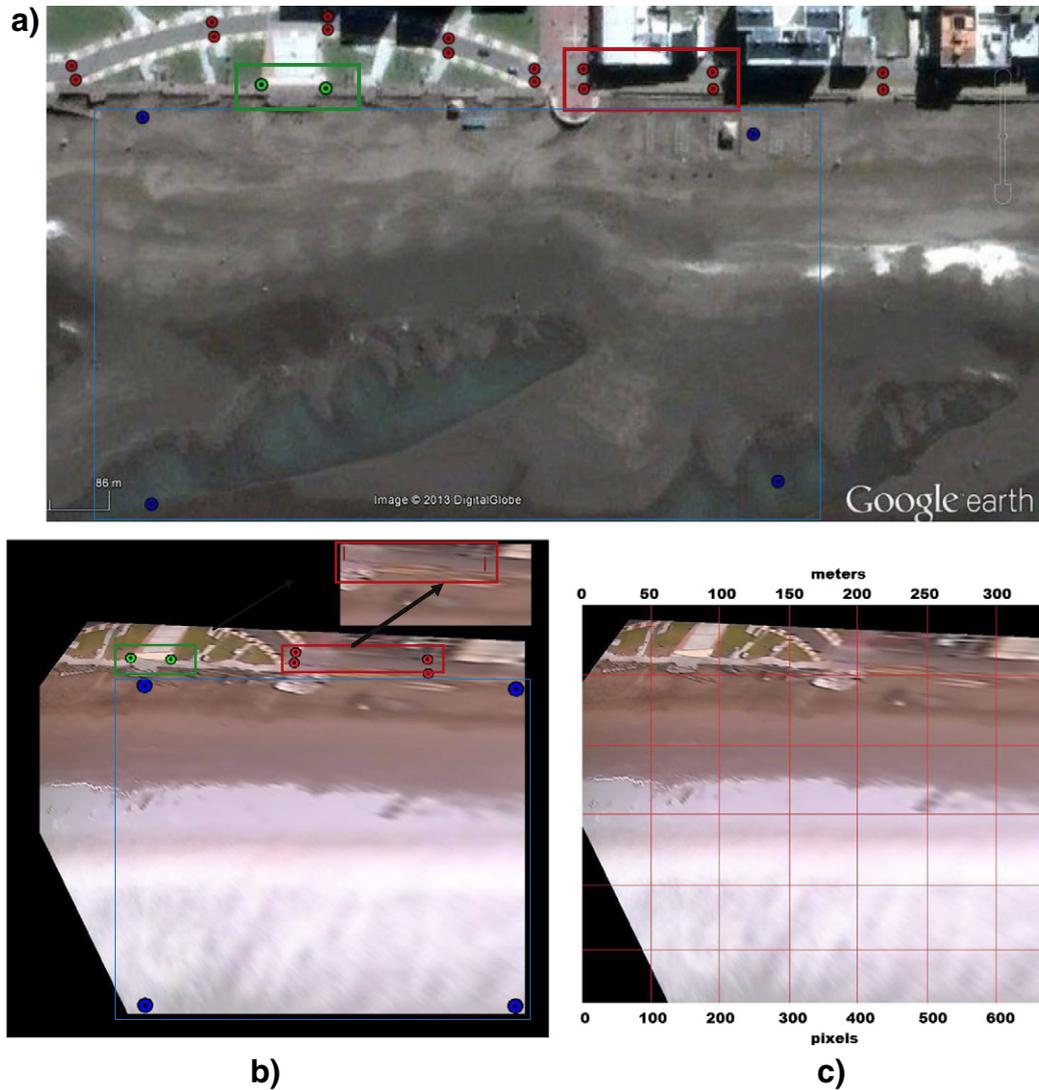
Fig. 16. Accuracy of the image rectification. a) Distances between points in terrain. b) Distances of the ground points over the rectified image. c) Image scale in pixels and their corresponding values in meters.

distances were also calculated among the corresponding points over the rectified image (Fig. 16b). After the reprojection, the pixel resolution on the image was about ≈0.5 m/pixel with a standard deviation of 0,025 m/pixel. A grid was superimposed on the image with a scale in pixels and their corresponding values in meters (Fig. 16c). For this image we provide the values of both real and estimated areas with the corresponding absolute error and percentage of matching for each segmented zone (Table 1).

The validation of the proposed pipeline requires measurements made by other techniques, in our case, human experts in recognizing beach zones from images. Direct, real-time inspection and measurement

Table 1
Area estimation, accuracy and error.

	Actual area	Estimated area	Accuracy	Error
Dry zone	3940m^2	3735m^2	94.0%	−215m^2
Wet zone	7429m^2	7463m^2	99.0%	+34m^2
Breaker zone	14085m^2	13888m^2	98.0%	−197m^2
Water	10810m^2	10611m^2	93.0%	−199m^2

is simply not feasible since beach zones are doomed to change during on-site inspection. In other words, there is no feasible "ground truth" to test against, and, therefore, a different validation methodology is required, which consists on the manual outlining of the different zones, performed by human experts using a GIS software. These supervised segmentations were used for testing the quality of the results of our unsupervised classification workflow. For this, three different segmentations were independently performed by three different experts (Fig. 17). Fig. 18 shows a zoom out of a comparison between the experts' segmentations and the automatic method.

The experts' evaluations also differ from each other, and, therefore, a methodological consensus index that takes into account all the mutual agreements and disagreements should be considered. A similar situation arises in the context of medical imaging, where also there is no ground truth against which an automated segmentation algorithm can be tested, and the quality assessment is related to experts' opinion. Often used for validation in this context is the Dice Similarity Coefficient (DSC) (Dice, 1945), which is based on the ratio of the intersection of two segmentations to their mean area. DSC, however, is reported to be less sensitive than newer validation methods, like the Validation Index (VI) (Juneja et al., 2013).
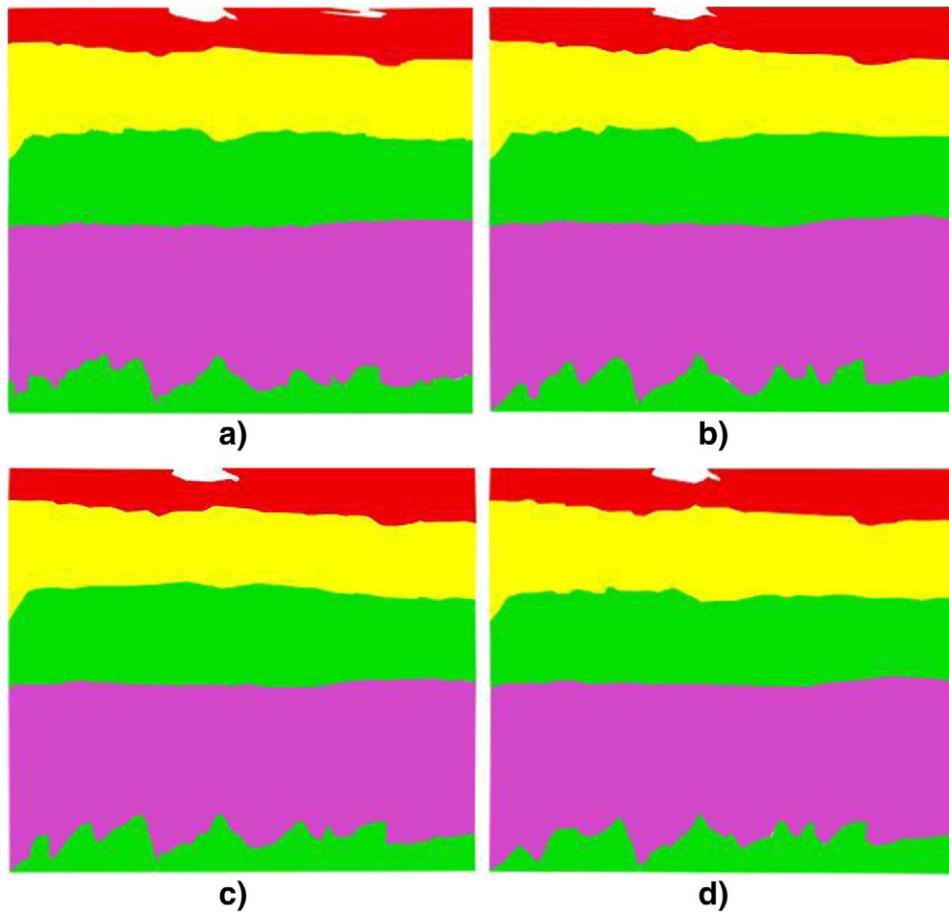
Fig. 17. Comparision of segmentation results. a) Automatic. b) Expert 1. c) Expert 2. d) Expert 3.

VI is based on different levels of agreement between multiple experts. In our case, the quality of the unsupervised segmentation can be evaluated by measuring the spatial coincidence or overlaps, of the different automatically and manually segmented zones. Hits (resp. misses) where more experts agree have a stronger positive (resp. negative) influence in the VI than the ones where fewer experts agree.

Given N human experts, a *membership value* $m = \frac{k}{N}$ is defined over every pixel, where k is the number of human experts classifying that given pixel as a specific zone type. For each area of overlap V_k (which k experts agreed that pertain to the same zone), the correlative area S_k segmented by the algorithm is defined as follows:

$$S_k = \begin{cases} S & \text{if } k = N \\ S_{k+1} - (V_{k+1} \cap S) & \text{if } k \in [1, N-1] \end{cases} \qquad (5)$$

The DSC of the automatically segmented zone S_k is then computed as follows:

$$DSC(k) = 2 \frac{V_k \cap S_k}{|V_k| + |S_k|} \qquad (6)$$

The validation measure (VM) of a specific point is defined as the product of its membership value m and the associated $DSC(k)$ at that point. Finally, the validation index (VI) is the average of the validation measures.

$$VM_i = m_i DSC(k_i) \qquad (7)$$

$$VI = \sum_{k=1}^{N} mDSC(k) \qquad (8)$$

VI equals zero if segmentation and expert outlines are disjoint and equals 1 if segmentation and all expert outlines overlap perfectly. VI for the segmentation is given in Table 2, the last column, indicates values greater than 0.75 in all the cases. These values indicate a near optimal performance of the automatic measurement of beach zones.

In addition, a confusion matrix was used to summarize the results of the automatic and manual segmentation (Fig. 19). For this we considered a sample of 24 images which were segmented manually. The diagonal values represent the percentage of accuracy for each zone with a mean of 95% of accuracy and an error of ± 200 m^2.

4. Conclusions and further work

Image and video processing of natural phenomena using inexpensive equipment posses a significant challenge and opportunity in environmental studies. In this work, we implemented a coastal monitoring and feature estimation system using image processing techniques to segment and measure some basic features of the beach. Specifically, the methodology was applied to segment beach zones during different tidal conditions. Our processing framework acquires videos from a static vantage point in a building. The results show that our proposed unsupervised processing algorithm classifies and measures the zones
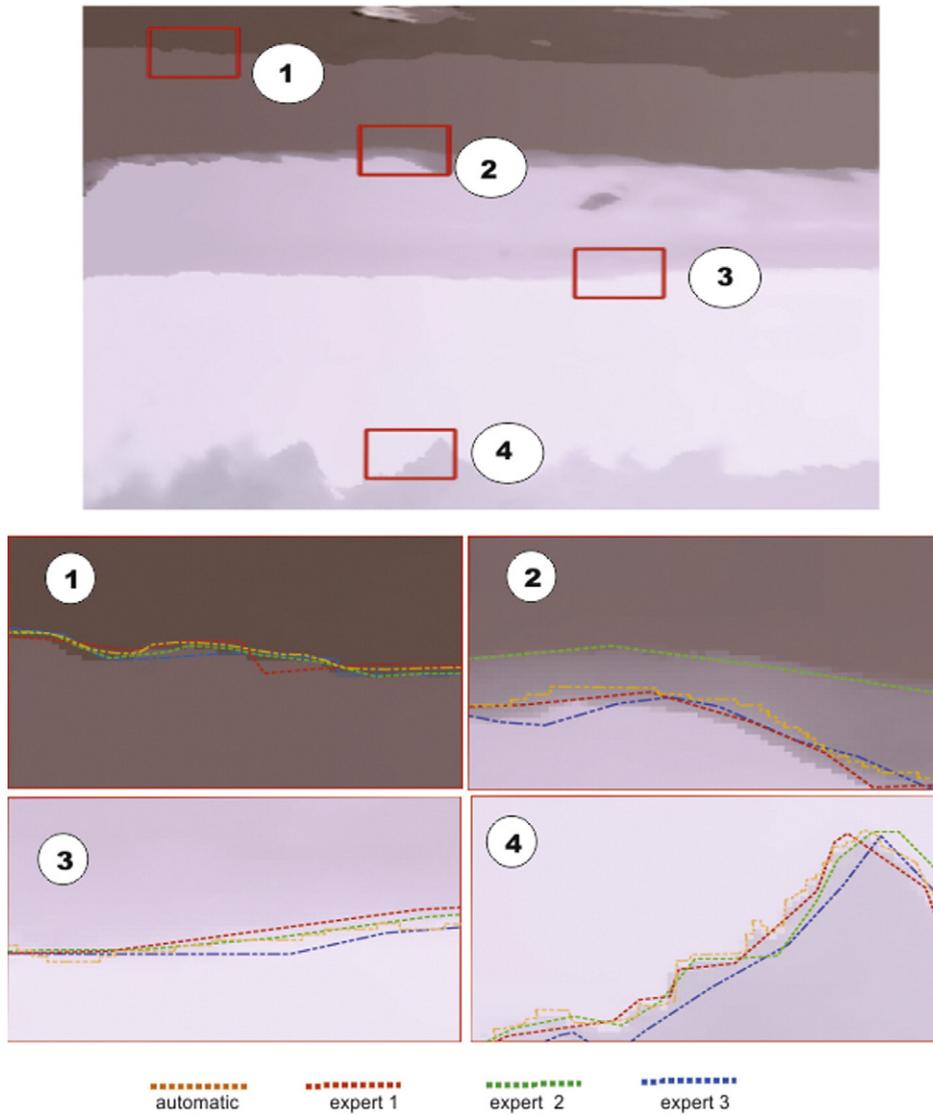
Fig. 18. Zoom out of the zones segmented showing the users segmentation versus the automatic segmentation.

automatically with high accuracy. The processing pipeline requires the callibration of only a small set of externally fixed parameters (the position and projection of four geographic coordinates and the selection of small training areas in the images) which typically require less than an hour to set up.

We are currently developing an autonomous embedded system that integrates the complete video-based monitoring system and is able to automatically upload the monitoring data through GSM connectivity. In that way we will be able to develop low-cost and easy to deploy video based sensor networks that will be able to operate autonomously in large geographical areas. Also, we are studying adaptive, automated color clustering methods in a way such that the zone identification is performed without human supervision.

Table 2
Validation measures according to the experts' segmentations.

Beach zone	VM$_1$	VM$_2$	VM$_3$	VI
Dry zone	0.328	0.651	0.991	0.766
Wet zone	0.33	0.688	0.99	0.783
Breaker zone	0.33	0.646	0.99	0.772
Water	0.328	0.687	0.999	0.780

Fig. 19. Confusion matrix of the segmentation algorithm. The diagonal values represents the percentage accuracy for each class considered.

Acknowledgements

References

Aarninkhof, S., 2003. Nearshore Bathymetry Derived from Video Imagery (PhD thesis) Delft University of Technology.

Aarninkhof, S.G.J., Holman, R.A., 1999. Monitoring the nearshore with video. Backscatter 10 (2), 8–11.

Cheng, Y., 1995. Mean shift, mode seeking, and clustering. IEEE Trans. Pattern Anal. Mach. Intell. 17 (8), 790–799.

Cipolletti, M.P., Delrieux, C.A., Perillo, G.M.E., Piccolo, M.C., 2012. Superresolution border segmentation and measurement in remote sensing images. Comput. Geosci. 40 (0), 87–96.

Comaniciu, D., Peter, M., 2002. Mean shift: a robust approach toward feature space analysis. IEEE Trans. Pattern Anal. Mach. Intell. 24 (5), 603–619.

Davidson, M., Koningsveld, M.V., de Kruif, A., Rawson, J., Holman, R., Lamberti, A., Medina, R., Kroon, A., Aarninkhof, S., 2007. The coastview project: developing video-derived coastal state indicators in support of coastal zone management. Coast. Eng. 54 (67), 463–475.

Dice, L., 1945. Measures of the amount of ecologic association between species. Ecology 1 (26), 297–302.

EMAC, 2003. Estación de monitoreo ambiental costero. http://emac.iado-conicet.gob.ar/.

Faugeras, O., 1993. Three-dimensional Computer Vision: A Geometric Viewpoint. MIT Press, Cambridge, MA, USA.

Freeman, H., 1970. Boundary encoding and processing. In: Lipkin, B.S., Rosenfeld, A. (Eds.), Picture Processing and Psychopictorics, pp. 241–263 (New York. Symposium on Psychopictorics held at Arlington).

Fukunaga, K., Hostetler, L., 1975. The estimation of the gradient of a density function, with applications in pattern recognition. Trans. Inf. Theory 21 (1), 32–40.

Girard, C.M., Girard, M.C., 1999. Processing of Remote Sensing Data. Dunod, Paris.

Hartley, R.I., Zisserman, A., 2004. Multiple View Geometry in Computer Vision. second edition. Cambridge University Press 0521540518.

Holman, R.A., 1981. Infragravity energy in the surf zone. J. Geophys. Res. Oceans 86 (C7), 6442–6450.

Holman, R., Stanley, J., 2007. The history and technical capabilities of argus. Coast. Eng. 54 (67), 477–491 (The CoastView Project: Developing coastal video monitoring systems in support of coastal zone management).

Hoonhout, B., Radermacher, M., Baart, F., van der Maaten, L., 2015. An automated method for semantic classification of regions in coastal images. Coast. Eng. 105, 1–12.

Huamantinco Cisneros, M.A., 2012. Efecto de la Variabilidad Climática del Balneario Monte Hermoso sobre su Geomorfología Costera y el Confort Climático (PhD thesis) Universidad Nacional del Sur.

Jensen, J.R., 2000. Introductory Digital Image Processing. A Remote Sensing Perspective. second edition. Prentice-Hall, Inc., Upper Saddle River, N.J.

Juneja, P., Evans, P., Harris, E., 2013. The validation index: a new metric for validation of segmentation algorithms using two or more expert outlines with application to radiotherapy planning. IEEE Trans. Med. Imaging 32 (8), 1481–1489.

Kroon, A., Davidson, M., Aarninkhof, S., Archetti, R., Armaroli, C., Gonzalez, M., Medrid, S., Osorio, A., Aagaarda, T., Holman, R., Spanhoff, R., 2007. Application of remote sensing video systems to coastline management problems. Coast. Eng. 54 (6-7), 493–505.

Lillesand, T., Kiefer, R., 2000. Remote Sensing and Image Interpretation. 4th. ed. Willey & Sons, New York.

Lippmann, T., Holman, R., 1989. Quantification of sand bar morphology: a video technique based on wave dissipation. J. Geophys. Res. 94 (C1), 75–95.

Madsen, A., Plant, N., 2001. Intertidal beach slope predictions compared to field data. Mar. Geol. 173 (14), 121–139.

Osorio, A., Medina, R., Gonzalez, M., 2012. An algorithm for the measurement of shoreline and intertidal beach profiles using video imagery: PSDM. Comput. Geosci. 46, 196–207.

Perez-Muñoz, J., Ortiz-Alarcón, C., Osorio, A., Mejía, C., Medina, R., 2013. Environmental applications of camera images calibrated by means of the levenberg-marquardt method. Comput. Geosci. 51, 74–82.

Plant, N.G., Holman, R.A., 1997. Intertidal beach profile estimation using video images. Mar. Geol. 140 (12), 1–24.

Quartel, S., Addink, E., Ruessink, B., 2006. Object-oriented extraction of beach morphology from video images. Int. J. Appl. Earth Obs. Geoinf. 8 (4), 256–269.

Schowengerdt, R., 1997. Remote Sensing Models and Methods for Image Processing. 2th. ed. Academic Press, San Diego.

Szeliski, R., 2010. Computer Vision: Algorithms and Applications. 1st edition. Springer-Verlag New York, Inc., New York, NY, USA.

Takewaka, S., Nakamura, T., 2001. Surf Zone Imaging with a Moored Video System (chapter 92) ASCE, pp. 1211–1216.

Teh, C.H., Chin, R., 1989. On the detection of dominant points on digital curves. IEEE Transactions on Pattern Analysis and Machine Intelligence 11(8), pp. 859–872.

Van Enckevort, I., Ruessink, B., 2001. Alongshore Uniform and Nonuniform Bar Crest Change (chapter 66) ASCE, pp. 656–665.

Williams, J.J., Arens, B., Davidson, M.A., Dias, J.M.A., Howa, H., O'Connor, B., Sarmento, A., Smith, J., Voulgaris, G., 1998. India: inlet dynamics initiative algarve. OCEANS '98 Conference Proceedings vol. 3, pp. 1540–1546.

Yatabe, S., Fabbri, A., 1986. The application of remote sensing to Canadian petroleum exploration: promising and yet unexploited. Comput. Geosci. 12 (4), 597–609 (Proceedings of the 14th Annual Geochautauqua).