



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

UNIVERSIDAD NACIONAL  
AUTÓNOMA DE  
MEXICO

**PROGRAMA DE MAESTRÍA Y DOCTORADO  
EN INGENIERIA**

Determinación de parámetros costeros con técnicas de visión  
computacional

**T E S I S**

Como requisito para obtener el grado de  
**Maestro en Ingeniería (Ingeniería Civil-Hidráulica)**

**P r e s e n t a**

**JAIME ALONSO ARRIAGA GARCÍA**

DIRECTOR DE TESIS

**Dr. EDGAR GERARDO MENDOZA BALDWIN**

MÉXICO, D.F.

NOVIEMBRE 2011





# Agradecimientos

Me gustaría agradecer a la UNAM y al Instituto de Ingeniería por brindarme un Oasis dentro de esta urbe, han sido metafóricamente y literalmente una segunda casa para mí. La Universidad Nacional Autónoma de México representa los principios más nobles y puros en México.

A mi familia, que sin su apoyo y cariño estaría perdido en el mundo. Madre, para ti tus hijos siempre serán los mejores en lo que sea y tu siempre serás la mejor, irrepetible e inigualable. Padre, gracias por tus consejos, hasta por los que están de sobra y muy repetidos, no olvidaré los principios que me inculcaste. Hermanos, gracias por ser tan diferentes a mí y por hacer de esas discrepancias de pensar un motivo más para que seamos unidos.

Al Dr. Edgar Mendoza, no solo ha sido Tutor y Profesor ha sido también un muy buen amigo, por decir las cosas sin tacto y con razón.

Al Dr. Rodolfo Silva por profesar su filosofía de trabajo con el ejemplo, hasta este día me continúa sorprendiendo su entusiasmo con la profesión.

Al Dr. Ismael Mariño por enseñarme que puede existir un balance entre el trabajo y la vida personal, y por brindarme en Mérida la oportunidad de ver una manera no ingenieril y muy útil de observar los procesos costeros.

Al Dr. Adrian Pedrozo por apoyarme durante este proceso final, por otorgarme un punto de vista diferente y alentarme a continuar con esta bella profesión.

Al Dr. Gabriel Echávez por compartir sus experiencias personales y profesionales. “Tres bolas a tres bandas más dos oponentes es igual a un juego de reyes.”

A mis compañeros del Instituto de Ingeniería por formar parte de este camino que se abre a muchos senderos, y en especial a los que se convirtieron en amigos, sin nombres para no ofender ni alabar, tan malo uno como otro. Un ejemplo de las lecciones aprendidas con ustedes.

***Por mi raza hablará el espíritu.***



# Resumen

El presente trabajo se enfoca en desarrollar técnicas de visión computacional para obtener parámetros costeros a partir de fotos y video. En general se desarrollan 3 aplicaciones: la extracción de la línea de costa; mediciones de runup; y un velocímetro de imágenes.

Diferentes algoritmos para extraer la línea de costa son programados y probados, y se desarrolla uno nuevo basado en la probabilidad de Bayes. Este proceso se trabaja sobre imágenes promedio calculadas a partir de una serie de trecientas que están previamente calibradas. Se utilizan transformaciones de espacio de color RGB a HSV, operaciones morfológicas y operadores tipo sobel para extraer el contorno del mar.

Las mediciones de ascenso máximo se toman a partir de 2 formas: una manual y una automática. La forma manual aplica la técnica de imágenes apiladas temporalmente, se selecciona un perfil (línea de píxeles) y se extrae ese mismo perfil de cada cuadro y se une en una nueva imagen, a partir de características ópticas se identifican los ascensos máximos. La forma automática hace uso de la segmentación por movimiento, el movimiento es definido a partir de la diferencia entre píxeles de distintos cuadros y si esta diferencia supera un umbral, el píxel se encuentra en movimiento, se usa el promedio corrido para disminuir las inconsistencias de iluminación causadas por la rotura del oleaje. El área de cálculo se limita a la pendiente de la playa y la punta de la lamina de agua que sube por la pendiente es identificada como el ascenso.

El velocímetro de imágenes se usa a partir de videos provenientes de una cámara de alta velocidad. Se basa en el coeficiente de correlación normalizado para la detección del desplazamiento, el tamaño de la ventana es variable, al igual que los cuadros de comparación.

## Abstract

The present work focuses in developing computer vision techniques to obtain coastal parameters from pictures and video. In general 3 applications are developed: Coastline extraction; Runup measurements; and an Image Velocimetry.

Different algorithms to extract the coastline are programmed and tested, a new algorithm based on Bayes probability is explained and tested. This process is worked on timex images calculated from a series of 300 pictures previously calibrated. Transformations from RGB color space to HSV, morphological operators and sobel operators are used in order to extract the countour of the sea.

The runup measurements are taken in two ways: a manual and an automatic. In the manual form the time slicing technique is applied, a profile is selected (a line of pixels) and from every frame this same profile is extracted and united in a new picture, from optical characteristics is possible to define the runup manually. The automatic form uses motion segmentation, the motion is defined by the subtraction of pixels and if the result of this operation is above a defined threshold this pixel is in motion, an average mean filter is used to diminish the effect of the light inconsistency produced by the breaking waves. The calculus area is limited to the beach slope, the tip of the water sheet is identified as the runup.

The image velocimetry uses a video collected by a high-speed cam. It's based in the normalized correlation coefficient for displacement detection, the interrogation widow size is variable, as well as the comparison frame.



# Índice

1	INTRODUCCIÓN .....	1
1.1	Objetivos.....	4
1.2	Estructura de la tesis .....	4
2	ASPECTOS FUNDAMENTALES DE LA VISIÓN COMPUTACIONAL .....	5
2.1	Aplicaciones y herramientas .....	5
2.2	Formación de la Imagen .....	11
2.3	Mejoramiento de la Imagen .....	13
2.4	Segmentación .....	19
2.5	Procesamiento Morfológico .....	22
2.6	Procesamiento en Color .....	23
2.7	Calibración de la Cámara .....	28
2.8	Transformaciones Geométricas .....	31
3	EXTRACCIÓN DE LA LÍNEA DE COSTA .....	35
3.1	ARGUS.....	35
3.2	IBIS.....	37
3.3	Tipos de Imágenes.....	38
3.4	Revisión de los Algoritmos y Modelos Existentes .....	39
3.4.1	Antecedentes.....	39
3.4.2	Aproximación Sencilla a la detección de la línea de costa .....	40
3.4.3	SLIM (Shore Line Intensity Maximum).....	41
3.4.4	Método “RatiosRGB” .....	43
3.4.5	Modelo SDM (Shoreline Detection Model) .....	44
3.4.6	Clasificador Bayesiano .....	46
3.5	Comparación de las técnicas .....	51
3.6	Geo-referencia.....	55
3.7	Conclusiones.....	61
4	MEDICIONES DEL ASCENSO MÁXIMO .....	63
4.1	Introducción .....	63
4.2	Secuencias de imágenes.....	67
4.3	Segmentación por Movimiento.....	72
4.3.1	Procedimiento Básico .....	72

4.3.2	Promedio Corrido .....	76
4.4	Conclusiones.....	79
5	VELOCÍMETRO DE IMÁGENES.....	81
5.1	Historia .....	81
5.2	Correlación .....	83
5.3	Aplicación de la técnica a casos reales.....	87
5.4	Conclusiones.....	93
6	CONCLUSIONES GENERALES.....	95
6.1	Recomendaciones .....	96
6.2	Futuras Líneas de Investigación.....	96
7	Bibliografía.....	97



# Lista de Figuras

Figura 2.1 Burbujas de jabón, tomada de (Geissler y Jahne, 1995).	6
Figura 2.2 Imágenes binarias del patrón de flujo con burbujas de dos cámaras posicionadas ortogonalmente. La figura muestra un retraso de 10 cuadros consecutivos de una medición de PIV. Las flechas delineadas manualmente muestran las trayectorias del flujo. (a) Plano Y-Z (b) Plano X-Z (Hassan <i>et al.</i> , 1998) $\lambda$	7
Figura 2.3 Concepto de cálculo virtual de ola, se usan 2 cámaras que toman al mismo tiempo la misma escena desde 2 perspectivas distintas (Bechle y Wu, 2011)	7
Figura 2.4 El viento sopla hacia la derecha. De izquierda a derecha se observa la pendiente de oleaje, componente filtrado y vectores de propagación, tomadas de (Zhang, 2003)	8
Figura 2.5 Extracción de contorno para arena, lodo y roca, tomadas de (Moldem <i>et al.</i> , 2008)	8
Figura 2.6 Diagrama de las tareas del Procesamiento de Imágenes Digitales	10
Figura 2.7 Representación del ojo viendo a una palmera, el punto C es el centro óptico del lente, tomada de (Gonzalez y Woods, 2002)	11
Figura 2.8 Proceso de la adquisición de una imagen digital, modificada de (Gonzalez & Woods, 2002)	12
Figura 2.9 Convención de coordenadas usada para representar imágenes, modificada de (Gonzalez y Woods, 2002)	13
Figura 2.10 de las funciones básicas en los niveles de gris, tomada de (Gonzalez & Woods, 2002)	14
Figura 2.11 Imagen Negativa de una ola	14
Figura 2.12 El histograma de una ola rompiendo.	15
Figura 2.13 El histograma ecualizado y su efecto sobre la imagen.	15
Figura 2.14 Imagen Instantánea y la Varianza de 300 imágenes, tomadas desde el hotel Royal Sunset.	16
Figura 2.15 Operador gradiente Sobel en X y en Y	18
Figura 2.16 Laplaciano simple y tomando en cuenta las diagonales, de tamaño 3x3.	18
Figura 2.17 Gotas de lluvia en un vidrio (izq.), el Gradiente (centro) de la imagen y el Laplaciano (der.), tomada de Michael Kiel.	18
Figura 2.18 Imagen de una huella digital, se muestra el histograma. Se binariza la imagen con un umbral de 70.	19
Figura 2.19 Imagen original, imagen segmentada con umbral simple y con umbral adaptativo.	20
Figura 2.20 a) Histograma de una imagen b) Segmentación por divisoria, locales mínimos representan la cuenca, niveles máximos al dique de la divisoria.	21
Figura 2.21 Imagen de cuadros de tamaños 1, 3, 5, 7, 9, 15 píxeles (izquierda), erosión de la imagen con un cuadro de 13 píxeles (en medio) y dilatación de la imagen erosionada con el mismo elemento (derecha), tomada de (Gonzalez y Woods, 2002).	22
Figura 2.22 Espectro de color visto por un rayo de luz blanco a través de un prisma.	24
Figura 2.23 Cubo de color RGB	24
Figura 2.24 Una imagen de varias frutas y su división en los 3 colores primarios.	25
Figura 2.25 Modelo de Color HSI.	26
Figura 2.26 Los canales HSI de la variedad de frutas.	28
Figura 2.27 Modelo Pinhole	29
Figura 2.28 Imagen de un tablero de ajedrez sin rectificar (izquierda) y rectificada (derecha).	31
Figura 2.29 Una imagen tipo tablero de ajedrez (arriba), transformación afín (izquierda) y proyección perspectiva (derecha).	34

Figura 3.1 Configuración del sistema Argus, Egmondaan, Países Bajos.....	36
Figura 3.2 Imágenes de muestra tomadas por el sistema ARGUS instalado en Milford en Sea, Reino Unido: a) Fotografía instantánea; b) Imagen promedio; c) Imagen varianza .....	37
Figura 3.3 Imagen promedio de muestra tomadas por el sistema Cam-Era instalado en Taihura, _Nueva Zelanda. ....	37
Figura 3.4 Sistema IBIS instalado en Cancún.....	38
Figura 3.5 Imágenes de muestra tomadas por el sistema IBIS en Cancún, México: a) Fotografía instantánea; b) Imagen Promedio; c) Imagen Varianza. ....	39
Figura 3.6 Aproximación sencilla, idealizada, para localizar la línea de costa. ....	40
Figura 3.7 (I) Playa de Coolangata en RGB con el perfil señalado (D) La componente V del espacio HSV .....	41
Figura 3.8 Zoom de la zona de interés y su perfil de intensidades. ....	42
Figura 3.9 (I) Imagen de Cancún con el perfil señalado (D) La componente V del espacio HSV .....	42
Figura 3.10 Zoom de la zona de interés y su perfil de intensidades. ....	42
Figura 3.11 El ratio a lo largo de la zona de interés (Cancún). ....	44
Figura 3.12 Los canales Hue (SD) Saturation (II) e Intensidad de la imagen (SI). ....	44
Figura 3.13 Histograma de Intensidades de Pixeles y la función discriminante, $p_1=5$ y $p_2=1$ . ....	45
Figura 3.14 a) Imagen binarizada; b) Imagen resultante de aplicar operadores morfológicos; c) Imagen resultante de aplicar un detector de bordes.....	46
Figura 3.15 La línea de costa resultante con el modelo SDM, parámetros $p_1=5$ y $p_2=1$ .....	46
Figura 3.16 Canal HUE. ....	49
Figura 3.17 Histograma de la componente HUE. ....	49
Figura 3.18 Imagen binaria a partir de la clasificación bayesiana de la componente HUE.....	50
Figura 3.19 a) Imagen resultante al aplicar erosión a 3.18; b) Imagen resultante de aplicar dilatación a 3.19a. .....	50
Figura 3.20 a) Detección de 2 áreas de playa; b) Eliminación del área falsa con vecinos de conectividad 8. ...	51
Figura 3.21 Contorno encontrado a partir del detector de bordes tipo Sobel de tamaño $3*3$ . ....	51
Figura 3.22 Imágenes de entrenamiento, la línea negra representa la interface agua-tierra, vista 1-7. ....	52
Figura 3.23 Línea de costa encontrada, Vistas 8-14 .....	53
Figura 3.24 Comparación entre SDM, el clasificador bayesiano y la línea ubicada manualmente, Vista 11. ...	53
Figura 3.25 Comparación entre SDM, el clasificador bayesiano y la línea ubicada manualmente, Vista 11. ...	54
Figura 3.26 Comparación entre SDM, el clasificador bayesiano y la línea ubicada manualmente, Vista 13. ...	55
Figura 3.27 Comparación entre SDM, el clasificador bayesiano y la línea ubicada manualmente, Vista 13. ...	55
Figura 3.28 Proyección de las esquinas y su extracción. ....	56
Figura 3.29 Acercamiento de la Figura 1.28. ....	56
Figura 3.30 Parámetros de calibración para la cámara de Royal Sunset. ....	57
Figura 3.31 Con parámetros de calibración, errores de proyección en pixeles de una imagen (I) y de todas (D). .....	57
Figura 3.32 Sin parámetros de calibración, errores de proyección en pixeles de una imagen (I) y de todas (D). .....	58
Figura 3.33 Corrección de una imagen con los parámetros hallados. ....	58
Figura 3.34 Ejemplo de punto de control de una vista, los 4 puntos dibujados. ....	59
Figura 3.35 Proyección Perspectiva.....	59
Figura 3.36 Proyección perspectiva en coordenadas UTM. ....	60
Figura 3.37 Línea de costa de la figura 1.36, localizada con el clasificador bayesiano. ....	61
Figura 4.1 Esquema de la sobreelevación del nivel del mar. ....	64

Figura 4.2 Esquema de las zonas cercanas a la costa.....	65
Figura 4.3 Imagen tipo apilado temporal "time-stack" tomada de (Salmon <i>et al.</i> , 2007). .....	66
Figura 4.4 Esquemmatización de la técnica de rebanar en el tiempo.....	68
Figura 4.5 Diferencia en pixeles de la anchura del canal. ....	69
Figura 4.6 Estimación del error en pixeles de la cámara robótica. ....	70
Figura 4.7 Líneas definidas por el usuario. ....	71
Figura 4.8 Imagen apilada "time-stack" .....	71
Figura 4.9 Ascenso máximo encontrado por 3 diferentes usuarios y uno de comparación (rojo). ....	72
Figura 4.10 a) Cuadro # 105; b) Imagen máscara con las coordenadas de la pendiente; c) Pixeles detectados en movimiento. ....	74
Figura 4.11 Errores cuadráticos con diferentes umbrales para los casos 1 y 2.....	75
Figura 4.12 La línea negra corresponde a las coordenadas en x del ascenso máximo observado, la azul representa al encontrado a través de la segmentación para los primeros 500 cuadros equivalentes a los primeros 17 segundos. ....	75
Figura 4.13 Variación de la iluminación en un punto debajo de la pendiente de la playa. ....	76
Figura 4.14 Error Cuadrático Medio para cada alfa con 15 umbrales.....	77
Figura 4.15 Acercamiento de la figura anterior en la zona con menores errores.....	77
Figura 4.16 La variación del error para los 3 alfas con mejor desempeño. ....	78
Figura 4.17 Ascenso máximo observado (negro) contra ascenso máximo encontrado (Azul), con un alfa de 0.17 y umbral 16.....	78
Figura 5.1 Esquema de la Correlación entre una ventana menor y una mayor. ....	84
Figura 5.2 (I) El abecedario representado en una imagen (D) La sub-ventana que representa la letra A.....	85
Figura 5.3 Representación numérica de la Figura 1.2 (letra A) .....	86
Figura 5.4 Superficie de Correlación. ....	86
Figura 5.5 Representación visual del campo de correlación. ....	87
Figura 5.6 Velocidades encontradas en una ola antes de romper. ....	89
Figura 5.7 Velocidades encontradas después de que la rotura ocurrió.....	89
Figura 5.8 Velocidades de gotas cayendo provocadas por la salpicadura de la ola.....	90
Figura 5.9 Definición del área de cálculo.....	90
Figura 5.10 Velocidades encontradas en un instante dado. ....	90
Figura 5.11 Velocidades promedio encontradas con un tamaño de ventana de 15 pixeles. ....	91
Figura 5.12 Velocidades promedio encontradas con una ventana de 33 pixeles.....	92
Figura 5.13 Velocidades promedio encontradas con los parámetros óptimos.....	92



# 1 INTRODUCCIÓN

El mantenimiento efectivo de las múltiples funciones de la costa, como protección, recreación, hábitat, navegación y pesca requiere de la comprensión y caracterización de los complejos sistemas y procesos físicos costeros. Una de las principales áreas de oportunidad, por no decir deficiencias, en los proyectos de ingeniería de costas, es la falta de información base para que los gestores puedan evaluar el impacto de diversas alternativas de actuación humana a fin de ponderar la ejecución de alguna o varias. Normalmente, la información que se recolecta para una zona de estudio se puede clasificar en 2 tipos: una relacionada con la topografía/batimetría y otra con el fluido, la primera ha de diagnosticar procesos de erosión, rebase y embrechamiento de dunas, variaciones morfológicas de barras y variaciones de la línea de costa, mientras que la segunda busca identificar los agentes marinos y meteorológicos causantes de la dinámica costera. En países como México, es muy raro contar con información suficiente y confiable de ambos tipos.

La morfología de la zona litoral es muy variable en un amplio rango de escalas espacio-temporales, existen variaciones morfo-dinámicas en una escala temporal de horas a siglos y en la escala espacial de metros a cientos de kilómetros (De Vriend, 1997), la zona cercana a la costa es probablemente la más dinámica de todos los ambientes costeros. La variabilidad morfológica de pequeña escala, de días a meses, puede impedir una correcta observación de la evolución de la costa en escalas de tiempo más largas, de años a décadas, y a la inversa, provocando la necesidad de desarrollar técnicas de mediciones permanentes y de alta resolución, que permitan tomar datos en escalas espaciales de metros a kilómetros y temporales de días hasta décadas.

Tradicionalmente, las estrategias del manejo costero se han basado en escalas temporales de años a décadas. Holanda, por ejemplo, cuya principal medida de intervención es la regeneración de playas, diseña los rellenos de playa de manera preventiva con base en una tendencia de diez años en la evolución del valor de un indicador de la línea de costa (Hamm *et al.*, 2002) determinado a partir de medidas anuales de la batimetría inter-mareal a lo largo de 250 m. En años recientes ha despertado un mayor interés en las escalas espaciales de metros a kilómetros y en las temporales de días a meses, debido a las intervenciones correctivas que se tienen que hacer para mitigar problemas locales de erosión (e.g. caso Cancún) adaptando así la forma de los perfiles de largo plazo en semanas.

Existen muy pocos datos morfológicos de alta resolución durante largos periodos, (Southgate *et al.*, 2003) mismos que son necesarios para resolver problemas costeros locales, y que permitan el desarrollo de la investigación en temas de dinámica costera. En México, lamentablemente, dicha información es prácticamente nula debido a la pobre cultura de monitoreo que aún prevalece. Los métodos tradicionales de medición en campo incluyen el uso de embarcaciones e instrumentos de

medición como ecosondas, boyas y perfiladores de corriente etc., entre otros. Estos medios de recopilar información, aunque proporcionan buenos datos, necesitan de grandes recursos humanos y económicos, pero quizás su mayor limitante es la resolución en las distintas escalas espacio-temporales, ya que no es posible mantener grupos de personas permanentemente en los sitios de estudio. Algunos centros de investigación y desarrollo, de los que las más notable es la Universidad Estatal de Oregón, han trabajado en crear herramientas para obtener datos de manera continua y en escala local como el ARGUS, cuyos orígenes se remontan a los años ochentas (Holman & Stanley, 2007) y ha sido usado en varios países como Holanda, Reino Unido, Estados Unidos, España, Brasil y Colombia. Uno de los programas más importantes de monitoreo en mar aparece con el nombre de CoastView, el cual tiene como base al ARGUS y fue apoyado por la Unión Europea (Davidson *et al.*, 2007). El principal objetivo de este proyecto consistió en simplificar las tareas de los manejadores costeros por medio del desarrollo de parámetros derivados de la observación continua de la costa con cámaras de video, conocidos como indicadores del estado costero (CSI, por sus siglas en inglés).

De forma paralela, y por lo que respecta al conocimiento de la hidrodinámica costera, existen también limitaciones en el conocimiento, como resultado de la naturaleza turbulenta de los flujos en la zona de rotura. Actualmente existen pocas técnicas no intrusivas para medir la velocidad del flujo cuando hay una densidad importante de burbujas, este tipo flujo bifásico (aire-agua) presente durante la rotura de oleaje, la interacción de oleaje con estructuras, en chorros de agua y saltos hidráulicos, representa uno de los retos más importantes en la ingeniería costera.

Diversas investigaciones se han realizado para caracterizar los flujos en la zona de rompientes y de transformación de la playa utilizando técnicas de medición que incluyen al LDV (velocímetro laser dopplery al PIV (velocímetro de partículas basado en imágenes), sin embargo los resultados obtenidos no han sido del todo satisfactorios (Greated y Emarat, 2000).

En 1977 tres grupos de investigación diferentes demostraron la factibilidad de medir el perfil parabólico del flujo laminar con la técnica LSV (velocímetro de tintes láser), originalmente desarrollada para medir desplazamientos de superficies de sólidos. En 1984 se le llamó PIV a la técnica LSV cuando la densidad de las partículas en las imágenes era tal que en lugar de observar tintes se apreciaban partículas, sin embargo el proceso para encontrar las velocidades es el mismo (Adrian R. J., 2004).

El PIV representa hoy día la técnica no intrusiva por excelencia, el diseño de su operación basado en el seguimiento de partículas artificialmente agregadas al flujo para revelar su naturaleza, nos transporta en el tiempo a la primera vez en la que una persona observó un puñado de hojas moverse en la superficie de un río. En la modernidad el PIV significa, mediciones cuantitativas de la velocidad de un fluido en un gran número de puntos, sin embargo, su uso está restringido de acuerdo a la naturaleza del flujo que se desea medir.

El PIV es una técnica robusta debido, principalmente, a su capacidad de registrar espacio-temporalmente las variaciones de un fluido de forma no intrusiva y su evolución ha pasado por varias etapas, algunas debidas a las limitaciones de la tecnología. El uso de computadoras en 1985, con memoria RAM de 128 kilobits limitaba el uso de la Transformada de Fourier 2D y el análisis de correlación, impulsando así la curiosidad de varios grupos de investigación para el cálculo de la correlación por métodos análogos (Morck *et al.*, 1993). La limitación de la toma de imágenes, que se encuentra en constante mejora, no permitía capturar imágenes con diferencias pequeñas en tiempo. Louis Lourenco convenció a la empresa Kodak para que se crearan cámaras de transferencia interlineal que pudieran soportar la grabación de 2 imágenes en sucesión rápida al transferir la primera imagen a un chip de almacenamiento y grabar la segunda imagen (Lourenco *et al.*, 1994), haciéndose notorio el impulso entre ramas distintas de la ciencia.

La técnica llega a presentar inconsistencias en zonas altamente aireadas, ya que pueden existir problemas de dispersión de la fuente de luz al encontrar zonas densas en burbujas (Adrian R. J., 1984). Entre las principales desventajas del PIV, se ha hecho evidente la alta sensibilidad del equipo a reflexiones no planeadas del laser que fácilmente llega a quemar el equipo óptico.

Recientemente, se desarrolló otra técnica llamada BIV (Ryu *et al.*, 2007), cuyo funcionamiento es similar al descrito para los sistemas LSV/PIV con algunas pequeñas diferencias. En esta técnica, se incorpora la técnica de iluminación de flujos conocida como shadowgraph, que tiene como función proyectar sombras de partículas trazadoras en el flujo, en un plano paralelo al movimiento del fluido. Las cuales se generan por medio de un potente equipo de iluminación, la determinación del campo de velocidades se estima al igual que en el PIV, por medio de un algoritmo de correlación cruzada. La combinación de la técnica de shadowgraph con algoritmos de correlación cruzada se debe a trabajos realizados por Hassan *et al.* (1998) y Lindken y Merzkirch (2001). La aplicación de esta técnica a flujos en rotura sobre estructuras ha sido ampliamente documentada por Ryu *et al.* (2007) o Ryu y Chang (2008). Recientemente su aplicación para la determinación del campo de velocidades en condiciones de rotura sobre una playa en laboratorio, fue reportado por Pedrozo-Acuña *et al.* (2011) y Rivillas-Ospina *et al.* (2011).

La automatización del proceso para obtener información por medio de imágenes, requiere implementar algoritmos para el tratamiento de imágenes y es esto, precisamente, el enfoque de la presente Tesis. Esta área es muy amplia y compleja, trata temas como la óptica, el reconocimiento de patrones, la robótica, la computación, visión artificial, entre otros. La aplicación que se le ha dado en el ámbito de ingeniería de costas es muy reciente y aun quedan por resolver muchas incógnitas, sin embargo, se considera que los beneficios que se pueden llegar a obtener son superiores que los retos a sortear.

## 1.1 Objetivos

El objetivo general de esta tesis es desarrollar técnicas de visión artificial para determinar parámetros costeros con distintos tipos de imágenes de video. Para alcanzar de manera satisfactoria el objetivo principal de la tesis, se proponen los siguientes objetivos específicos:

- Desarrollar un modelo integral para la detección y extracción de la línea de costa, junto con una orto-rectificación de las imágenes para poder geo-referenciar los datos obtenidos.
- Implementar una herramienta para la determinación de las velocidades en flujos aireados con base en el coeficiente de correlación normalizado.
- Dar seguimiento a la punta de un flujo en ascenso sobre un talud para comparar las mediciones realizadas en el laboratorio en distintos estados de oleaje con las ecuaciones que se encuentran en la literatura para predecir el run-up.

## 1.2 Estructura de la tesis

En general, en el presente trabajo se trata de mostrar la importancia y utilidad de la Visión Computacional/Artificial para determinar parámetros costeros; se muestra el desarrollo, implementación y los resultados de distintas técnicas.

En el segundo capítulo se repasan los aspectos fundamentales del procesamiento de imágenes digitales y la visión computacional, así mismo, se revisa el funcionamiento del ojo humano para un mayor entendimiento.

En el capítulo 3 se hace una revisión de los modelos de detección de línea de costa, así como el desarrollo de uno nuevo y la implementación de uno ya existente. El capítulo se complementa con el proceso necesario para tener datos confiables y medibles por lo que se incluye la calibración de la cámara y la orto rectificación de las imágenes.

En el capítulo 4 se desarrolla un programa, basado en el flujo óptico, que sigue la punta del run-up, obteniendo así gráficas de ascenso máximo vs tiempo, con las cuales se verifica la validez de las distintas ecuaciones para estimar el run-up.

En el capítulo 5 se implementa una herramienta para la medición de flujos aireados con base en el coeficiente de correlación normalizado, se muestran las gráficas de la evolución de la velocidad de un chorro de agua.

El capítulo 6 aborda las conclusiones de cada herramienta, el posible mejoramiento de las mismas y las distintas vertientes de investigación que se pueden seguir, aprovechando las bondades de la complementación de la ingeniería costera y la visión computacional.



## 2 ASPECTOS FUNDAMENTALES DE LA VISIÓN COMPUTACIONAL

### 2.1 Aplicaciones y herramientas

Desde los inicios de la ciencia, la observación visual ha sido crucial; aún cuando tradicionalmente la única manera de documentar experimentos era a través de la descripción verbal y dibujos. El siguiente gran paso en la visión fue la invención de la fotografía lo que favoreció la objetividad, sin embargo, los procedimientos de evaluación eran en su mayoría manuales y solo en casos semi-automáticos existían aplicaciones muy específicas, por lo que la evaluación cuantitativa de imágenes no tuvo gran difusión. Generalmente las imágenes se usaban para documentación, descripción cualitativa e ilustración de los fenómenos.

En los últimos años, el uso de las imágenes está tomando mayor relevancia debido al progreso en la tecnología de video y de computación además, la tecnología se encuentra al alcance de cualquier científico, ingeniero y estudiante. En consecuencia, el procesamiento de imágenes se ha expandido convirtiéndose en una herramienta científica estándar que se aplica en casi todas las ciencias naturales y disciplinas técnicas.

Para ejemplificar el poder de la información visual basta con tratar de describir un rompeolas, actividad que requeriría un gran esfuerzo sin el apoyo de imágenes y dibujos técnicos, sin dejar de lado una experiencia frustrante y tiempo perdido. Con apoyo visual se tendría una idea inmediata del rompeolas y se podrían estudiar detalles no otorgados en la descripción.

Desde su nacimiento, la ciencia del procesamiento de imágenes ha sido interdisciplinaria en varios aspectos y la cadena de procesos involucra varias áreas, como la física, la ciencia computacional y las matemáticas; incluyendo sub-disciplinas como óptica, física del estado sólido, diseño de chips, arquitectura computacional, algebra, estadística, teoría del algoritmo, teoría grafica, teoría de sistemas y matemática numérica. Desde un punto de vista ingenieril hay contribuciones de la ingeniería óptica, ingeniería eléctrica, fotónica e ingeniería de software.

El procesamiento de imágenes, además, se traslapa con otras disciplinas, como la metrología o el reconocimiento de patrones y tiene conexiones similares con redes neuronales, inteligencia artificial y percepción visual. La visión computacional, que Marr, (Marr, 1982), definió como “Descubrir a partir de imágenes qué hay en el mundo y dónde se encuentra” trata de imitar a la visión biológica a través del uso del procesamiento de imágenes y la geometría proyectiva.

Existen otras disciplinas que por razones históricas se desarrollaron de manera independiente. La fotogrametría, que se enfoca en la realización de mapas e inspección. Los sensores remotos en imágenes aéreas y satelitales. La astronomía y el área especializada de imágenes médicas.

Entre los aspectos importantes de la naturaleza interdisciplinaria del análisis de imágenes se cuenta su aplicación generalizada. No hay, prácticamente, ningún campo en las ciencias naturales o disciplinas técnicas donde el procesamiento de imágenes no encuentre aplicación, de ahí que el campo es tierra fértil para un progreso acelerado debido a la afluencia de técnicas, ideas y logros. Dicha naturaleza no es una extensión que se da por gusto o casualidad sino por necesidad. La falta de conocimiento, ya sea en el área de aplicación o en la del procesamiento de imágenes, inevitablemente lleva a soluciones sub-óptimas o mediocres y algunas veces al completo fracaso.

El procesamiento de imágenes digitales (DIP por sus siglas en inglés) permite el estudio de fenómenos complejos que no pueden ser evaluados adecuadamente con técnicas de medición convencionales. Una tarea clásica del DIP es el conteo de partículas y su medición. Por ejemplo, el conteo de burbujas Figura 2.1, la medición de burbujas de jabón o la dimensionalización de la forma de partículas pigmentadas. Estas tareas se pueden extrapolar a sistemas de inspección visual para control de calidad.

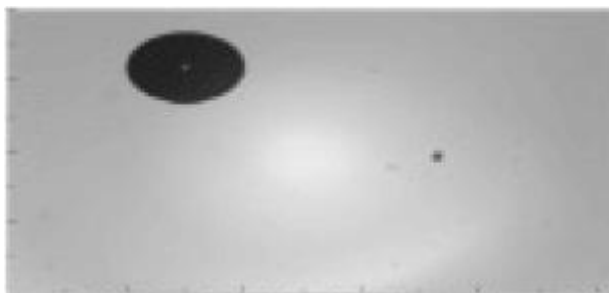


Figura 2.1 Burbujas de jabón, tomada de (Geissler y Jahne, 1995).

Una de las deficiencias que se encuentra al realizar mediciones con fotografías es la pérdida de profundidad de campo debido a que el espacio 3D es proyectado a un plano 2D. Para resolver esta pérdida se han desarrollado técnicas especiales para recuperar la topografía de imágenes volumétricas, generalmente se ocupan dos cámaras para poder reconstruir la escena en 3D, a lo que se le conoce como visión estéreo. En imágenes médicas, por ejemplo, se puede observar, a partir del tercer mes, al feto en 3 dimensiones, (Anquez *et al.*, 2007), mientras que en hidráulica se ha investigado la estructura tridimensional de flujos turbulentos (véase Figura 2.2). Respecto de problemas costeros, se ha logrado obtener las dimensiones 3D del oleaje con el concepto de cálculo virtual de olas, en donde la superficie del agua en el punto de interés debe existir en una línea vertical entre los niveles máximos y mínimos de superficie, esta línea se proyecta en las imágenes como una línea recta (véase Figura 2.3).

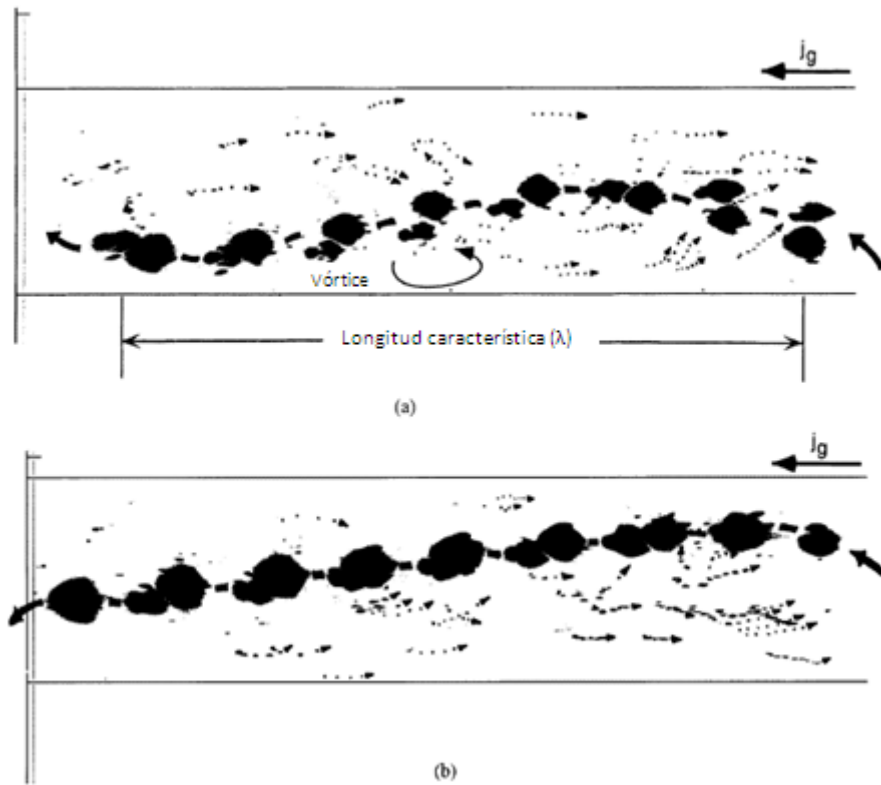


Figura 2.2 Imágenes binarias del patrón de flujo con burbujas de dos cámaras posicionadas ortogonalmente. La figura muestra un retraso de 10 cuadros consecutivos de una medición de PIV. Las flechas delineadas manualmente muestran las trayectorias del flujo. (a) Plano Y-Z (b) Plano X-Z (Hassan *et al.*, 1998)  $\lambda$

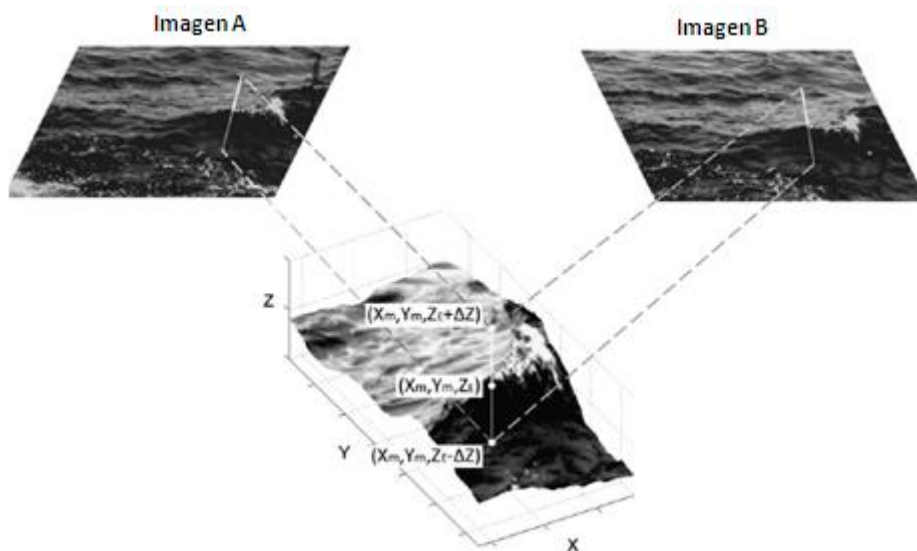


Figura 2.3 Concepto de cálculo virtual de ola, se usan 2 cámaras que toman al mismo tiempo la misma escena desde 2 perspectivas distintas (Bechle y Wu, 2011)

Los procesos dinámicos también han sido analizados a través de secuencias de imágenes. Los procesos de pequeña escala que toman lugar en la vecindad de la superficie del océano son de difícil medición por la ondulación de las olas, con mediciones puntuales es imposible conocer la

estructura de las ondas en la superficie, usando iluminación basada en el paradigma “forma de sombra” la pendiente de la ola es visible, por lo que es posible conocer el movimiento con tratamiento de imágenes.



Figura 2.4 El viento sopla hacia la derecha. De izquierda a derecha se observa la pendiente de oleaje, componente filtrado y vectores de propagación, tomadas de (Zhang, 2003)

Otra tarea importante de la visión computacional es la clasificación de objetos, que lleva, a su vez, al dilema de la extracción de características. Una técnica común es el uso de imágenes adquiridas con equipo acústico para clasificar el sedimento del fondo del mar tomando en cuenta la caracterización visual del mismo. Esta técnica se apoya en la forma del contorno del horizonte y en la textura que se registra debajo del mismo (véase Figura 2.5). La granulometría es una tarea tediosa y repetitiva, el sistema WipFrag es un sistema granulométrico automatizado que usa fotografías y videos para determinar la distribución de tamaño de grano de rocas (Maerz *et al.*, 1996), eventualmente es posible extrapolar la idea a tamaños de grano más pequeños.

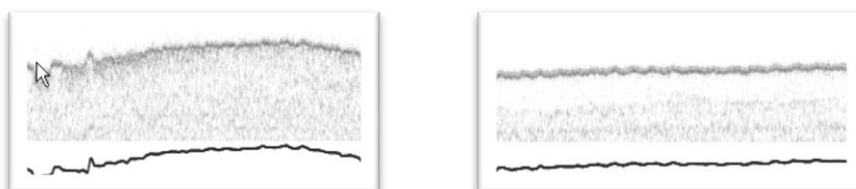


Figura 2.5 Extracción de contorno para arena, lodo y roca, tomadas de (Moldem *et al.*, 2008)

El DIP es un proceso que se puede separar en varios pasos con cierta secuencia para poder extraer información de interés, en la Figura 2.6 se muestra un diagrama que otorga una perspectiva general del proceso.

Se comienza con una captura, no necesariamente óptica, adecuada. Se selecciona el sistema de iluminación, la mejor longitud de onda y se seleccionan las opciones para capturar las características deseadas en la imagen. Una vez capturada la imagen se debe transformar para que las computadoras digitales puedan trabajar con ella, a lo que se le llama digitalización.

El pre-procesamiento es una serie de operaciones independientes de la imagen, se deben al tipo de sensor y lente de la cámara, por ejemplo, si el sensor tiene características no lineales, deben ser

corregidas. En otros casos la imagen puede necesitar mejoras en la brillantez y el contraste. Otro ejemplo de pre-proceso es la transformación de coordenadas para corregir las distorsiones geométricas introducidas durante la formación de la imagen.

Cuando se necesita analizar e identificar objetos, se requiere de una cadena de procesamientos. Inicialmente se requiere un filtro para discriminar el objeto de interés de otros y del fondo. Se prosigue con la extracción de características, requiriendo de operaciones y procesos como el promedio, la detección de bordes, el análisis de vecindades y de patrones complejos conocidos como textura. Otra característica de interés es el movimiento.

La segmentación es el siguiente paso esencial; dicha tarea es fácil si el objeto es contrastante con las características circundantes, cuando no es así, se emplean estrategias para disminuir la desviación entre la información de la imagen y una función que incorpora el conocimiento previo de los objetos en la imagen.

Existen también alteraciones conocidas, las cuales se pueden eliminar o reducir con algoritmos matemáticos, entre ellas se cuentan el enfoque inadecuado, el movimiento que produce borrosidad y errores en el sensor o en la transmisión de la señal.

Una vez conocida la forma geométrica del objeto se pueden usar operadores morfológicos para analizar y modificar la forma o extraer otro tipo de información como valor medio, área, perímetro y otros parámetros de la forma. Se usa dicha información para clasificar objetos.

Un problema que al principio resulta amplio y complejo, puede ser apreciado y resuelto con mayor facilidad si se divide en sus componentes más pequeñas y se da, a cada una de ellas, una secuencia lógica.

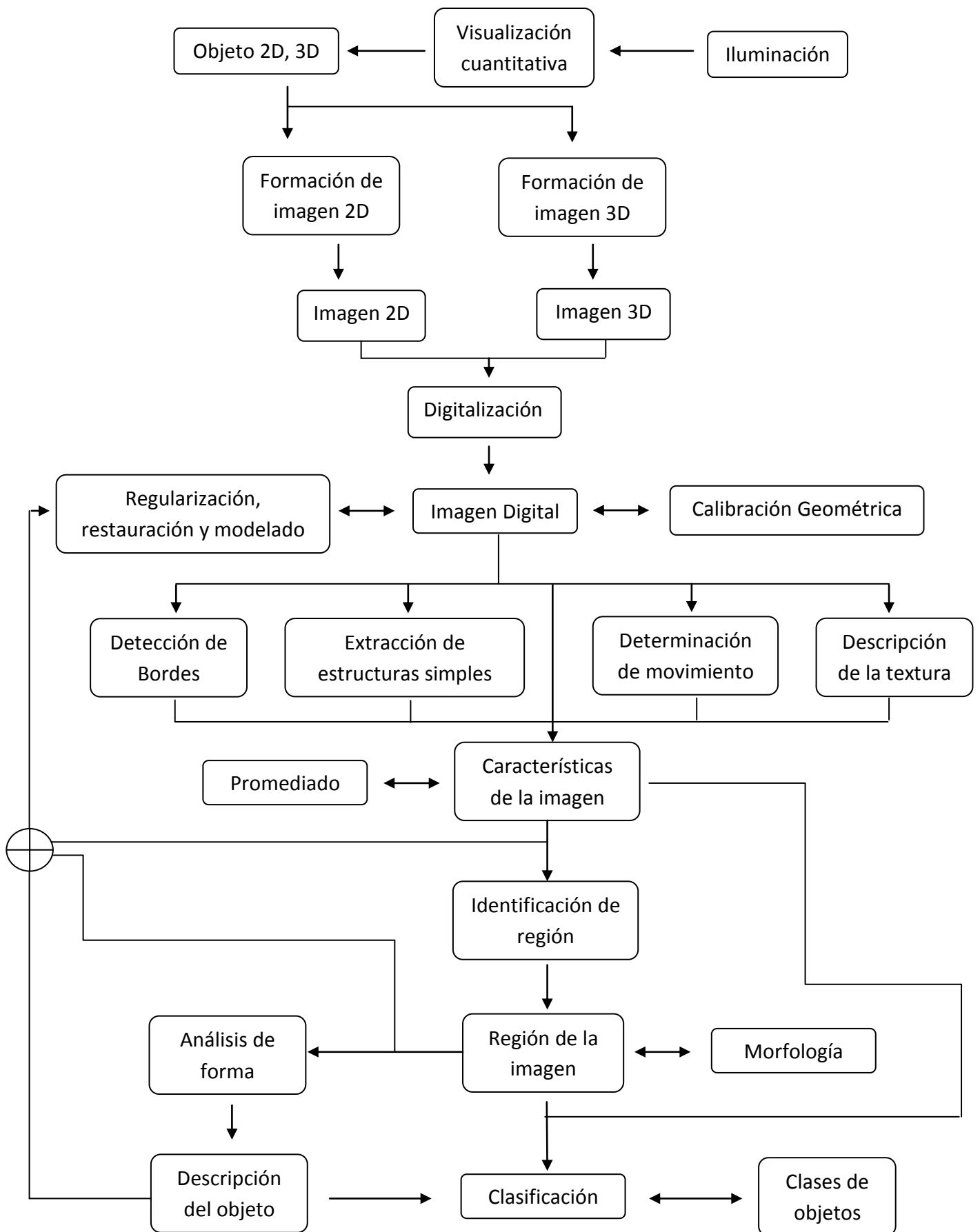


Figura 2.6 Diagrama de las tareas del Procesamiento de Imágenes Digitales

## 2.2 Formación de la Imagen

Las técnicas que se emplean para resolver problemas de visión, a pesar de estar basadas en formulaciones matemáticas y probabilísticas, generalmente se escogen subjetivamente con base en criterios humanos visuales, por lo que resulta importante entender mejor el funcionamiento del ojo humano y la forma en la que se adquieren las imágenes.

El ojo humano se puede aproximar a una esfera de diámetro promedio de 20 mm. Tiene 3 membranas principales: la cornea y esclerótica; la coroides y la retina. La cornea es un tejido duro y transparente que cubre la parte posterior de la superficie del ojo, la esclerótica es opaca y cubre el resto del globo óptico, la coroides se encuentra debajo de la esclerótica y nutre al ojo. La coroides se divide en el cuerpo ciliar y el diafragma del iris, el segundo controla la entrada de luz. La retina es la capa interior, aquí es donde se forma la imagen, el patrón que se forma se debe a diferentes receptores de luz: los conos se localizan en la parte central de la retina y son sensibles al color; los bastones están distribuidos en toda la retina y son sensibles a bajos niveles de iluminación.

El lente está hecho con capas concéntricas de fibras celulares y se encuentra suspendido por fibras que lo sujetan al cuerpo ciliar. La principal diferencia entre el lente humano y el lente de una cámara es que el primero es flexible. La forma del lente se controla por las fibras del cuerpo ciliar y para enfocar objetos distantes.

La selección adecuada de la distancia entre el lente y el objeto de estudio brinda una mejor adquisición de imagen. Dada una distancia de trabajo constante, ente más grande el campo de visión menor tendrá que ser la distancia focal. Si se aumenta la apertura del lente disminuye la profundidad de campo.

Tomando como ejemplo básico la Figura 2.7, se calcula el tamaño de la imagen que se proyecta en la retina o, para el caso de una cámara, de la imagen proyectada en el CCD (Charged-Coupled Device), si tenemos un árbol de 5 metros de alto a una distancia de 100 metros y lo enfocamos con una distancia de 17 mm, el tamaño de la imagen sería de  $(5/100=h/17)$   $h=2.55$  mm.

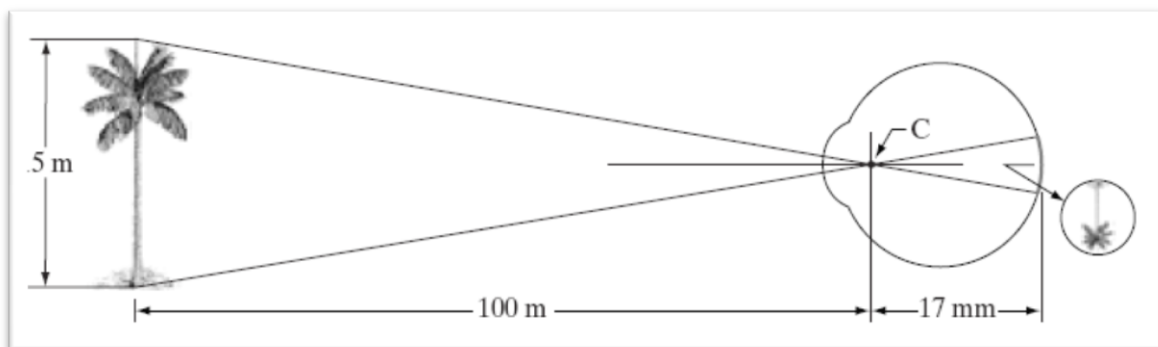


Figura 2.7 Representación del ojo viendo a una palmera, el punto C es el centro óptico del lente, tomada de (Gonzalez y Woods, 2002)

Las cámaras digitales contienen un gran número de sensores, acomodados en forma de matriz. Los sensores, que pueden tener distintas características, responden proporcionalmente a la energía proyectada en su superficie. Su funcionamiento básico se puede observar en la Figura 2.8

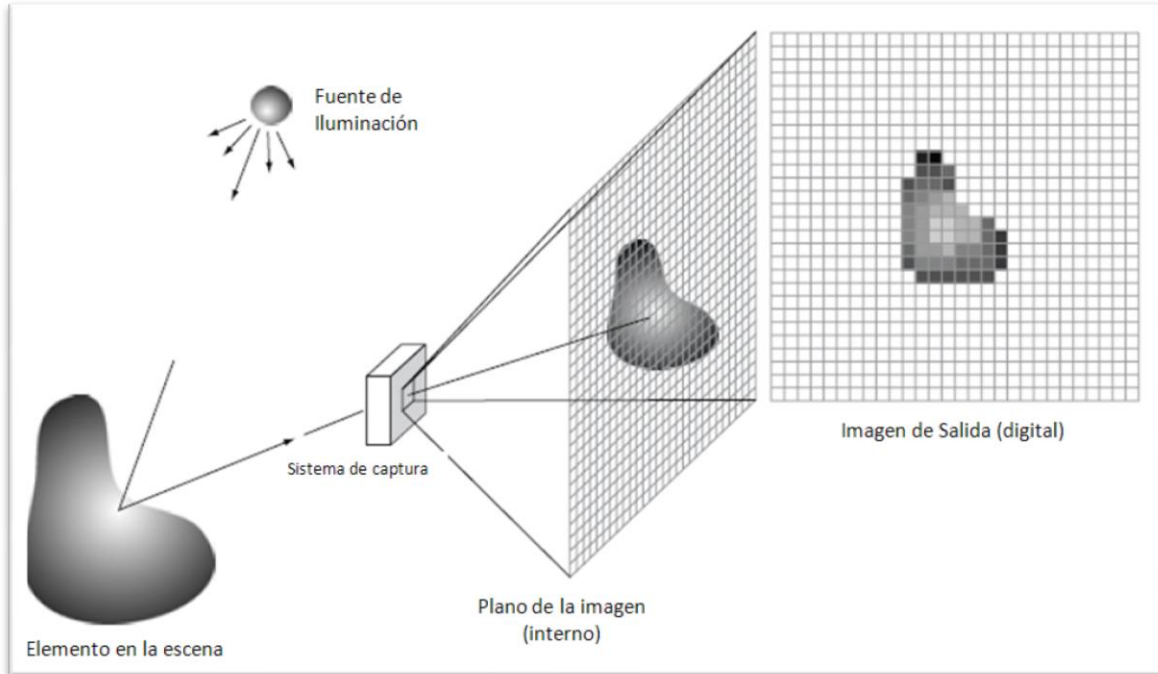


Figura 2.8 Proceso de la adquisición de una imagen digital, modificada de (Gonzalez & Woods, 2002)

Una imagen se puede ver como una función bidimensional de la forma  $f(x,y)$ , el valor de la amplitud de  $f$  en las coordenadas espaciales  $(x,y)$  es una cantidad positiva cuyo significado físico es determinado por la fuente de energía, luz, de la imagen; por lo tanto, tiene un valor teórico máximo que tiende al infinito.

La función  $f(x,y)$  se puede ver como el producto de 2 componentes: la cantidad de luz incidente en la escena observada, iluminación  $i(x,y)$ , y la cantidad de luz reflejada por los objetos en la escena, reflectancia  $r(x,y)$ . La reflectancia se limita al rango  $[0, 1]$ , donde 0 significa absorción total y 1 reflectancia total, la nieve alcanza un valor de 0.93

Una imagen en tonos de gris tiene una escala que se puede ver como un intervalo  $[L_{min}, L_{max}]$ , transformándolo al intervalo numérico  $[0, L-1]$  el primer elemento se considera negro y el segundo blanco, todos los valores intermedios son grises de distinta intensidad, las imágenes de 8 bits manejan 256 niveles. La digitalización de las coordenadas se conoce como muestreo y la digitalización de la amplitud se conoce como cuantización, en la parte derecha de la Figura 2.8 se aprecia el concepto.

El resultado del muestreo y la cuantización es una matriz de números reales de tamaño  $M$  filas por  $N$  columnas, donde cada elemento se llama pixel (véase Figura 2.9).



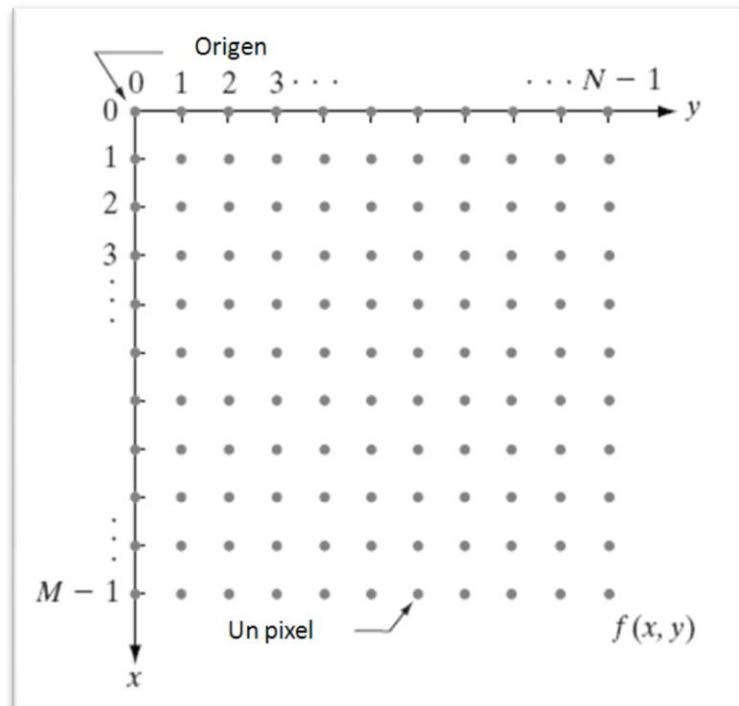


Figura 2.9 Convención de coordenadas usada para representar imágenes, modificada de (Gonzalez y Woods, 2002)

## 2.3 Mejoramiento de la Imagen

El principal objetivo del mejoramiento es procesar una imagen para volverla más adecuada que la imagen original para una aplicación específica, en este sentido, un método que es adecuado para mejorar imágenes de rayos X no necesariamente es el apropiado para mejorar imágenes de la costa.

Es posible mejorar la imagen en el dominio espacial y en el dominio de la frecuencia. El dominio espacial se refiere a la que se da en el plano mismo, la manipulación directa de los píxeles, mientras que aquellas implementadas en el dominio de la frecuencia se basan en modificar la transformada de Fourier de la imagen. El desarrollo del presente trabajo se limita al dominio espacial por lo que solo se profundizará en dichas técnicas.

Las transformaciones más sencillas son las que se dan en los niveles de grises, se dividen en lineales, logarítmicas y de potencia. Su comportamiento se esquematiza en la Figura 2.10. La transformación negativa lineal produce el equivalente a un negativo de fotografía (véase Figura 2.11), esta es adecuada cuando el negro es dominante en la imagen. Las transformaciones logarítmicas aumentan el rango de niveles bajos de gris, lo opuesto sucede con niveles altos de gris. Las transformaciones de potencia son conocidas como corrección gamma, al variar el coeficiente de la potencia se puede obtener una familia de curvas con las que se puede modificar la imagen dependiendo de la necesidad.

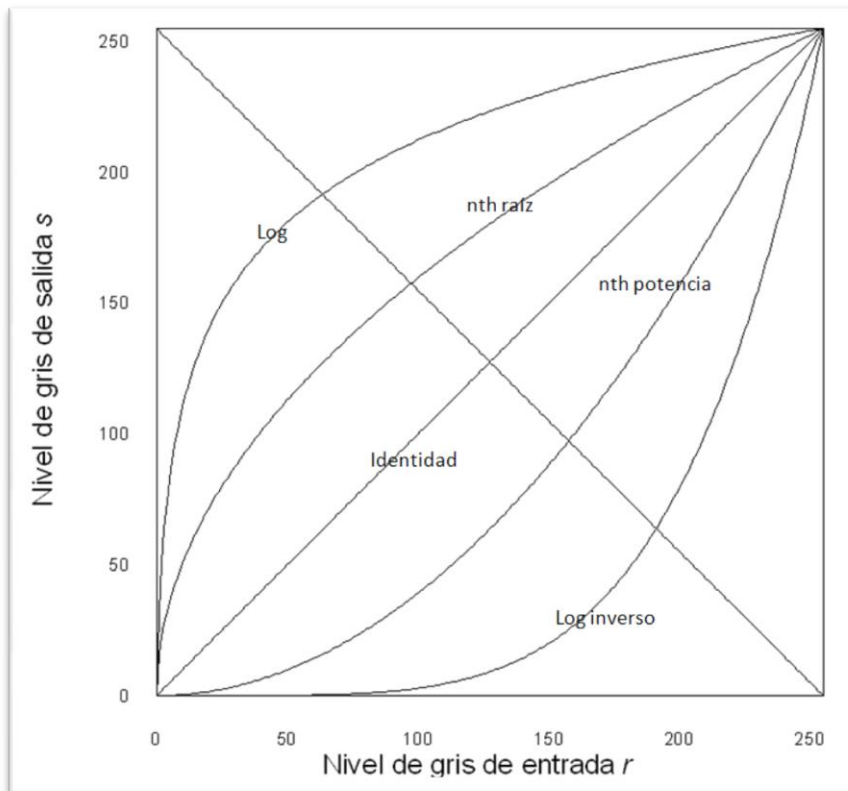


Figura 2.10 de las funciones básicas en los niveles de gris, tomada de (Gonzalez & Woods, 2002)

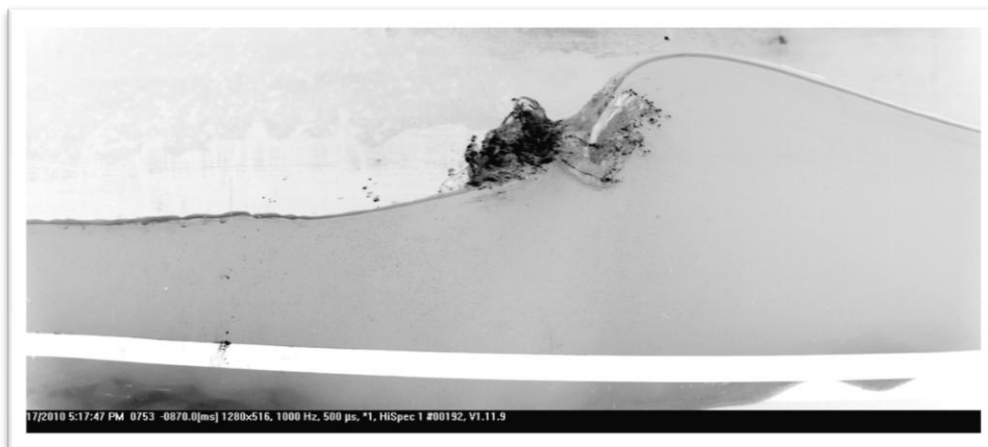


Figura 2.11 Imagen Negativa de una ola

El histograma de una imagen digital es muy útil para conocer la distribución de los niveles de grises. Las imágenes contrastantes cubren un amplio rango de niveles con pocos picos y tienen una distribución uniforme, para imágenes que no son contrastantes (véase Figura 2.12) la ecualización del histograma puede ser muy útil ya que automáticamente distribuye los niveles de gris de manera uniforme (véase Figura 2.13).

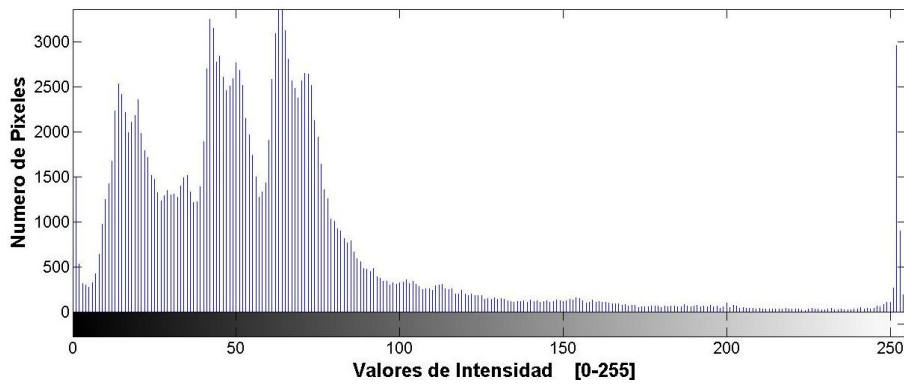
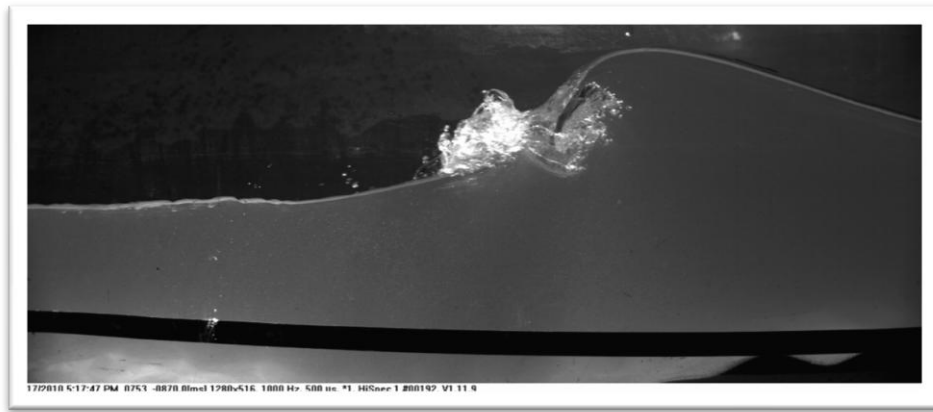


Figura 2.12 El histograma de una ola rompiendo.

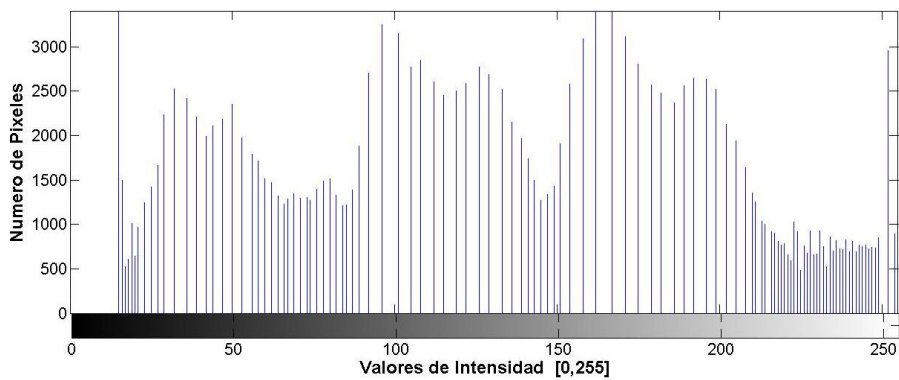
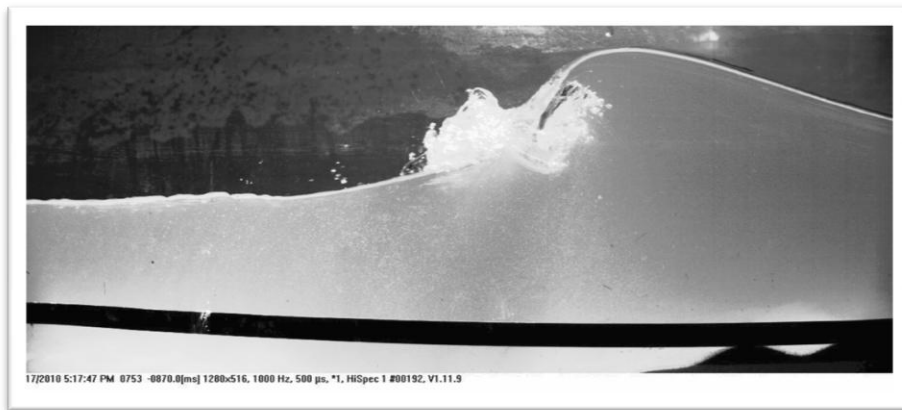


Figura 2.13 El histograma ecualizado y su efecto sobre la imagen.

Las operaciones aritméticas/lógicas son aquellas que se realizan entre píxeles equivalentes de 2 o más imágenes. Las operaciones lógicas como AND, OR y NOT generalmente se usan para delimitar regiones de interés. Las operaciones aritméticas básicas son la suma, resta, división y multiplicación, cuyas combinaciones definen imágenes promedio, varianza, etc. Estas operaciones son útiles para diferenciar el fondo del objeto, eliminar ruido en las imágenes, realizar flujo óptico o resaltar características importantes, por ejemplo, en la Figura 2.14 se puede observar el promedio y la varianza estimada de una serie de 300 imágenes distintas de playa Marlin en Cancún, en la que los píxeles en los que el canal rojo es predominante indican la presencia de barras que provocan la rotura del oleaje.

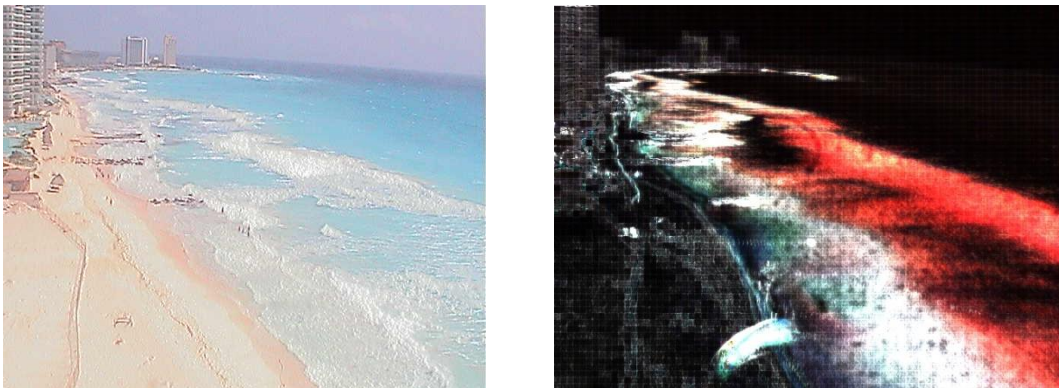


Figura 2.14 Imagen Instantánea y la Varianza de 300 imágenes, tomadas desde el hotel Royal Sunset.

Se pueden realizar otro tipo de operaciones en el dominio espacial como el filtrado, que opera con los píxeles de vecindad y los correspondientes valores de una sub-imagen. También se conoce al filtro como máscara, kernel, template o ventana, los valores en una sub-imagen filtro se conocen como coeficientes en lugar de píxeles.

Hay distintos tipos de filtros. Los filtros pasa-bajos atenúan los componentes de altas frecuencias y resultan en imágenes borrosas. Los filtros pasa-altos atenúan los componentes de bajas frecuencias y resultan en el realce de bordes y otros detalles marcados o finos. Los filtros pasa-banda remueven una fracción de frecuencias seleccionada entre frecuencias bajas y altas, se usa normalmente para restauración de imágenes.

El proceso de filtrado espacial consiste en mover el filtro punto a punto sobre la imagen. En cada punto  $(x,y)$  la respuesta del filtro en ese punto se calcula utilizando una relación predefinida.

Para una máscara de tamaño  $3 \times 3$ , el resultado,  $R$ , del filtro lineal con el filtro con origen en el punto  $(x,y)$ , Ecuación 2.1, es la suma de los productos de los coeficientes del filtro con los píxeles correspondientes de la imagen que están directamente bajo la máscara,  $w(s,t)$ , este paso se conoce también como convolución:

Ecuación 2.1

$$R = w(-1, -1)f(x - 1, y - 1) + w(-1, 0)f(x - 1, y) + \dots + w(0, 0)f(x, y) + \dots \\ + w(1, 0)f(x + 1, y) + w(1, 1)f(x + 1, y + 1)$$

Se asegura que la máscara sea de tamaño impar para poder centrar la máscara en el punto  $(x, y)$  el filtrado lineal de una imagen  $f$  de tamaño  $M \times N$  con un filtro de tamaño  $m \times m$  está dado por:

Ecuación 2.2

$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t)f(x + s, y + t)$$

La convolución puede ser derivativa, cuya respuesta es proporcional al grado de discontinuidad de la imagen en el punto en la que se aplica, por lo tanto, realiza los bordes y atenúa las áreas que tienen baja variabilidad. Las derivadas de primer orden responden a bordes gruesos, las segundas derivadas tienen una respuesta mayor a detalles finos.

Las primeras derivadas en el procesamiento de imágenes son implementadas utilizando la magnitud del gradiente. Para una función  $f(x, y)$  el gradiente está definido como el vector columna en dos dimensiones:

Ecuación 2.3

$$\nabla F = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

Y la magnitud es:

Ecuación 2.4

$$\nabla f = \text{mag}(\nabla f) = [G_x^2 + G_y^2]^{1/2}$$

Las derivadas parciales no son invariantes a la rotación pero la magnitud del vector gradiente sí, a pesar de no ser estrictamente correcto, a la magnitud se le conoce como gradiente. El operador Sobel utiliza un peso de 2 y se usa para obtener un mayor suavizado dando más importancia a los píxeles centrales.

1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

Figura 2.15 Operador gradiente Sobel en X y en Y

Los filtros que son invariantes a la rotación se llaman isotrópicos, el operador derivativo isotrópico más simple es el Laplaciano, Ecuación 2.5, y la máscara tiene la forma de la Figura 2.16. En la Figura 2.17 se puede apreciar el resultado de aplicar el gradiente y el Laplaciano a una imagen en escala de grises.

Ecuación 2.5

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

0	1	0
1	-4	1
0	1	0

1	1	1
1	-8	1
1	1	1

Figura 2.16 Laplaciano simple y tomando en cuenta las diagonales, de tamaño 3x3.



Figura 2.17 Gotas de lluvia en un vidrio (izq.), el Gradiente (centro) de la imagen y el Laplaciano (der.), tomada de Michael Kiel.

## 2.4 Segmentación

La segmentación divide una imagen en sus regiones u objetos. El nivel de la subdivisión a realizar depende del problema a resolver. La segmentación de imágenes no triviales es una de las tareas más laboriosas en PDI. La precisión de la segmentación determina el eventual éxito o fracaso de la mayoría de los análisis.

Los algoritmos se basan en una de dos propiedades básicas de valores de intensidad: discontinuidad o similitud. La primera se basa en partir la imagen en cambios abruptos de la intensidad. La segunda se basa en partir la imagen en regiones que son similares, de acuerdo con criterios predefinidos. El resultado de la segmentación son imágenes binarias a las que generalmente se aplican Operaciones Morfológicas, sub capítulo 2.5.

Debido a sus propiedades intuitivas y simplicidad de implementación, la técnica de umbrales tiene una posición muy relevante en la segmentación de imágenes. Se supone un histograma de grises correspondiente a una imagen  $f(x,y)$  que se compone de objetos con luz en un fondo oscuro, de tal manera que se agrupan los pixeles en 2 grupos:

Ecuación 2.6

$$\begin{array}{ll} f(x,y) < T & \text{representa al objeto} \\ f(x,y) > T & \text{representa el fondo} \end{array}$$

O viceversa, dependiendo del tipo de fondo, en la Figura 2.18 se muestra una huella digital que tiene bajos niveles de gris, mientras que el fondo aparenta ser más brillante.

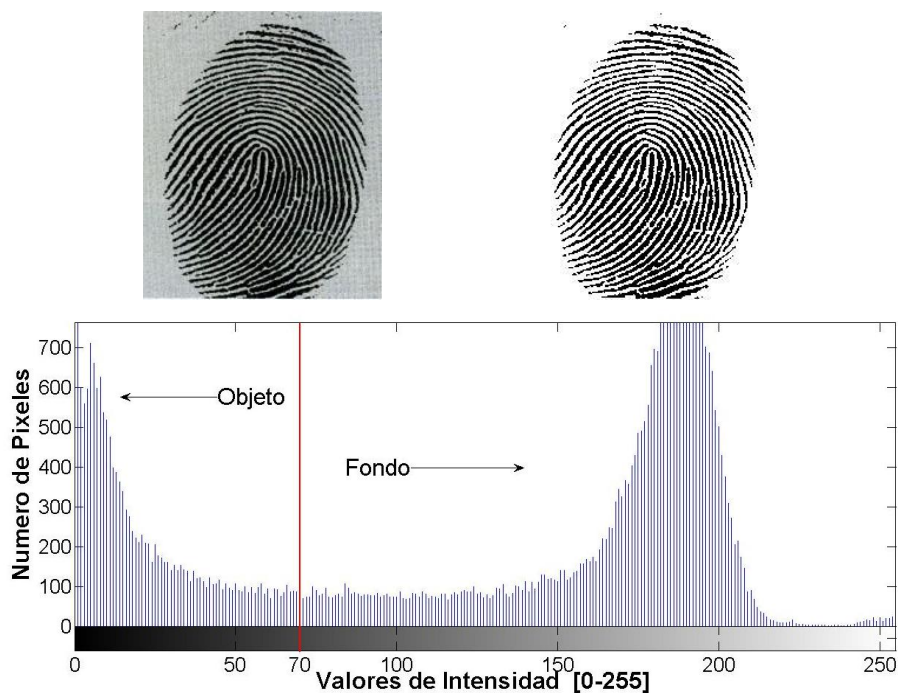


Figura 2.18 Imagen de una huella digital, se muestra el histograma. Se binariza la imagen con un umbral de 70.



Para obtener un umbral adecuado se puede fijar un valor  $T$  inicial, se segmenta y se obtienen 2 grupos de píxeles,  $G_1$  y  $G_2$ , se calculan los niveles de gris promedios de cada grupo, se toma como nuevo umbral el promedio de los promedios y se repite el proceso hasta que  $T$  cambie muy poco.

El umbral anterior solo funciona cuando el histograma tiene 2 modos bien separados, por lo que una iluminación irregular puede complicar el problema. Para evitarlo se divide la imagen en sub-imágenes y se escoge un umbral diferente para cada sub-imagen. La subdivisión escogida en la Figura 2.19 fue de 71 píxeles.

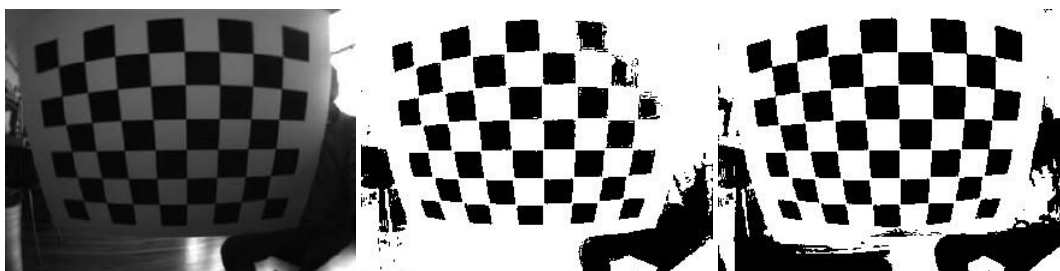


Figura 2.19 Imagen original, imagen segmentada con umbral simple y con umbral adaptativo.

El objetivo de la segmentación es dividir la imagen en regiones, por lo que se pueden usar métodos más directos. Para lograrlo se divide la imagen en regiones conexas, existen 2 enfoques principales: crecimiento por regiones y división-fusión.

En el crecimiento por regiones se empieza a partir en regiones pequeñas y se extienden, se empieza con puntos “semillas” y se agregan puntos vecinos que tengan características similares. Para el criterio de similitud es posible usar niveles de gris, color, textura, movimiento o forma y tamaño de la región formada.

A diferencia del crecimiento por regiones, la división-fusión divide la imagen en regiones hasta que cada región tenga características homogéneas. Luego fusiona las regiones vecinas con características similares.

Otra forma de segmentar es la divisoria de aguas, donde una imagen puede ser vista como un mapa topográfico donde los niveles de gris representan una altura, se explica la delimitación de las regiones usando una analogía hidrográfica. Si un manantial surge en un punto cualquiera, el agua fluye hacia un punto que corresponde a un mínimo local, la cuenca. Para encontrar la divisoria se llenan las cuencas subiendo progresivamente el nivel del agua, cuando 2 cuencas están a punto de fusionarse se construye la divisoria (véase Figura 2.20).



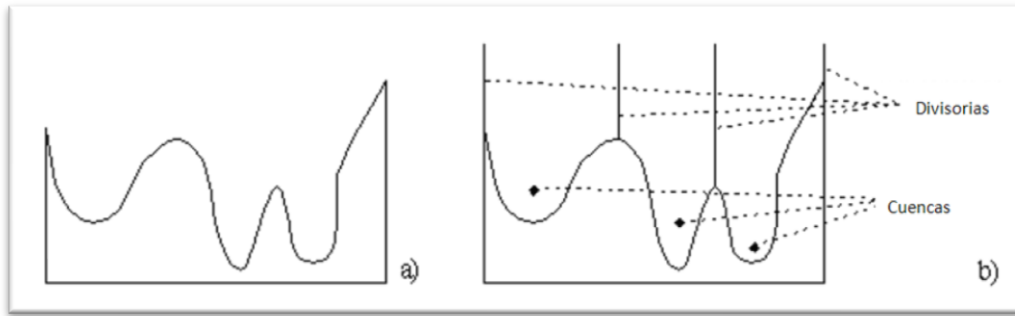


Figura 2.20 a) Histograma de una imagen b) Segmentación por divisoria, locales mínimos representan la cuenca, niveles máximos al dique de la divisoria.

El movimiento es una poderosa clave usada por los humanos y los animales para extraer objetos de interés de un fondo de detalle irrelevante. En aplicaciones de imágenes, el movimiento se deriva de un desplazamiento relativo entre la escena vista y el sistema de captura.

Uno de los enfoques más simples para detectar cambios entre dos cuadros de imágenes tomadas en el tiempo  $t_1$  y  $t_2$ , respectivamente, es comparar la imagen pixel por pixel. Una forma de hacerlo es formar una imagen de diferencias. Si se tiene una imagen que contiene solamente componentes estacionarios y se resta esta imagen de una tomada con el mismo fondo pero con un objeto en movimiento, el resultado será la cancelación de los objetos estacionarios dejando pixeles, no igual a 0, pertenecientes a los componentes no estacionarios.

Una imagen diferencial entre dos imágenes tomadas en los tiempos  $t_1$  y  $t_2$  se define como:

Ecuación 2.7

$$d_{ij}(x, y) = \begin{cases} 1 & \text{si } |f(x, y, t_1) - f(x, y, t_2)| > T \\ 0 & \text{de otra manera} \end{cases}$$

Donde  $T$  es un umbral, así todos los pixeles que tengan valor de 1 se considerarán en movimiento, este procedimiento funciona correctamente si la iluminación es constante relativa a los valores de  $T$ . Hay pixeles con valores 1 que son resultado de ruido, generalmente estos componentes son aislados haciendo un análisis de componentes conectados.

Aunque este ruido se puede atenuar, este proceso también puede remover objetos pequeños. Una manera de mejorar el filtrado es considerar cambios en el pixel de varios cuadros. La idea es ignorar cambios que ocurren esporádicamente en una secuencia de cuadros y que pueden ser atribuidos a ruido.

Considerando una secuencia de imágenes y siendo la primera el cuadro de referencia, una imagen de diferencias acumulativa se forma al comparar el cuadro de referencia con los demás. Un contador por cada pixel en la imagen acumulativa es incrementado cada vez que una diferencia

ocurre, así da el número de veces que el pixel difiere en la secuencia. Con este procedimiento se atenúan los problemas de iluminación no constante, que se da en ambientes no controlados.

## 2.5 Procesamiento Morfológico

La palabra morfología se asocia con la rama de la biología que trata con la forma y estructura de animales y plantas, ese significado se puede extrapolar matemáticamente para la extracción de componentes de la imagen que son útiles para la representación y descripción de regiones. Las 2 operaciones fundamentales del procesamiento Morfológico son la dilatación y la erosión, que generalmente se realizan sobre imágenes binarias.

Siendo  $A$  y  $B$  conjuntos en un espacio bidimensional, la dilatación de  $A$  por  $B$ , es:

Ecuación 2.8

$$A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \emptyset\}$$

La Ecuación 2.8 se basa en obtener la reflexión, que es la proyección tipo espejo de  $B$  respecto a su origen y moverla por  $z$ . La dilatación de  $A$  por  $B$  es un conjunto de todos los desplazamientos,  $z$ , tal que la reflexión y  $A$  coinciden por lo menos en un elemento. El proceso es análogo a la convolución. Al conjunto  $B$  se le conoce como elemento estructural.

Siendo  $A$  y  $B$  conjuntos en un espacio bidimensional, la erosión de  $A$  por  $B$ , es:

Ecuación 2.9

$$A \ominus B = \{z \mid (B)_z \subseteq A\}$$

La ecuación indica que la erosión de  $A$  por  $B$  es el conjunto de todos los puntos  $z$  tal que  $B$ , trasladado por  $z$ , es contenido por  $A$ .

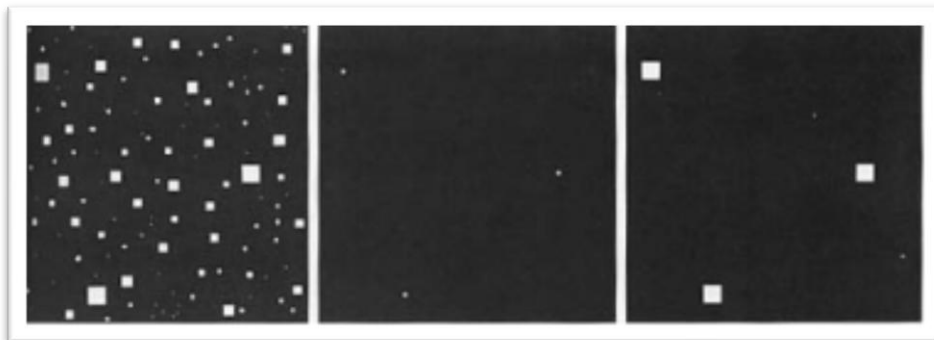


Figura 2.21 Imagen de cuadros de tamaños 1, 3, 5, 7, 9, 15 pixeles (izquierda), erosión de la imagen con un cuadro de 13 pixeles (en medio) y dilatación de la imagen erosionada con el mismo elemento (derecha), tomada de (Gonzalez y Woods, 2002).

En palabras simples la dilatación expande una imagen y la erosión la encoge. Las demás operaciones morfológicas se derivan de una combinación de estas dos y de operaciones lógicas y de conjuntos. La operación apertura, que suaviza el contorno de un objeto, rompe estrechos istmos y elimina delgadas protuberancias, consta de una erosión seguida de una dilatación. La operación cerrado también tiende a suavizar secciones del contorno pero elimina pequeños huecos y llena espacios en el contorno, consta de una dilatación seguida de una erosión. La transformación pega-o-falla se usa para la detección de formas, usa dos elementos estructurales, el segundo es el contorno del primero (cualquiera que sea este) con un ancho de por lo menos un pixel. La operación se puede ver como la intersección del conjunto  $C$  (la erosión de  $A$  por el primer elemento) con el conjunto  $D$  (el complemento de  $A$  erosionado por el segundo elemento, véase Ecuación 2.10).

Ecuación 2.10

$$A \otimes B = \{(A \ominus B_1) \cap [A^c \ominus B_2]\}$$

Entre los usos prácticos de las operaciones morfológicas se encuentra la extracción de componentes que son útiles para la representación y descripción de la forma. En particular se extraen contornos, componentes conectados, cascos convexos y esqueletos de la región. Para información detallada se puede consultar el capítulo 9 de (Gonzalez y Woods, 2002).

## 2.6 Procesamiento en Color

El uso de color en el procesamiento es motivado por dos factores principales. Primero, el color es un poderoso descriptor que simplifica la identificación y extracción de objetos. Segundo, los humanos pueden discernir miles de tonos de color e intensidades, comparados con solamente doce tonos de gris.

El proceso del cerebro para percibir e interpretar los colores, todavía no es comprendido en su totalidad, es un fenómeno físico-sicológico cuya naturaleza física puede ser comprendida por resultados teóricos y experimentales.

En 1666, Sir Isaac Newton descubrió que cuando un rayo de luz solar pasa a través de un prisma de vidrio, Figura 2.22, el rayo que sale no es blanco sino consiste de un espectro continuo de colores desde violeta a rojo.

Básicamente, los colores que los humanos percibimos en un objeto son determinados por la naturaleza del rayo que el objeto refleja. A un cuerpo que refleja luz que se encuentra balanceada en todas las longitudes de onda visibles lo percibimos como blanco. Una imagen acromática, vacía de color, es aquella que se aprecia en blanco y negro.

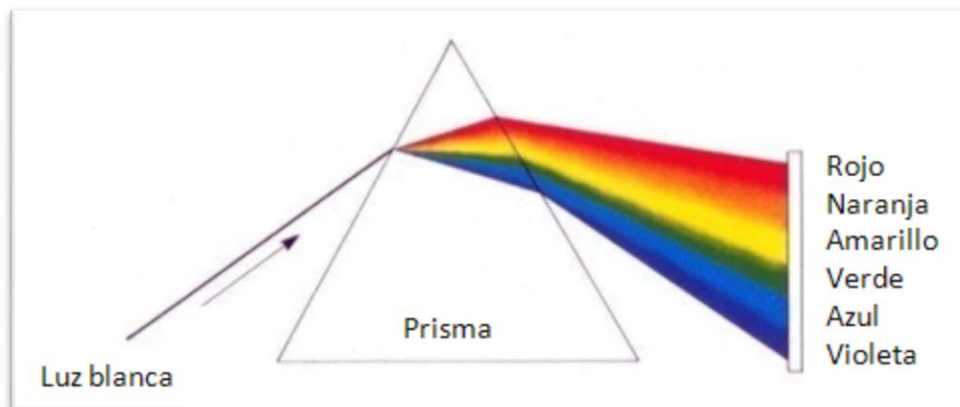


Figura 2.22 Espectro de color visto por un rayo de luz blanco a través de un prisma.

Un modelo que se usa para comprender el funcionamiento de la percepción del color es el RGB (rojo, verde y azul) debido a que 65% de los conos, del ojo humano, son sensibles al rojo, 33% al verde y 2% al azul. El modelo se basa en coordenadas cartesianas. En el cubo RGB, Figura 2.23, los 3 colores en las esquinas son los colores primarios y en las contra esquinas los colores secundarios (cian, magenta y amarillo), la esquina no visible es la negra, si se dibuja una línea de la esquina negra a la blanca se tienen las intensidades de gris.

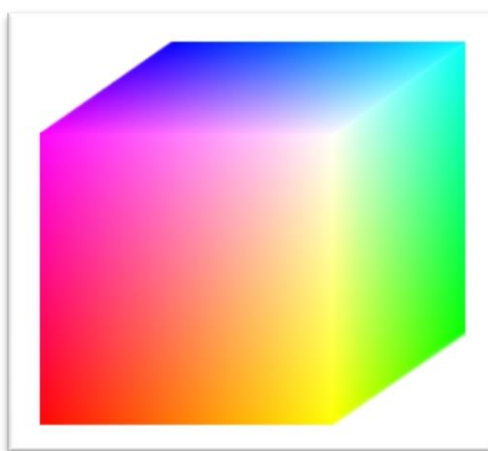


Figura 2.23 Cubo de color RGB

Ya que la mayoría de las cámaras usa este modelo para captar los colores, las imágenes se pueden ver como 3 matrices sobre puestas, cada una correspondiente a la componente de un color, puede llegar a ser muy útil cuando se identifican objetos, por ejemplo en la Figura 2.24 se observa que son fácilmente identificables las fresas si se mezcla la componente roja con la negativa de la componente verde.



Figura 2.24 Una imagen de varias frutas y su división en los 3 colores primarios.

Desafortunadamente el modelo RGB no es adecuado para describir colores en términos prácticos para la interpretación humana. Esto es, el color de un objeto cualquiera nunca se expresa como los porcentajes de cada uno de los colores primarios.

Cuando se observa un objeto, se le describe por su tono (Hue), saturación (Saturation) y brillo (Brightness). El tono es un atributo que se asocia con la longitud de onda dominante en una mezcla de ondas de luz, representa el color dominante que se percibe. La saturación se refiere a la relativa pureza o cantidad de luz blanca mezclada con un tono, los colores puros del espectro están totalmente saturados, el grado de saturación es inversamente proporcional a la cantidad de luz añadida. El brillo es un descriptor subjetivo que es prácticamente imposible de medir, por lo que se asocia con la intensidad (Intensity), que es el descriptor más útil en imágenes monocromáticas.

El modelo HSI, Figura 2.25, transforma las matrices del modelo RGB en las matrices de tono, saturación e intensidad. Separa la intensidad de una imagen de la información de color. Es una herramienta que se usa para desarrollar algoritmos basados en la descripción de color que es natural e intuitiva para los humanos.

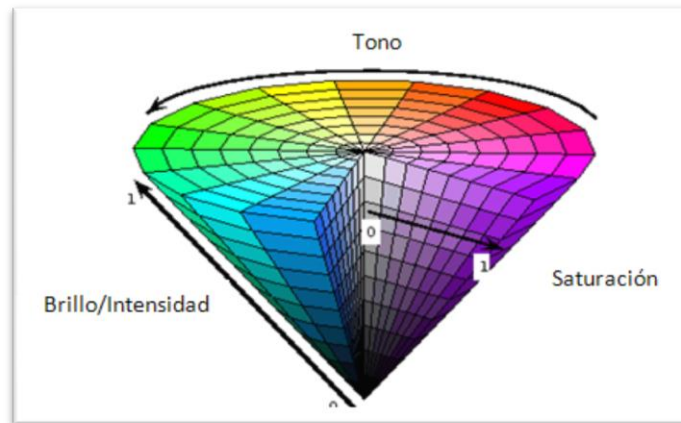


Figura 2.25 Modelo de Color HSI.

Para obtener la componente H de una imagen en formato RGB, cada pixel debe convertirse con las siguientes ecuaciones, los valores RGB deben ser normalizados previamente:

Ecuación 2.11

$$H = \begin{cases} \theta & \text{si } B \leq G \\ 360 - \theta & \text{si } B > G \end{cases}$$

Con

Ecuación 2.12

$$\theta = \cos^{-1} \left\{ \frac{\frac{1}{2}[(R - G) + (R - B)]}{[(R - G)^2 + (R - B)(G - B)]^{1/2}} \right\}$$

La componente de la saturación está dada por:

Ecuación 2.13

$$S = 1 - \frac{3}{(R + G + B)} [\min(R, G, B)]$$

Finalmente, la componente de la intensidad es:

Ecuación 2.14

$$I = \frac{1}{3}(R + G + B)$$

Para regresar al espacio RGB, las ecuaciones a usar dependen de la zona en la que se encuentre H, hay 3 intervalos de  $120^\circ$  y comienza en la zona roja.

Sector RG ( $0^\circ \leq H < 120^\circ$ ):

Ecuación 2.15

$$B = I(1 - S)$$

Ecuación 2.16

$$R = I \left[ 1 + \frac{S \cos H}{\cos(60^\circ - H)} \right]$$

Ecuación 2.17

$$G = 1 - (R + B)$$

Sector GB ( $120 \leq H < 240^\circ$ ):

Ecuación 2.18

$$H = H - 120^\circ$$

Ecuación 2.19

$$R = I(1 - S)$$

Ecuación 2.20

$$G = I \left[ 1 + \frac{S \cos H}{\cos(60^\circ - H)} \right]$$

Ecuación 2.21

$$B = 1 - (R + G)$$

Sector GB ( $120 \leq H < 240^\circ$ ):

Ecuación 2.22

$$H = H - 120^\circ$$

Ecuación 2.23

$$R = I(1 - S)$$

Ecuación 2.24

$$G = I \left[ 1 + \frac{ScosH}{\cos(60^\circ - H)} \right]$$

Ecuación 2.25

$$B = 1 - (R + G)$$

En la Figura 2.26 se muestra la variedad de frutas y sus componentes de tono, saturación e intensidad mostradas monocromáticamente.



Figura 2.26 Los canales HSI de la variedad de frutas.

## 2.7 Calibración de la Cámara

El objetivo de la calibración de la cámara es determinar una serie de parámetros que describen el mapeo entre las coordenadas 3D y 2D. Se determinan geometrías internas de la cámara y características ópticas (parámetros intrínsecos), la posición 3D y orientación de la cámara relativa a un sistema de coordenadas (parámetros extrínsecos).

El problema de la calibración se puede resolver minimizando el error de una función no lineal (Slama, 1980), con simplificaciones del modelo de la cámara o con una combinación de estos métodos. Existen 2 métodos: los explícitos que se basan en las características físicas y los implícitos en donde los parámetros físicos se remplazan por no físicos que se usan para interpolar entre puntos conocidos.

Los parámetros físicos se dividen en extrínsecos e intrínsecos. Los extrínsecos son necesarios para transformar las coordenadas del objeto a las coordenadas de la cámara. El modelo de la cámara pinhole es simple, sin embargo, es de gran utilidad para comprender mejor el proceso de calibración, se basa en el principio de la colinealidad, cada punto en el espacio es proyectado por una línea recta en el plano de la imagen. El origen del sistema de coordenadas está en el centro de la proyección con respecto a las coordenadas del objeto y el eje z del cuadro de la cámara es perpendicular al plano de la imagen. La rotación se representa con los ángulos Euler  $\omega$ ,  $\kappa$  y  $\varphi$ . Para expresar un punto en coordenadas de la imagen  $(X_i, Y_i, Z_i)$  es necesario transformarlas a



coordenadas de la cámara  $(x_i, y_i, z_i)$ . Esta transformación consiste en una traslación y una rotación y puede ser resuelta con la siguiente ecuación:

Ecuación 2.26

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$$

Ecuación 2.27

$$\begin{aligned} m_{11} &= \cos\varphi\cos\kappa & m_{12} &= \sin\omega\sin\varphi\cos\kappa - \cos\omega\sin\kappa & m_{13} &= \cos\omega\sin\varphi\cos\kappa + \sin\omega\cos\kappa \\ m_{21} &= \cos\varphi\sin\kappa & m_{22} &= \sin\omega\sin\varphi\sin\kappa + \cos\omega\cos\kappa & m_{23} &= \cos\omega\sin\varphi\sin\kappa - \sin\omega\cos\kappa \\ m_{31} &= -\sin\varphi & m_{32} &= \sin\omega\cos\varphi & m_{33} &= \cos\omega\cos\varphi \end{aligned}$$

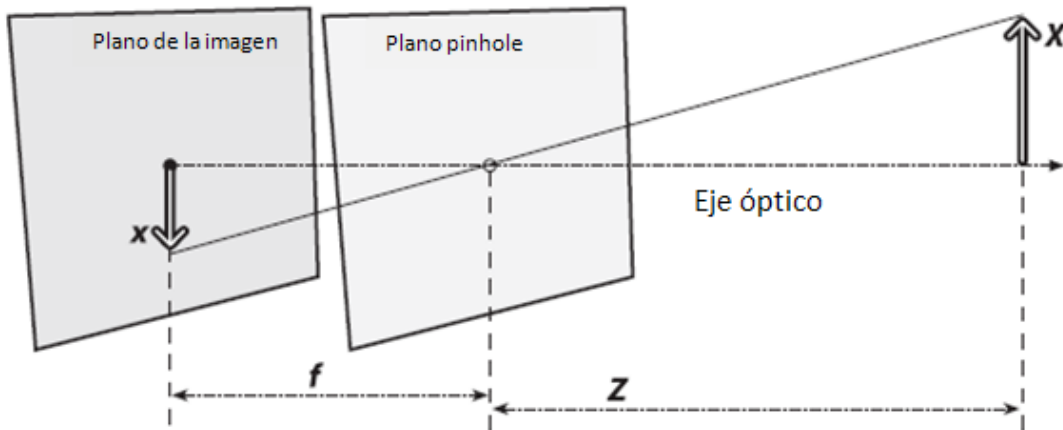


Figura 2.27 Modelo Pinhole

Los parámetros intrínsecos incluyen la longitud focal efectiva,  $f$ , factor de escala,  $s_0$ , y el centro de la imagen  $(u_0, v_0)$  también llamado punto principal. Las unidades de las coordenadas de la imagen están en píxeles por lo que se necesitan los coeficientes  $D_u$  y  $D_v$  para cambiar las unidades a métricas, los coeficientes son linealmente dependientes de la distancia focal y el factor de escala. Usando el modelo pinhole, Figura 2.27, la proyección de un punto al plano de la imagen se expresa:

Ecuación 2.28

$$\begin{bmatrix} \tilde{u}_i \\ \tilde{v}_i \end{bmatrix} = \frac{f}{z_i} \begin{bmatrix} x_i \\ y_i \end{bmatrix}$$

Las correspondientes coordenadas de la imagen  $(u_i', v_i')$  en píxeles se obtienen de la proyección  $(\tilde{u}_i, \tilde{v}_i)$  al aplicar la siguiente transformación:

Ecuación 2.29

$$\begin{bmatrix} u_i' \\ v_i' \end{bmatrix} = \begin{bmatrix} D_u s_u \tilde{u}_i \\ D_v \tilde{v}_i \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}$$

El modelo pinhole es una aproximación, un método útil que permite una formulación matemática entre el objeto y las coordenadas de la imagen. No es válido cuando se necesita alta precisión, por lo que el modelo pinhole constituye una base que se extiende con algunas modificaciones.

Existen correcciones para la distorsión radial del lente que causa al punto de imagen que se desplace radialmente en el plano de la imagen.

Ecuación 2.30

$$\begin{bmatrix} \delta u_i^{(r)} \\ \delta v_i^{(r)} \end{bmatrix} = \begin{bmatrix} \tilde{u}_i (k_1 r_i^2 + k_2 r_i^4 + \dots) \\ \tilde{v}_i (k_1 r_i^2 + k_2 r_i^4 + \dots) \end{bmatrix} \quad \begin{array}{l} k_1, k_2 \dots \text{coeficientes de distorsión radial} \\ r_i = \sqrt{\tilde{u}_i^2 + \tilde{v}_i^2} \end{array}$$

Debido a que los centros de la curvatura de los lentes no son estrictamente colineales se genera una distorsión con componentes radiales y tangenciales. La última se expresa:

Ecuación 2.31

$$\begin{bmatrix} \delta u_i^{(t)} \\ \delta v_i^{(t)} \end{bmatrix} = \begin{bmatrix} 2p_1 \tilde{u}_i \tilde{v}_i + p_2 (r_i^2 + 2\tilde{u}_i^2) \\ p_1 (r_i^2 + 2\tilde{v}_i^2) + 2p_2 \tilde{u}_i \tilde{v}_i \end{bmatrix} \quad \begin{array}{l} p_1, p_2 \text{ coeficientes de} \\ \text{distorsión tangencial} \end{array}$$

Hay otras correcciones para la distorsión lineal o para la distorsión del prisma, que no son tan comunes (Melen, 1994). Un modelo apropiado debe combinar al modelo pinhole con las distorsiones radiales y tangenciales:

Ecuación 2.32

$$\begin{bmatrix} u_i \\ v_u \end{bmatrix} = \begin{bmatrix} D_u s_u (\tilde{u}_i + \delta u_i^{(r)} + \delta u_i^{(t)}) \\ D_v (\tilde{v}_i + \delta v_i^{(r)} + \delta v_i^{(t)}) \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}$$

Los parámetros intrínsecos que son necesarios para la rectificación de las imágenes son la distancia focal, el punto principal, el coeficiente de torsión y los coeficientes de distorsión radiales y tangenciales. Los parámetros extrínsecos, como las rotaciones y traslaciones, no son necesarios ya que el modelo es independiente del factor de escala. En la Figura 2.28 se aprecia claramente en la parte inferior del tablero la distorsión radial y su respectiva corrección.

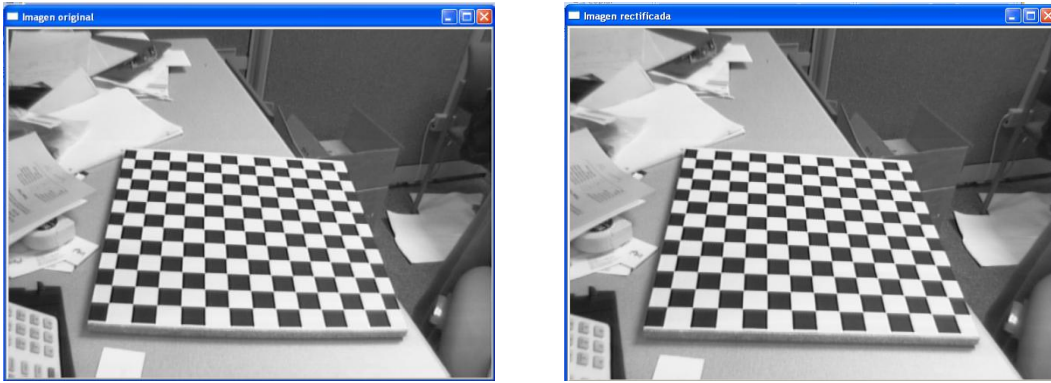


Figura 2.28 Imagen de un tablero de ajedrez sin rectificar (izquierda) y rectificada (derecha).

## 2.8 Transformaciones Geométricas

Muchas aplicaciones involucran dos vistas del mismo objeto tomadas en condiciones diferentes, por lo que el aspecto geométrico del objeto puede cambiar por deformación del objeto, cambio de posición y orientación del objeto. El cambio se traduce por una deformación de la imagen (proyección del objeto).

Las deformaciones se pueden explicar matemáticamente al entender una imagen como un espacio vectorial de dos dimensiones, donde cada punto es un vector  $X = (x, y)^T$  y una función  $f$  transforma cada punto de la imagen original en un punto de la imagen final.

Las transformaciones geométricas definen la relación entre los puntos comunes de dos imágenes. Esta relación puede ser expresada de dos maneras: las coordenadas de la imagen de salida,  $X'$ , pueden ser especificadas como una función de las coordenadas de entrada,  $X$ , o viceversa:

Ecuación 2.33

$$X' = M(X) \text{ o } X = M^{-1}(X')$$

Donde  $M$  especifica la función de mapeo y  $M^{-1}$  su inversa. Las dos expresiones anteriores dan lugar a dos principales tipos de transformaciones espaciales: el mapeo y el mapeo inverso.

El mapeo de un pixel en una imagen de entrada a una imagen de salida tiene el problema que, generalmente, el pixel de la imagen de entrada cae entre los pixeles de la imagen de salida, por lo que una técnica apropiada distribuye el valor de un pixel de entrada en varios pixeles de salida. El procedimiento más fácil es ver pixeles como cuadros y tomar la fracción del área del pixel de entrada que cubre el pixel de salida como el factor de peso. Cada pixel de salida acumula las fracciones correspondientes de los pixeles de salida que se suman para cubrir al pixel entero. El mapeo inverso consiste en revertir el proceso anterior.

Una transformación afín es una transformación de coordenadas lineales que incluye las transformaciones elementales traslación, rotación, escala, estirado y cortado. Puede ser expresado como un vector de adición y una multiplicación de matriz.

**Ecuación 2.34**

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

Es conveniente cuando se trabaja con transformaciones afines usar coordenadas homogéneas; donde un punto en 2D:

**Ecuación 2.35**

$$P = [x, y]^T \text{ son } [sx, sy, s]^T \text{ donde } s \text{ es un factor de escala}$$

Por lo que se puede expresar la transformación afín como:

**Ecuación 2.36**

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Una transformación afín tiene 6 grados de libertad: dos para la traslación ( $t_x, t_y$ ) y uno para la rotación, escala, estirado y cortado. La transformación afín mapea un triángulo en otro triángulo y un rectángulo en un paralelogramo.

La proyección perspectiva es la base de la óptica de imágenes. La transformación afín corresponde a una proyección paralela y su uso se limita a un pequeño campo de visión. Se expresa convenientemente la proyección perspectiva con coordenadas homogéneas:

**Ecuación 2.37**

$$\begin{bmatrix} w'x' \\ w'y' \\ w' \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & 1 \end{bmatrix} \begin{bmatrix} wx \\ wy \\ w \end{bmatrix}$$

Los dos coeficientes adicionales  $a_{31}$  y  $a_{32}$  describen la proyección perspectiva, escrita en coordenadas estándares, la transformación es:

Ecuación 2.38

$$\begin{aligned}x' &= a_{11}x + a_{12}y + a_{13} - a_{31}xx' - a_{32}yy' \\y' &= a_{21}x + a_{22}y + a_{23} - a_{31}xy' - a_{32}yy'\end{aligned}$$

En contraste a la transformación afín, la proyección perspectiva es no lineal. Sin embargo, se reduce a lineal usando coordenadas homogéneas. Una transformación perspectiva mapea líneas en líneas pero solamente las líneas paralelas al plano de proyección permanecen paralelas. Un rectángulo es mapeado en un cuadrilátero.

Generalmente, los coeficientes de una transformación no se conocen por lo que se tiene un conjunto de puntos correspondientes entre el objeto y la imagen. Para una transformada afín se necesitan tres puntos de control no co-lineales, por lo que se tiene que resolver un sistema de ecuaciones lineales de la siguiente forma:

Ecuación 2.39

$$\begin{bmatrix}x'_1 & x'_2 & x'_3 \\y'_1 & y'_2 & y'_3 \\1 & 1 & 1\end{bmatrix} = \begin{bmatrix}a_{11} & a_{12} & t_x \\a_{21} & a_{22} & t_y \\0 & 0 & 1\end{bmatrix} \begin{bmatrix}x_1 & x_2 & x_3 \\y_1 & y_2 & y_3 \\1 & 1 & 1\end{bmatrix} \quad \rightarrow \quad P' = AP$$

La inversa de la matriz P, necesaria para resolver la matriz, existe cuando tres puntos X1, X2, X3 son linealmente independientes. Esto significa, geoméricamente, que no deben encontrarse en una misma línea.

La determinación de los coeficientes de la proyección perspectiva es ligeramente más compleja. Dados cuatro o más puntos correspondientes, los coeficientes se pueden determinar generalizando N puntos de control con un sistema de ecuaciones lineales con 2N ecuaciones y 8 variables desconocidas de la forma:

Ecuación 2.40

$$\begin{bmatrix}x'_1 \\y'_1 \\x'_2 \\y'_2 \\\vdots \\x'_N \\y'_N\end{bmatrix} = \begin{bmatrix}x_1 & y_1 & 1 & 0 & 0 & 0 & -x_1x'_1 & -y_1y'_1 \\0 & 0 & 0 & x_1 & y_1 & 1 & -x_1y'_1 & -yy'_1 \\x_2 & y_2 & 1 & 0 & 0 & 0 & -x_2x'_2 & -y_2y'_2 \\0 & 0 & 0 & x_2 & y_2 & 1 & -x_2y'_2 & -yy'_2 \\\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\x_N & y_N & 1 & 0 & 0 & 0 & -x_Nx'_N & -y_Ny'_N \\0 & 0 & 0 & x_N & y_N & 1 & -x_Ny'_N & -yy'_N\end{bmatrix}$$

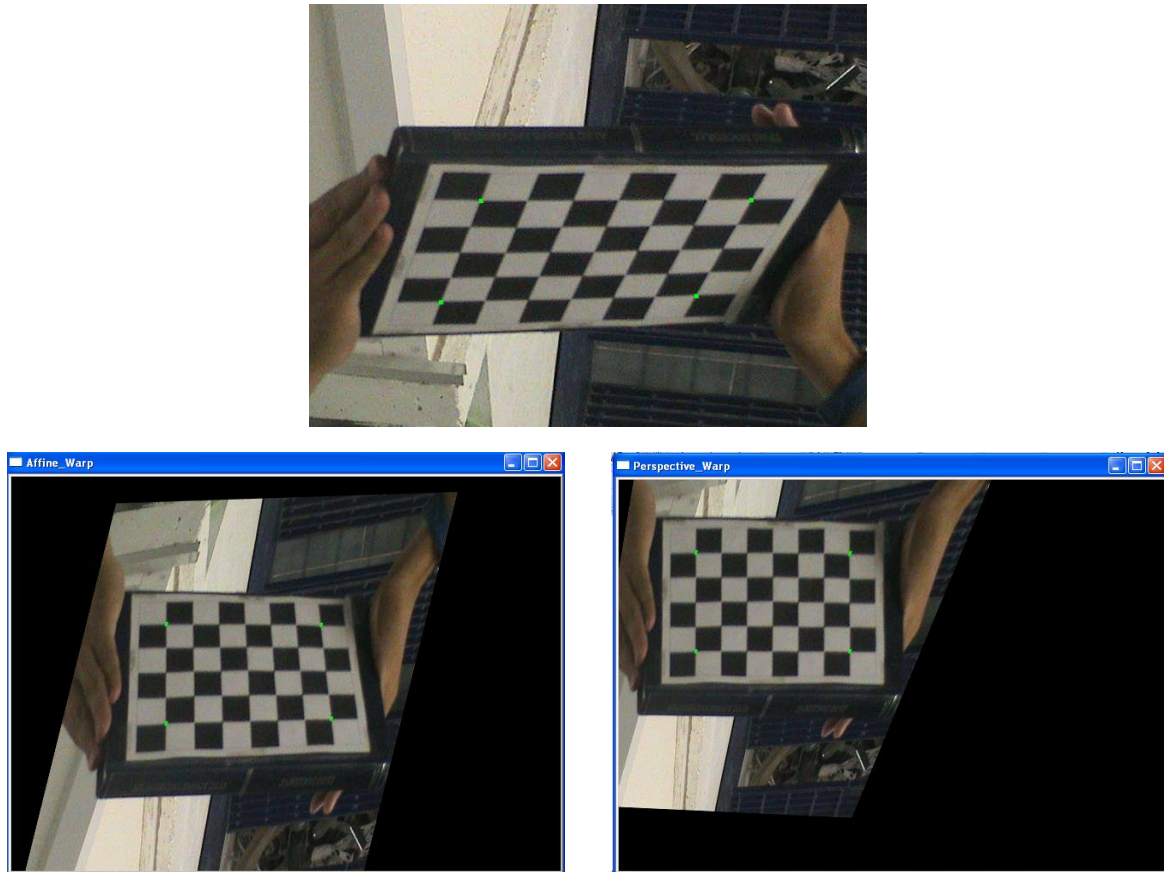


Figura 2.29 Una imagen tipo tablero de ajedrez (arriba), transformación afín (izquierda) y proyección perspectiva (derecha).

En la Figura 2.29 se observa el resultado de aplicar la transformación afín (usando 3 puntos de control) y la proyección perspectiva (4 puntos de control); los puntos de control son los pixeles verdes, la proyección perspectiva tiene un mejor resultado para proyectar el plano del tablero.

## 3 EXTRACCIÓN DE LA LÍNEA DE COSTA

### 3.1 ARGUS

En los últimos 25 años, el uso de cámaras de video para el monitoreo continuo de las costas y estuarios, se ha convertido en una poderosa herramienta para el muestreo periódico del ambiente cercano a la costa. La idea de colocar una video cámara hacia los ambientes costeros, permitió en una primera iteración la adquisición periódica y cualitativa de información que reveló el comportamiento de ciertos rasgos morfológicos característicos de la zona de rompientes ante variaciones en el clima marítimo. Este sistema desarrollado en 1986 por científicos de la Universidad Estatal de Oregon OSU (Oregon State University), es conocido en la actualidad como ARGUS (Holman y Stanley, 2007), nombre que deriva del gigante de la mitología griega caracterizado por tener 100 ojos que todo lo mira y que tenía como misión el cuidado de la ninfa lo (Conner, 2010).

El primer sistema ARGUS instalado en el mundo, se ubicó en Duck, Carolina del Norte, donde se instalaron cámaras para estudiar la variabilidad de las barras que se generan como resultado del balance energético que se presenta en la zona de rompientes. En este primer sistema se grababan 15 minutos de video al día, que luego eran post-procesados de tal suerte que se generaba una imagen digital promedio (timex) con resolución de 512 por 480 pixeles (Crowson *et al.*, 1988). Los datos registrados por esta estación durante el periodo de tiempo de 1986 a 1993 se conocen como ARGUS 0. Este sistema requería el uso de cintas magnéticas para la grabación de las imágenes, las cuales eran recogidas y transportadas del punto de registro hacia la oficina para ser procesadas.

La obtención de datos a una alta frecuencia y de amplia cobertura espacial comenzó a ser de suma importancia, sin embargo, el conjunto de tareas de recoger las cintas, enviarlas y post-procesarlas requería de una considerable cantidad de recursos humanos y de tiempo, por lo que su automatización se hizo necesaria. La primera estación automática se creó en Yaquina Head (Oregon) donde se transmitía la información a OSU, vía internet. La resolución de las imágenes obtenidas era de 640 por 480 pixeles y se le dio el nombre de Argus I.

Así como se implementó el uso de imágenes promedio durante un periodo de tiempo de grabación, a la par también se estableció el empleo de las imágenes de “varianza”, basadas en la desviación estándar de la intensidad registrada en cada pixel de la imagen registrada. En particular, este tipo de imágenes fueron de utilidad para identificar la ubicación de la zona de rompientes y variabilidad temporal y espacial del punto de rotura en una playa. Este tipo de imágenes presentan los pixeles de alta variabilidad con tonos blancos (en imágenes monocromáticas), mientras que los que permanecen más estables son identificados con un color negro. Ambos tipos de imágenes,



promedio y varianza, cambiaron considerablemente el concepto que se tenía de las barras y del comportamiento morfodinámico de la costa en escalas de largo plazo (Lippmann y Holman, 1990). Se empezó a explorar el uso de las series temporales, tomando como unidad de medida el pixel y como herramienta de cálculo el software MatLab. A partir de este momento la capacidad de las computadoras era limitada, lo que llevó a su fin la generación Argus I.

La actualización del sistema, denominada Argus II fue diseñada en torno a un servidor y con cámaras que captaban imágenes en color de 640 por 480 pixeles de resolución, pero de mejor calidad que las del sistema anterior. La nueva limitación era que solo tenía una entrada de video y, por lo tanto, solo podía digitalizar una señal al mismo tiempo. Para incorporar la información del conjunto de cámaras era necesario un conector que permitiera transformar dos entradas en una. El no poder adquirir datos de varias cámaras simultáneamente condujo al inicio de una nueva generación en el 2002, cuando se puso en marcha el proyecto CoastView (Davidson *et al.*, 2007).

La tercera generación fue desarrollada por Irv Elshoff de Delft y John Stanley de la Universidad estatal de Oregon. La resolución de las cámaras de Argus III es de 1024 por 768 pixeles, en escala RGB a 8 bits por canal, todas las cámaras están conectadas directamente al servidor y es posible la adquisición simultánea de datos. En la Figura 3.1 se muestra la configuración actual del sistema Argus (Holman y Stanley, 2007).



Figura 3.1 Configuración del sistema Argus, Egmondaan, Países Bajos.

La Figura 3.2 presenta un ejemplo del tipo de imágenes que se adquieren de forma automática con este sistema, en la que el panel a) presenta la fotografía instantánea tomada de una de las cámaras ubicadas en la estación de Milford on Sea, Reino Unido, el panel b) presenta la imagen promedio y el panel c) ilustra la imagen varianza. De la observación de estas imágenes, tomadas en bajamar, se revelan dos rasgos morfológicos evidentes en esta costa, por un lado la línea de rotura del oleaje (donde se asume la presencia de una barra exterior) y la línea de costa identificada por la variación



de la superficie libre en la zona de lavado, la cual es de mejor identificación en la imagen de varianza.



**Figura 3.2** Imágenes de muestra tomadas por el sistema ARGUS instalado en Milford en Sea, Reino Unido: a) Fotografía instantánea; b) Imagen promedio; c) Imagen varianza

El costo anual de la licencia de mantenimiento del sistema ARGUS (~80,000 pesos/año), ha propiciado que diversas universidades alrededor del mundo consideren más rentable la creación de sistemas propios. Tal y como lo demuestra el desarrollado por el Instituto Neozelandés para la Investigación del Agua y la Atmósfera NIWA, en conjunto con la Universidad de Bourdeaux en Francia conocido como Cam-Era. La Figura 3.3 presenta un ejemplo de una imagen promedio tomado en la playa de Taihura, en la que se aprecian las corrientes de rip generadas en esta playa. Este sistema ahora también comercial, es de menor costo al del sistema ARGUS.



**Figura 3.3** Imagen promedio de muestra tomadas por el sistema Cam-Era instalado en Taihura, \_Nueva Zelanda.

## 3.2 IBIS

En México, el desarrollo de un sistema similar al ARGUS o el Cam-Era está siendo desarrollado por investigadores del Instituto de Ingeniería (Rodolfo Silva) en colaboración con contrapartes del CINVESTAV (Ismael Mariño), a iniciativa del primero. El sistema bautizado con el nombre de IBIS

debe su nombre al ave marina que se caracteriza por ser la última especie en tomar abrigo al ocurrir un huracán y la primera en regresar, una vez que la tormenta ha pasado (Bildstein, 1993).

La diferencia entre los sistemas IBIS y ARGUS/Cam-Era es que los segundos utilizan cámaras fijas y el primero cámaras móviles. Las cámaras SNC-RZ30 cuentan con una robótica interna que les permite cambiar la zona de muestreo con cierta libertad, además de tener un zoom 1x a 25x y una resolución de 736 por 480 píxeles. La cámara tiene un mecanismo de barrido horizontal con un ángulo que varía de  $-170^{\circ}$  a  $170^{\circ}$  y un mecanismo de barrido vertical con un ángulo que varía de  $-90^{\circ}$  a  $25^{\circ}$ . Se programaron para obtener series de imágenes, con frecuencia de 1 hz durante 5 minutos con 14 posiciones distintas, el proceso de la toma de imágenes dura 70 minutos.



Figura 3.4 Sistema IBIS instalado en Cancún.

Cada imagen, en promedio, ocupa 700KB de memoria, por lo que al día se alcanza un total de 2.94GB, al mes se generan 88.2GB, en un año se alcanza un Terabyte de información. Tal cantidad de datos genera necesidades que en parte pueden ser resueltas usando un formato de compresión de imágenes, JPEG, reduciendo el tamaño de cada imagen a 70KB. Los métodos de muestreo se diseñan para reducir aún más la información y hacerla más digerible. Se logra con la producción de imágenes individuales que representen la estadística de las variaciones de intensidad de cada vista, la automatización de este proceso es parte de la programación desarrollada para el presente trabajo, dichas imágenes se describen a continuación.

### 3.3 Tipos de Imágenes

De forma similar a los sistemas ARGUS y Cam-Era, las cámaras utilizadas en el sistema IBIS son programadas para generar al día 3 imágenes por vista: instantánea, promedio y varianza. .

La imagen instantánea, Figura 3.3a, se guarda al inicio de cada muestreo para conocer las condiciones del sitio, no para obtener información cuantitativa sino para una evaluación visual cualitativa.

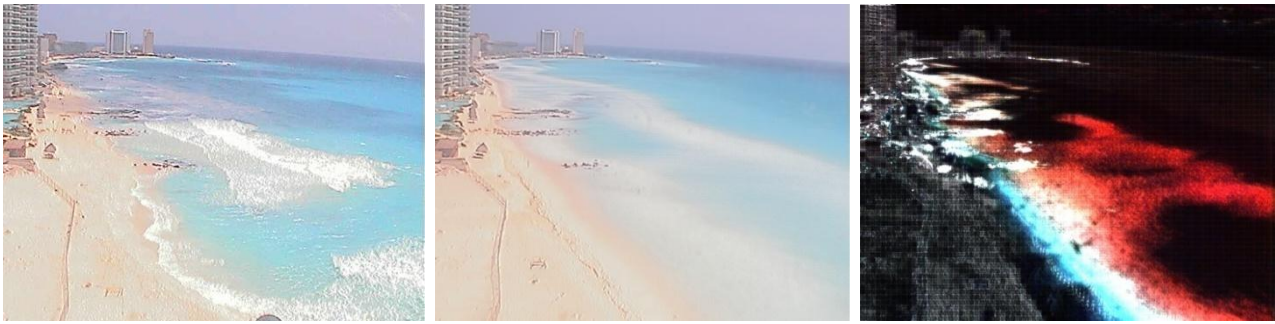


Figura 3.5 Imágenes de muestra tomadas por el sistema IBIS en Cancún, México: a) Fotografía instantánea; b) Imagen Promedio; c) Imagen Varianza.

La imagen promedio, Figura 3.3b, representa la media temporal de toda la serie. Los objetos que no se mueven se ven igual que en la toma instantánea. Su característica principal es la delineación de las olas rompiendo en la zona de surf como bandas blancas. Las barras de arena causan el rompimiento de la ola por lo que estas imágenes pueden ser usadas para encontrar la localización y la morfología de barras y canales de corriente (Lippmann y Holman, 1989).

La imagen varianza, Figura 3.3c, representa la desviación de la intensidad de cada pixel respecto a su promedio. En imágenes a color, se corresponde con la variabilidad respecto a los 3 canales (RGB). En el caso específico de Cancún, a las 9-11 de la mañana, se observa que cuando las olas rompen debido a la barra resalta el canal rojo y en la zona de lavado el azul.

## 3.4 Revisión de los Algoritmos y Modelos Existentes

### 3.4.1 Antecedentes

En la última década ha aumentado el uso de los sensores remotos y las técnicas de video para el estudio de la franja litoral y los procesos costeros. Una de las principales aplicaciones de estas técnicas es la obtención de la configuración del fondo marino en aquellas playas expuestas a rangos de marea de gran magnitud, que en la fase bajamar permite el monitoreo de una zona del fondo que durante esa fase aparece seca. Sin embargo, el problema básico consiste en cómo convertir un espacio bidimensional en uno tridimensional, es por eso que han surgido diversas formas de suplir estas dificultades físicas. Diversos investigadores han solucionado el problema de una manera acertada asumiendo que una de las coordenadas es conocida (Plant y Holman, 1997; Aarninkhof *et al.*, 2003) o en otras ocasiones empleando técnicas de estereoscopia (Holland y Holman, 1993; Holland y Holman, 1997).

Los modelos más comunes para obtener batimetrías intermareales, detectan la línea de costa a partir de imágenes oblicuas diferenciando las zonas de arena de las de agua y estimando la elevación con base en las condiciones hidrodinámicas (marea, oleaje, etc.).

Uno de los principales esfuerzos ha sido desarrollar modelos fiables para detectar la línea de costa, que está definida por Plant y Holman (1997) como la intersección del nivel medio de aguas con la superficie de la playa, en ausencia de olas. Entre los primeros modelos desarrollados se encuentra el conocido como SLIM (ShoreLine Intensity Maximum) que discrimina patrones de intensidad de los píxeles en una escala de grises (Plant y Holman, 1997). En playas reflejantes (con pendientes pronunciadas), los resultados obtenidos por medio de este método fueron calificados como buenos mientras que en playas disipativas (pendientes suaves), la escala de grises limitaba la discriminación de los píxeles, obligando a calibrar y corregir los resultados para cada lugar de aplicación.

Con el paso de los años y la mejora de la tecnología, se incluyeron las imágenes a color, con lo cual se incrementaron las posibilidades para localizar la interfaz agua-tierra, se desarrollaron nuevas técnicas en distintas tesis doctorales como el agrupamiento de los píxeles (Aarninkhof, 2003) o el "RatiosRGB" que trata de resaltar la banda roja a través de un ratio con las demás componentes (Osorio, 2005).

### 3.4.2 Aproximación Sencilla a la detección de la línea de costa

La forma más simple para determinar la posición de la línea de costa de una imagen, consiste en seleccionar los valores de intensidad asociados con las coordenadas (u,v) de cada píxel a lo largo de un perfil y observar las variaciones. En la Figura 3.6 se presenta una idealización del comportamiento de la intensidad de un transecto de píxeles, donde se registra un cambio abrupto en la intensidad en el cambio de agua a tierra (ej. línea de costa). Este tipo de aproximación funciona de forma satisfactoria en ambientes controlados, sin embargo, en playas reales se requiere de algoritmos más sofisticados.

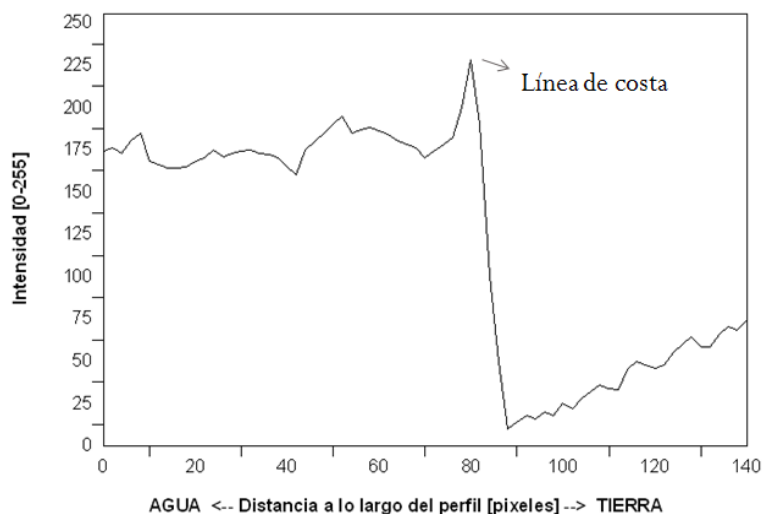


Figura 3.6 Aproximación sencilla, idealizada, para localizar la línea de costa.

### 3.4.3 SLIM (Shore Line Intensity Maximum)

En presencia de oleaje, la identificación del nivel medio del agua (MWL), que define la línea de costa es más complicada como resultado de los movimientos oscilatorios (de ascenso y descenso) que ocurren en la zona de lavado. Es precisamente en estas condiciones que la utilización de imágenes tipo promedio es ideal, dado que en ellas se muestran patrones con áreas blancas, correspondientes a la espuma producida por la rotura del oleaje, estas zonas pueden ser un indicativo de la presencia de barras en la zona de rotura y la identificación de la anchura de la zona de lavado. La transición de esa zona blanca a la zona seca (de arena) permite la fácil y correcta estimación de la línea de costa. A la intensidad de luz que corresponde a la línea de costa se le llama máxima intensidad de línea de costa (SLIM).

Para revisar el funcionamiento y la puesta a punto de este modelo, se aplicó a dos tipos de imágenes una tomada con el sistema ARGUS en la playa de Coolangata ubicada en Australia (véase Figura 3.7, imagen tomada de los archivos de imágenes de Argus <http://argus-data.wldelft.nl/index.html>) y otra tomada con el sistema IBIS en la playa de Cancún, México (véase Figura 3.9).

Se convierten las imágenes de RGB a HSV, donde la única componente de interés es aquella referente a la intensidad de la luz (V).

En la Figura 3.7 la línea roja identifica el transecto de pixeles seleccionado a evaluar, se puede escoger cualquier perfil con cualquier orientación, se escoge este por simplicidad ya que no se requiere manipular la imagen a través de rotaciones.



Figura 3.7 (I) Playa de Coolangata en RGB con el perfil señalado (D) La componente V del espacio HSV

La Figura 3.8 ilustra los resultados obtenidos para el transecto seleccionado por medio del método SLIM, en la que dos crestas son claramente identificables a lo largo del mismo. La primera de ellas corresponde a la franja de arena más clara, mientras que la segunda, con mayor intensidad ( $V=213$ ), identifica la zona de lavado y por ende la línea de costa.



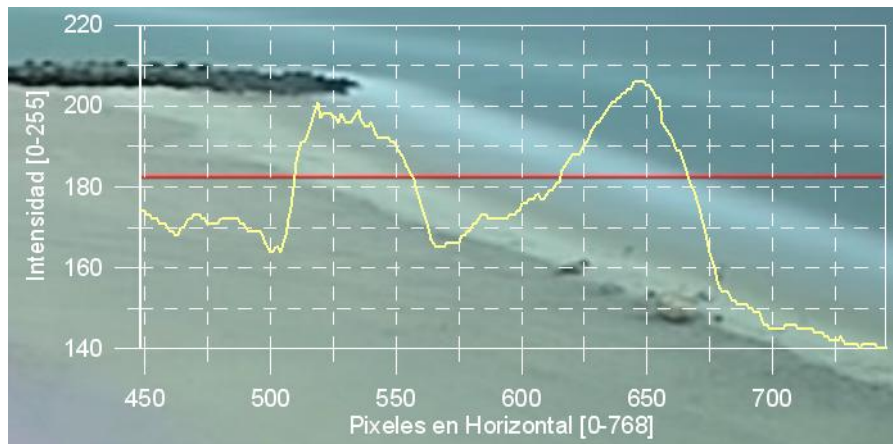


Figura 3.8 Zoom de la zona de interés y su perfil de intensidades.

Por su parte, los resultados de la aplicación del método SLIM al transecto seleccionado en la imagen del sistema IBIS (Figura 3.9), se presentan en la Figura 3.10. Donde se es posible identificar en la parte izquierda de la gráfica un comportamiento caótico. En esta imagen, es notable que el sistema muestra dos valles en las zonas de interés asociadas, rompientes y lavado. La diferencia entre los resultados obtenidos para ambas imágenes ARGUS e IBIS, se deben a la diferencia en la iluminación de la toma y la alta reflectancia de la arena de Cancun (ver Figura 3.9). Es precisamente en esa región donde se presentan los valores de intensidad más grandes para esta imagen.



Figura 3.9 (I) Imagen de Cancún con el perfil señalado (D) La componente V del espacio HSV

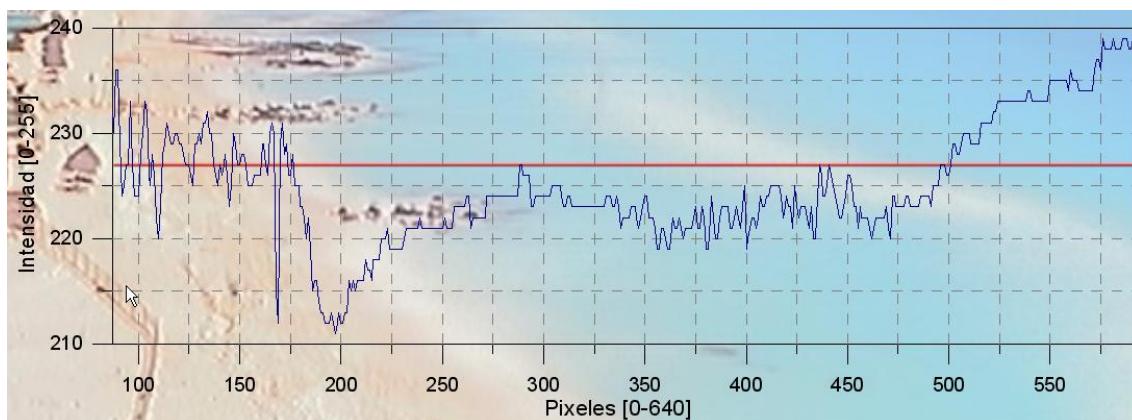


Figura 3.10 Zoom de la zona de interés y su perfil de intensidades.

Las diferencias aquí documentadas entre los resultados obtenidos para imágenes IBIS y ARGUS indican que el método SLIM, para el caso de la imagen de Cancún, no identificaría la línea de costa de forma adecuada, si se buscan los valores de intensidad máximos. Sin embargo, es posible argumentar que si se cambia la condición de búsqueda hacia píxeles de menor intensidad (mínimos) su implementación en Cancún es posible.

Un resultado notable de este análisis indica que la calidad e iluminación de la imagen, la cual está en función del sitio de toma y de la cámara, es dominante para determinar la calidad de los resultados que se obtienen con este método. En lo que resta del capítulo se trabajara solamente con imágenes de Cancún.

#### 3.4.4 Método “RatiosRGB”

Este método consiste en la separación la imagen a color en sus 3 componentes de color Rojo, Verde y Azul (RGB). Al igual que en el método anterior, se selecciona un transecto de análisis y se estima un cociente a partir de la combinación de la información en cada canal. Este procedimiento debe su origen a los fundamentos de teledetección, en particular a los trabajos de Lillesand y Kiefer (1984), quienes demostraron que a partir de un análisis de la banda del rojo las zonas de tierra y las de agua son fácilmente identificables. La expresión matemática que se utiliza para resaltar la banda roja en una imagen se presenta en la Ecuación 3.1, que expresa operaciones matemáticas simples entre las 3 bandas del espacio RGB.

Ecuación 3.1

$$Ratio = \frac{R - G - B}{R + G + B}$$

La Figura 3.11 presenta los resultados obtenidos por medio del método ratios RGB a la imagen tomada por el sistema IBIS (panel derecho Figura 3.9). En esta Figura se representa de forma gráfica el ratio estimado para cada pixel sobre el transecto ilustrado, en la que se observa un incremento al tocar agua pero no existe un gradiente o pico que permita diferenciar la línea de costa del resto de la imagen de forma automática.

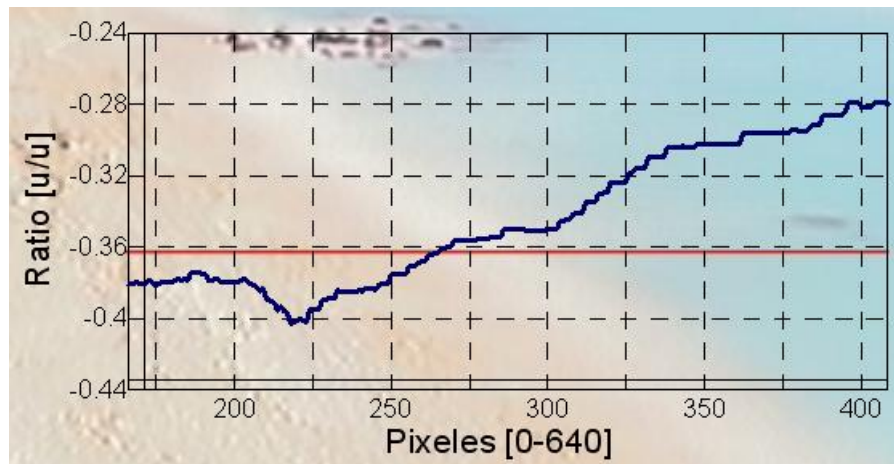


Figura 3.11 El ratio a lo largo de la zona de interés (Cancún).

### 3.4.5 Modelo SDM (Shoreline Detection Model)

El Modelo de detección de la Línea de Costa, (SDM) fue desarrollado por Aarninkhof (2003) para imágenes tomadas con el sistema ARGUS. Este método representa un algoritmo que busca hacer una diferenciación de color entre la zonas de la playa seca y húmeda. Esta técnica está basada en el método PIC (Pixel Intensity Clustering), que transforma el espacio RGB a HSV y utiliza los canales Hue (H), Saturación(S) para realizar el agrupamiento de píxeles. La Figura 3.12, presenta el ejemplo de la separación de la imagen RGB tomada por el sistema IBIS a sus tres componentes HSV.

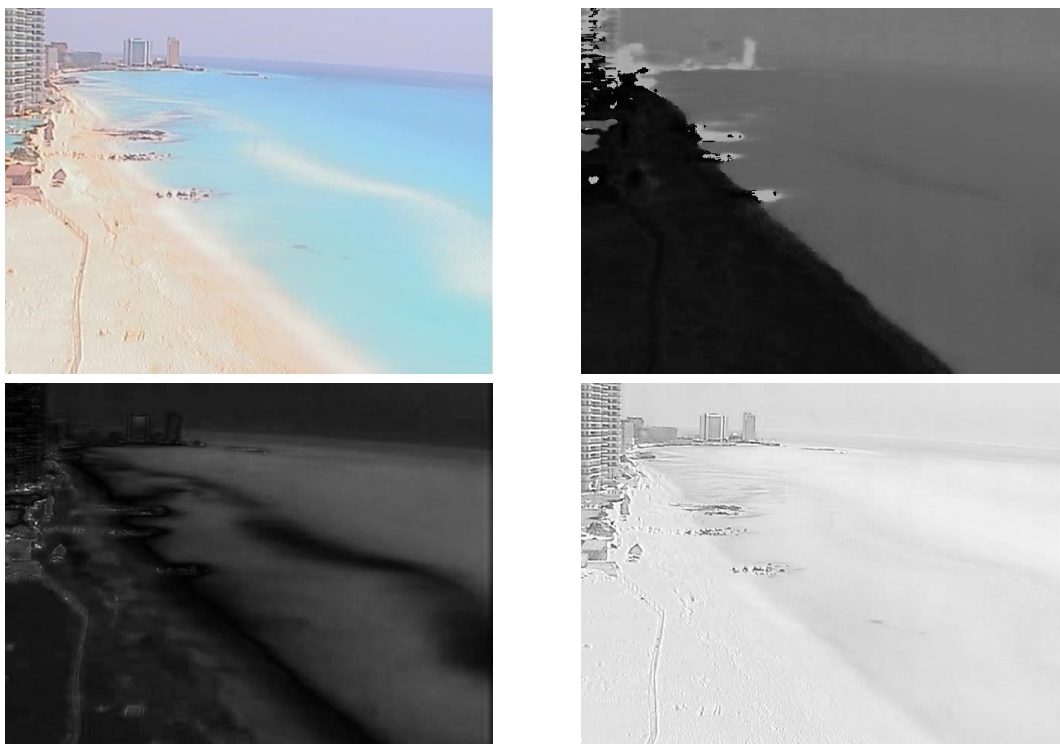


Figura 3.12 Los canales Hue (SD) Saturación (II) e Intensidad de la imagen (SI).



El método consiste en combinar la información del canal Hue (H) y del Saturation (S), de tal suerte que la separación de píxeles entre mojados y secos sea más clara. Para ello se requiere de la definición de una línea divisoria que está representada por una función discriminante que considera los valores registrados en los canales H y S, tal que:

Ecuación 3.2

$$f_{disc}(H, S) = p_1 * H + p_2 - S$$

Esta ecuación depende de dos parámetros  $p_1$  y  $p_2$ , que ajustan la pendiente y la ordenada al origen de la recta definida. En la figura Figura 3.13 se muestra la función discriminante y la separación de los tipos de píxeles.

La utilidad de la ecuación 3.2 consiste en crear una imagen binaria, definiendo a los píxeles negros como la playa y los blancos como el mar.

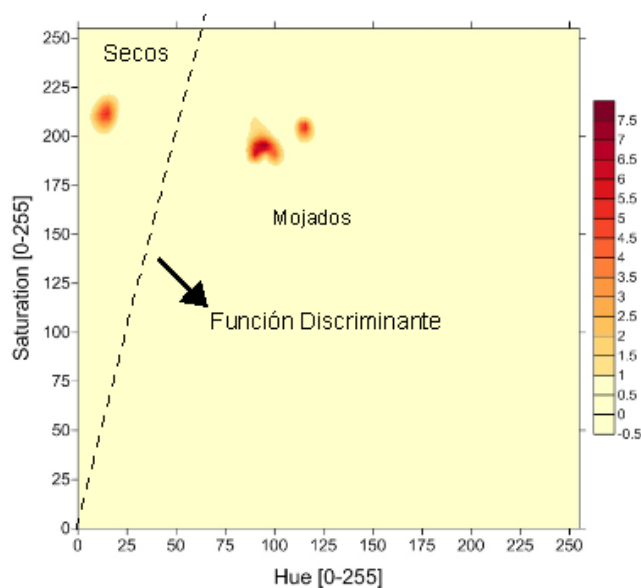


Figura 3.13 Histograma de Intensidades de Píxeles y la función discriminante,  $p_1=5$  y  $p_2=1$ .

De esta forma, se revisa la consistencia espacial, usando distintas operaciones como los operadores morfológicos de apertura y cierre, descritas en la sección 2.5, principalmente respecto a sus vecinos más cercanos (véase Figura 3.14). Por último, el método indica la utilización de un detector de bordes a la imagen resultante, obteniéndose un resultado final como el que se muestra en la Figura 3.15.



Figura 3.14 a) Imagen binarizada; b) Imagen resultante de aplicar operadores morfológicos; c) Imagen resultante de aplicar un detector de bordes.



Figura 3.15 La línea de costa resultante con el modelo SDM, parámetros  $p_1=5$  y  $p_2=1$

Diversos investigadores han intentado diferentes combinaciones de las intensidades HSV y RGB para crear nuevas funciones discriminantes, dando los mejores resultados Hue-Saturation y Hue-Green (Caljouw, 2000). La primera combinación trabaja mejor bajo condiciones de calma mientras que la segunda trabaja mejor bajo condiciones energéticas, cuando hay mucha disipación por rotura en la zona de cercanías de la costa (zonas blancas).

El inconveniente de este método consiste en la definición automática de los parámetros  $p_1$  y  $p_2$ , que deben ajustarse de tal suerte que la función discriminante divida los grupos de píxeles correctamente.

#### 3.4.6 Clasificador Bayesiano

El clasificador bayesiano se puede utilizar para eliminar la dependencia de parámetros que no tienen un sentido físico. De esta forma, este método incorpora conceptos de probabilidad para la discriminación de píxeles con el propósito de tratar eventos físicos aleatorios que generan los patrones observados en la imagen.

La probabilidad de que un patrón  $\mathbf{x}$  pertenezca a la clase  $w_i$  se denota  $p(w_i|\mathbf{x})$ . Si el patrón clasificador decide que  $\mathbf{x}$  pertenece a  $w_j$  cuando realmente proviene de  $w_i$  incurre en una pérdida denotada  $L_{ij}$ . Como el patrón  $\mathbf{x}$  puede pertenecer a cualquier clase  $W$  bajo consideración, la pérdida promedio incurrida al asignar  $\mathbf{x}$  a la clase  $w_j$  es:

Ecuación 3.3

$$r_j(\mathbf{x}) = \sum_{k=1}^W L_{kj} p(w_k|\mathbf{x})$$

Esta ecuación normalmente se conoce como riesgo condicional promedio o “pérdida” en la terminología de la teoría de decisiones.

De la teoría de probabilidad básica se sabe que

Ecuación 3.4

$$p(w_k|\mathbf{x}) = (p(\mathbf{x}|w_k)P(w_k))/p(\mathbf{x})$$

Sustituyendo en la Ecuación 3.3

Ecuación 3.5

$$r_j(\mathbf{x}) = \frac{1}{p(\mathbf{x})} \sum_{k=1}^W L_{kj} p(\mathbf{x}|w_k)P(w_k)$$

Donde  $p(\mathbf{x}|w_k)$  es la función de densidad de probabilidad de los patrones de la clase  $w_k$  y  $P(w_k)$  es la probabilidad de ocurrencia de la clase  $w_k$ . Ya que  $1/p(\mathbf{x})$  es común para todas las  $r_j(\mathbf{x})$  es posible obviar el término.

El clasificador tiene  $W$  posibles clases a asignar a cualquier patrón desconocido. Se calcula  $r_1(\mathbf{x}), r_2(\mathbf{x}) \dots r_W(\mathbf{x})$  para cada patrón  $\mathbf{x}$ , y se asigna a la clase con la menor pérdida,  $r(\mathbf{x})$ , significando que minimizará la pérdida total promedio. A este proceso se le conoce como el clasificador de Bayes.

A la pérdida para una decisión correcta se le asigna 0 y para una incorrecta 1. Bajo estas condiciones la función de pérdida se convierte en:

Ecuación 3.6

$$L_{ij} = 1 - \delta_{ij}$$

Sustituyendo

Ecuación 3.7

$$r_j(\mathbf{x}) = \sum_{k=1}^W (1 - \delta_{ij}) p(\mathbf{x}|w_k)P(w_k) = p(\mathbf{x}) - p(\mathbf{x}|w_j)P(w_j)$$

Y al comparar las  $r_j(\mathbf{x})$  se puede obviar nuevamente el término  $p(\mathbf{x})$ , se aprecia que el clasificador de Bayes para una función de perdida 0-1 representa el cálculo de las funciones de decisión ( $d_j$ ) a partir de la siguiente ecuación:

Ecuación 3.8

$$d_j(\mathbf{x}) = p(\mathbf{x}|w_k)P(w_k)$$

Debido al cambio de signo, al patrón  $\mathbf{x}$  se asigna a la clase cuya función de decisión entrega el número mayor.

Las funciones de decisión son óptimas en el sentido que minimizan la perdida promedio en los errores de clasificación. Para que se mantenga esta condición se deben conocer las funciones de probabilidad de densidad de los patrones y la probabilidad de ocurrencia. La última puede ser inferida con conocimiento del problema, pero el encontrar las funciones de probabilidad de densidad se complica para patrones n-dimensionales, requiriendo métodos de teoría de probabilidad multi-variable para su estimación. Por ello generalmente se supone una expresión analítica para las funciones y se estiman los parámetros necesarios a partir de patrones de muestra de cada clase. La forma aceptada para  $p(\mathbf{x}|w_k)$  es la función de densidad de probabilidad Gaussiana.

Para un problema 1D, los patrones  $\mathbf{x}$  se vuelven escalares, con 2 clases, medias  $m_1$  y  $m_2$  y desviaciones estándar  $\sigma_1$  y  $\sigma_2$  se tiene la forma:

Ecuación 3.9

$$d_j(x) = \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{(x-m_j)^2}{2\sigma_j^2}} P(w_j) \quad j = 1,2$$

Para el caso de las imágenes del sistema IBIS en Cancún, se decidió abordar el problema 1D (ecuación 3.9) tomando en cuenta solamente la componente HUE de las imágenes promedio. A

continuación se muestra el proceso a través de diversas imágenes, donde cada una muestra la aplicación de la técnica.

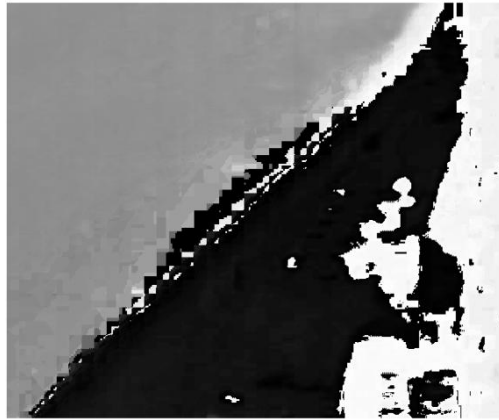
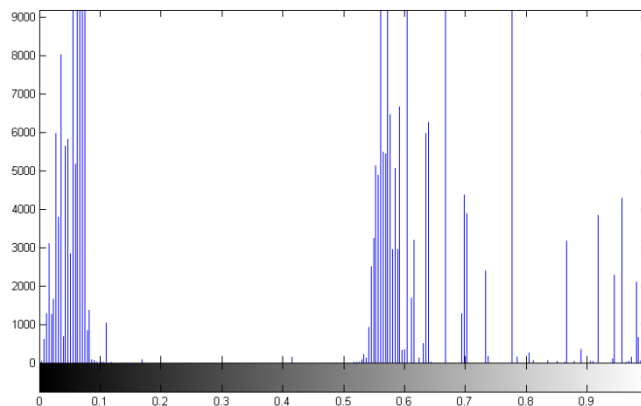


Figura 3.16 Canal HUE.

En la Figura 3.16 se muestra la componente HUE de la cual se calcula el histograma para cada imagen de entrenamiento, y se definen 4 clases, 2 pertenecientes a píxeles ubicados en zonas mojadas y 2 a zonas secas, y se calcula la varianza y media de cada clase. Un ejemplo de estos cálculos se muestra en la Figura 3.17.



Clase	1	2	3	4
Varianza	0.00043	0.002234	0.00285	0.0016293
Medias	0.05773	0.085934	0.57046	0.9680989

Figura 3.17 Histograma de la componente HUE.

Una vez definidos estos parámetros es posible implementar la clasificación Bayesiana, de la que resulta una imagen como la que se muestra en la Figura 3.18. En esta figura se identifican claramente dos regiones en la imagen asociadas a los píxeles en tierra y en mar, pero es posible observar cierto ruido en la zona correspondiente a la línea de costa y los edificios ubicados sobre la playa.



Figura 3.18 Imagen binaria a partir de la clasificación bayesiana de la componente HUE.

Para mejorar la calidad de la información que se presenta en la Figura 3.18, se utilizan operadores morfológicos, tales como erosión y dilatación, para eliminar pixeles dispersos y cerrar huecos, este proceso se usa para disminuir la probabilidad de que el detector de bordes clasifique pixeles de tierra como agua. Una vez aplicados estos operadores se obtienen imágenes como las que se presentan en la Figura 3.19.

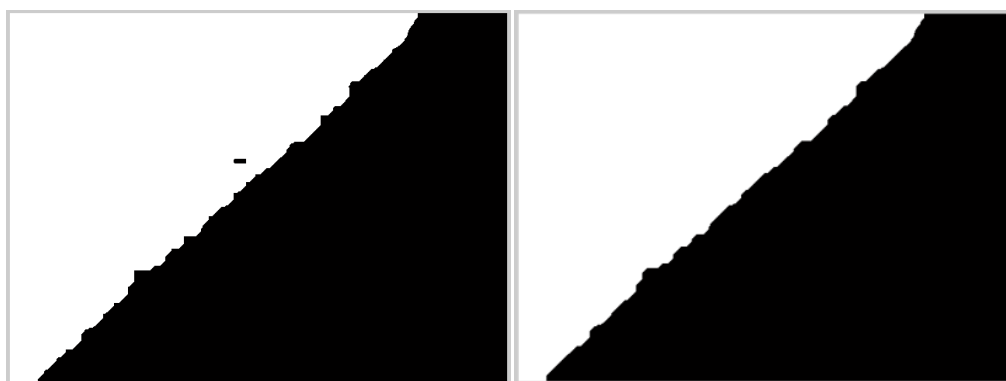


Figura 3.19 a) Imagen resultante al aplicar erosión a 3.18; b) Imagen resultante de aplicar dilatación a 3.19a.

Para evitar la detección de 2 posibles regiones de playa seca, como se aprecia en el panel izquierdo de la Figura 3.20, se utilizan vecinos de conectividad 8 (que incluyen los vecinos horizontales, verticales y las diagonales de la matriz). Con estos se estima el área de cada uno de los elementos, y se determina como playa seca al objeto con mayor área, la Figura 3.20 no es la misma que se ha estado usando para la explicación del proceso debido a que en la anterior no se da este fenómeno.

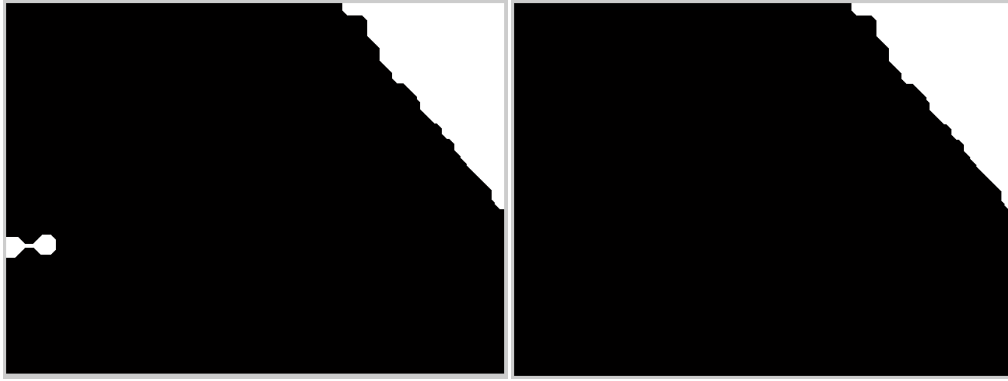


Figura 3.20 a) Detección de 2 áreas de playa; b) Eliminación del área falsa con vecinos de conectividad 8.

Por último, se utiliza un operador Sobel, definido en la sección 2.3, para detectar los bordes del área resultante y la línea de costa queda definida en la imagen como se observa en la Figura 3.21.

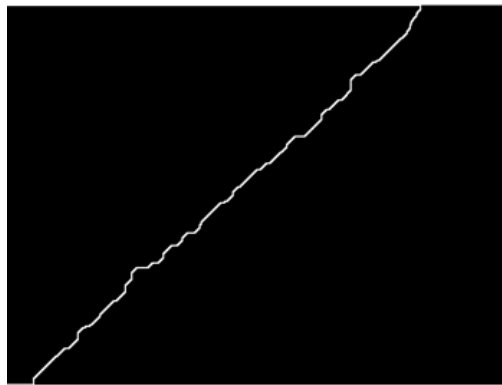


Figura 3.21 Contorno encontrado a partir del detector de bordes tipo Sobel de tamaño 3\*3.

### 3.5 Comparación de las técnicas

Con el fin de evaluar la aplicación de los métodos descritos en las imágenes promedio obtenidas por el sistema IBIS, en este apartado se utilizará la información de la cámara instalada en el hotel Royal Sunset de Cancún, que está programada para tomar 14 vistas de la playa. Para ello, se utilizan 7 imágenes con el propósito de obtener los parámetros estadísticos necesarios para calcular la función de densidad de probabilidad Gaussiana y se utilizan las siete restantes para estimar la línea de costa. La Figura 3.22 presenta la línea de costa estimada en las primeras 7 imágenes utilizadas para la obtención de los parámetros estadísticos. En estos resultados es posible determinar que esta primera estimación no es del todo adecuada, ya que en algunos casos se observa un comportamiento caótico de la línea de costa.



Figura 3.22 Imágenes de entrenamiento, la línea negra representa la interface agua-tierra, vista 1-7.

Sin embargo, si se utilizan los parámetros estadísticos de estas primeras siete figuras para encontrar la línea de costa de las 7 restantes, se obtiene el resultado que se presenta en la Figura 3.23. Donde se aprecia un mejor comportamiento del método.



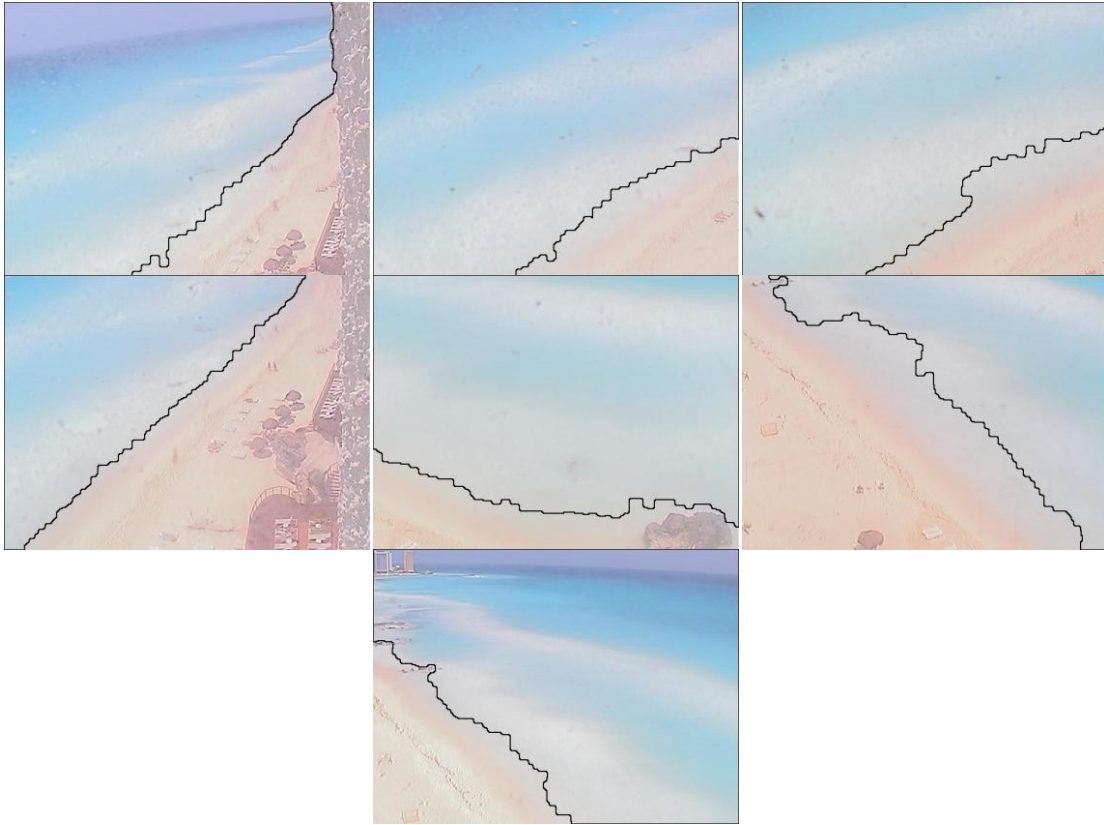


Figura 3.23 Línea de costa encontrada, Vistas 8-14

Para verificar la funcionalidad de la técnica implementada, se digitalizó de forma manual la línea de costa de una imagen promedio tipo y se comparó con aquellas encontradas por el SDM y el clasificador bayesiano. La Figura 3.24 presenta los resultados de esta comparación, donde los puntos azules representan la digitalización manual, los rojos son aquellos encontrados por el SDM y los verdes indican los obtenidos por el clasificador bayesiano. La definición de los parámetros  $p_1$  y  $p_2$  en el método SDM, se realizó a través de la prueba y error, conservando los valores que de forma visual generaban un mejor resultado para el método (los valores definidos fueron de 4 y 1 respectivamente).

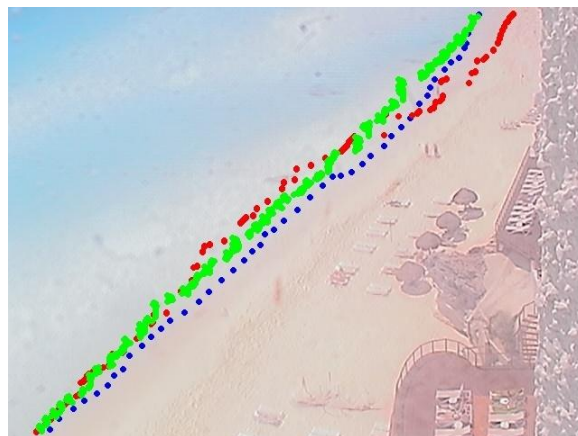


Figura 3.24 Comparación entre SDM, el clasificador bayesiano y la línea ubicada manualmente, Vista 11.

Si se dibujan las líneas de costa encontradas por estos tres métodos en el espacio de píxeles de la imagen, como se reporta en la Figura 3.25, se representan tres líneas invertidas en la vertical. Esto se debe a que el origen de las imágenes se encuentra en la esquina superior izquierda de la imagen.

El clasificador bayesiano presenta oscilaciones de alta frecuencia espacial. Sin embargo, tiene la misma tendencia, con un error cuadrático medio de 25.56 píxeles. El SDM comienza igual que el bayesiano, pero a partir del píxel 200 (x) se incrementa su diferencia con la línea manual y a partir del 400 acorta la diferencia, tiene un error cuadrático medio de 31.11 píxeles.

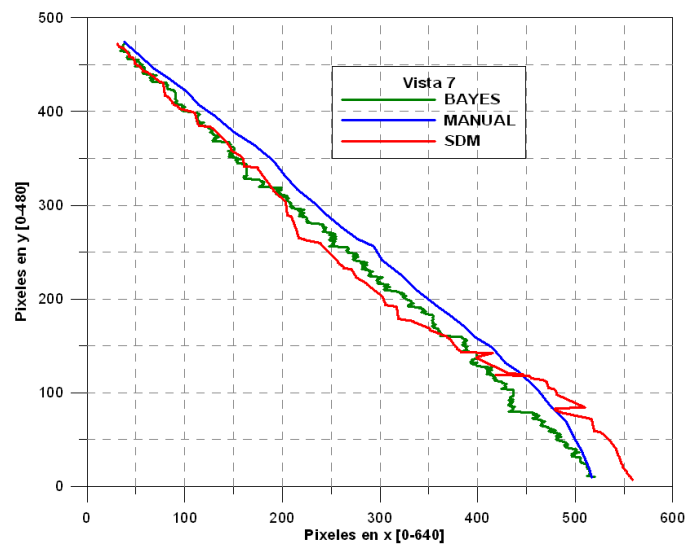


Figura 3.25 Comparación entre SDM, el clasificador bayesiano y la línea ubicada manualmente, Vista 11.

Con el propósito de observar el desempeño de estas técnicas en otra imagen, en la que de antemano se sabe que el clasificador bayesiano no funciona correctamente, se presentan los resultados en la Figura 3.26. En este caso, los parámetros del SDM que mejor ajustaron para la siguiente imagen son 10 y 1, el error promedio calculado para el método bayesiano fue superior al SDM (93.21 píxeles contra 63.89). Cabe señalar que en esta imagen, encontrar la interface entre la tierra y el agua fue complicado aun de forma manual.

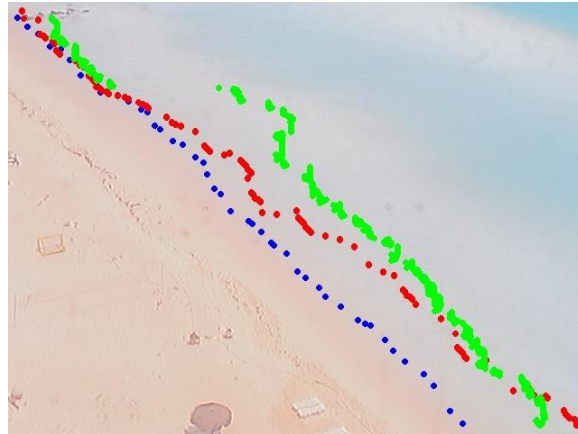


Figura 3.26 Comparación entre SDM, el clasificador bayesiano y la línea ubicada manualmente, Vista 13.

La Figura 3.27 presenta las líneas obtenidas en el espacio de píxeles de la imagen en donde se aprecia una diferencia notable entre la línea de costa derivada por ambos métodos con aquella obtenida de forma manual a partir del píxel 100.

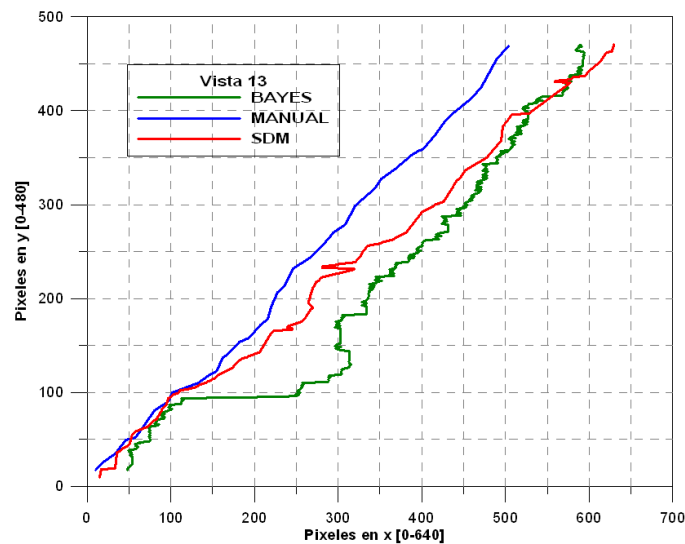


Figura 3.27 Comparación entre SDM, el clasificador bayesiano y la línea ubicada manualmente, Vista 13.

### 3.6 Geo-referencia

Las líneas de costa que se presentaron en el sub-capítulo pasado tienen propiedades cualitativas, para poder extraer características cuantitativas de las imágenes es necesario seguir 2 pasos: el primero corresponde a la corrección geométrica de las deformaciones propias del lente de la cámara (calibración) y el segundo consiste en realizar una proyección perspectiva (ambos procesos fueron descritos en el capítulo 2 y no se repiten aquí).

La corrección geométrica de la cámara se hace con ayuda del “Camera Calibration Toolbox” de MatLab, distribuido por Jean-Yves Bouguet y la cual se puede descargar en la siguiente dirección [http://www.vision.caltech.edu/bouguetj/calib\\_doc/](http://www.vision.caltech.edu/bouguetj/calib_doc/). El Toolbox se basa en el trabajo de Heikkilä y Silvén (1997), que se explicó en el sub-capítulo 2.7.

El programa proyecta puntos en líneas rectas y los compara con puntos conocidos en la imagen. Para encontrar los puntos conocidos se requiere un tablero de ajedrez, ya que es sencillo extraer las esquinas de los cuadros automáticamente. La parte manual consiste en seleccionar las 4 esquinas del tablero, con éstas se proyectan las esquinas en líneas rectas. Posteriormente se encuentran las esquinas reales. Se usan 14 imágenes con distintos ángulos.

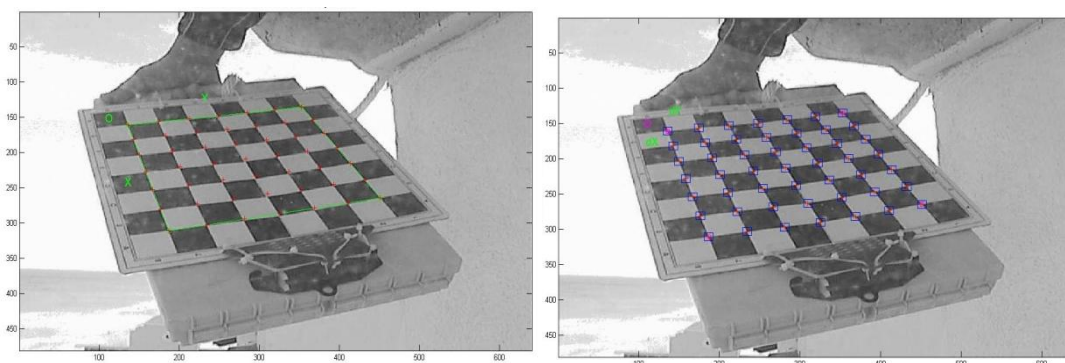


Figura 3.28 Proyección de las esquinas y su extracción.

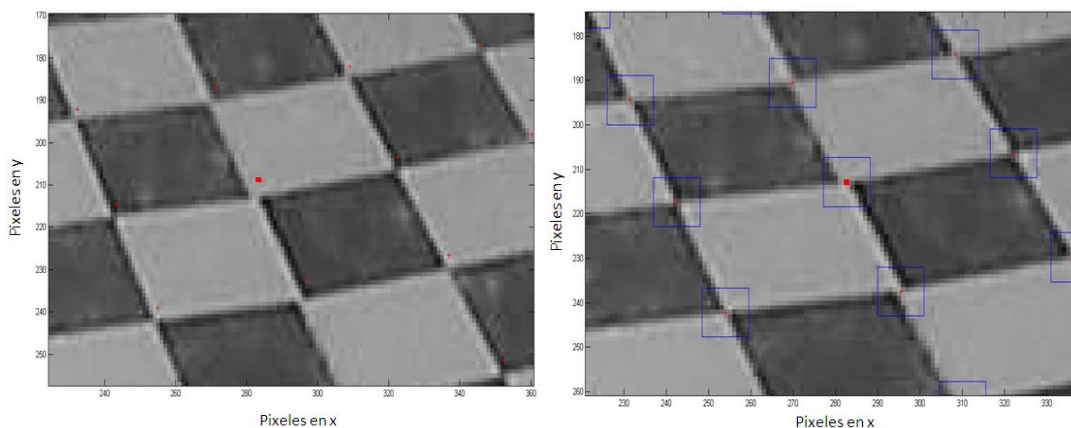


Figura 3.29 Acercamiento de la Figura 1.28.

Se continúa con el proceso de la calibración, que consiste en 2 pasos: la inicialización y la optimización no lineal. El primero calcula una solución de forma cerrada para los parámetros de calibración sin incluir las distorsiones por el lente, solamente la distancia focal y el punto principal. La optimización no lineal reduce el error total de la re-proyección con todos los parámetros de calibración y se realiza a través de un método iterativo de gradiente descendiente con un cálculo explícito de la matriz Jacobiana.

El programa entrega los siguientes parámetros internos: distancia focal (cc1 y cc2), punto principal (fc1 y fc2), coeficiente de torcimiento (sk), 2 coeficientes de distorsiones radiales (kc1 y kc2) y tangenciales (kc3 y kc4).

fcx	933.115
fcy	947.45
cc1	331.56
cc2	127.879
sk	0
r1	-0.359
r2	0.9096
t1	-0.01243
t2	0.002481

Figura 3.30 Parámetros de calibración para la cámara de Royal Sunset.

Con dichos parámetros se reduce la distorsión de las imágenes, sin embargo, no la desaparece por completo. Con un error promedio de 0.48 pixeles en X y 0.52 en Y. El rango del error en X es [-0.7,0.6] y en Y [-1.4,2] para la Figura 1.28 y tomando en cuenta todas las imágenes el rango en X es [-1.8,1.7] y en Y [-1.5,2.3]. Las Figuras 3.31 y 3.32 muestran el error de los pixeles proyectados con los encontrados en las esquinas de los cuadros del tablero.

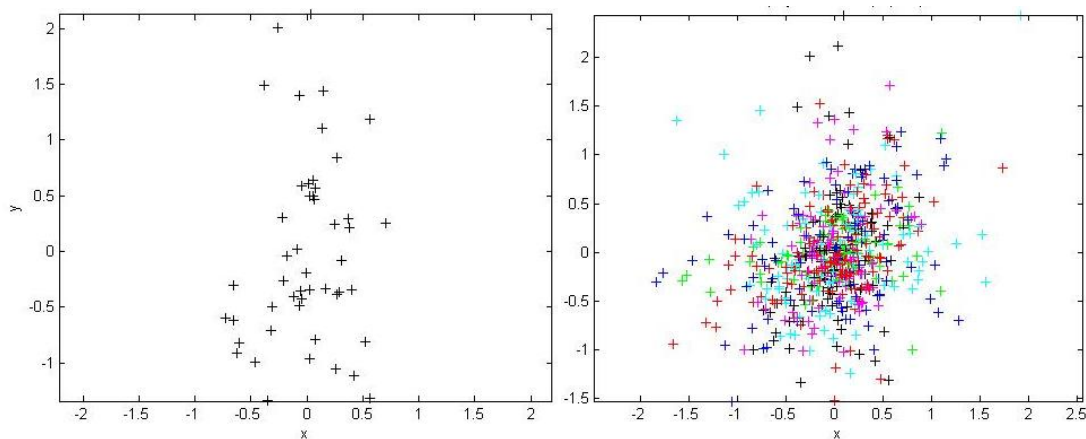


Figura 3.31 Con parámetros de calibración, errores de proyección en pixeles de una imagen (I) y de todas (D).

Para comparar qué sucede sin calibrar la cámara, se forzaron los coeficientes de distorsión a 0, con lo que se tiene un error promedio de 0.58 pixeles en X y 0.63 pixeles en Y. El rango de error en X es [-1.1,1.8] y en Y [-1.9,2.4] para una sola imagen y tomando en cuenta todas las imágenes el rango en X [-2.1,2.1] y en Y [-2.2,2.6].

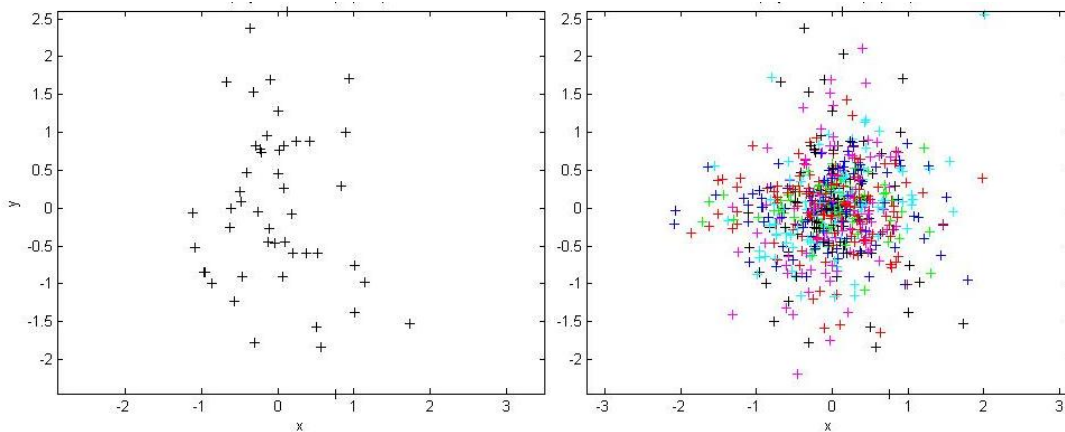


Figura 3.32 Sin parámetros de calibración, errores de proyección en pixeles de una imagen (I) y de todas (D).

El promedio del error se reduce en 0.1 pixeles en X y 0.12 pixeles en Y. Los rangos de error se reducen de 2.9 a 1.3 en X y de 4.3 a 3.4 tomando en cuenta una imagen, tomando todas en cuenta se reduce de 4.2 a 3.5 en X y de 4.8 a 3.8 en Y.

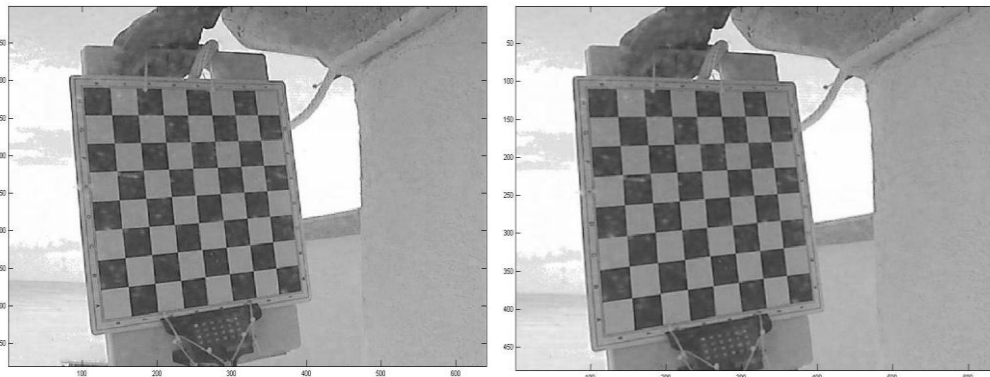


Figura 3.33 Corrección de una imagen con los parámetros hallados.

Con la reducción de las distorsiones en la imagen, se prosigue con la proyección perspectiva que requiere de 4 puntos de control, como se explicó en el sub-capítulo 2.8.

Para obtener las coordenadas de los puntos de control en Coordenadas UTM, se empleó el sistema de GPS diferencial (DGPS) que corresponde al modelo Leica 1200 de doble frecuencia que consiste de 2 aparatos, la estación base y la móvil. La estación base se corrige en posición geográfica con la estación MERI de GPS del INEGI, para posteriormente corregir la información de la estación móvil con la de la base (González Leija, 2009).

Se tomó el muestreo de cada punto a una frecuencia de 1 hz durante un minuto, haciendo lo posible por no mover el GPS, posteriormente se promedia la información en longitud y latitud para reducir errores no deseados. En la siguiente figura se dibujan, con círculos, los puntos de control, el centro es el pixel al que pertenece.



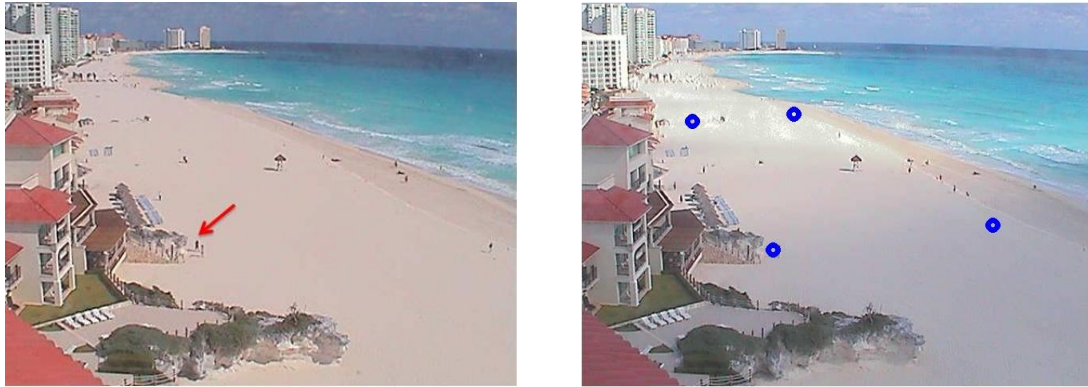


Figura 3.34 Ejemplo de punto de control de una vista, los 4 puntos dibujados.

Con los puntos de control definidos se aplican las ecuaciones del sub-capítulo 2.8, con lo que se observa la gran distorsión que hay en los hoteles, debido a que éstos no se encuentran en el mismo plano.

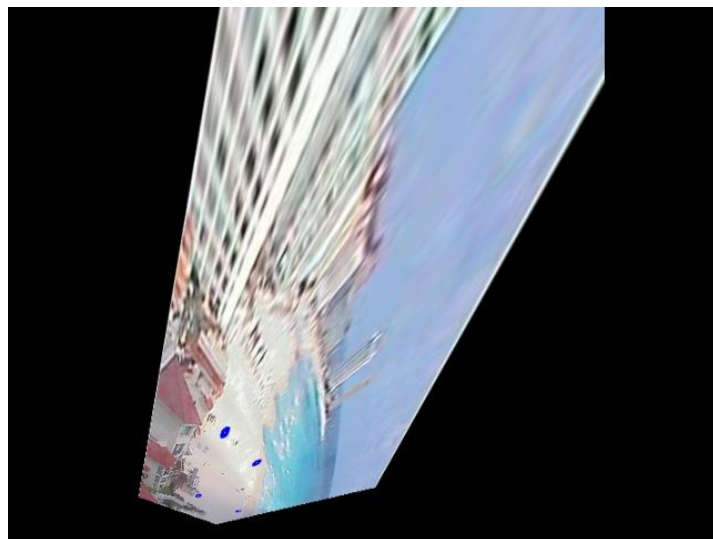
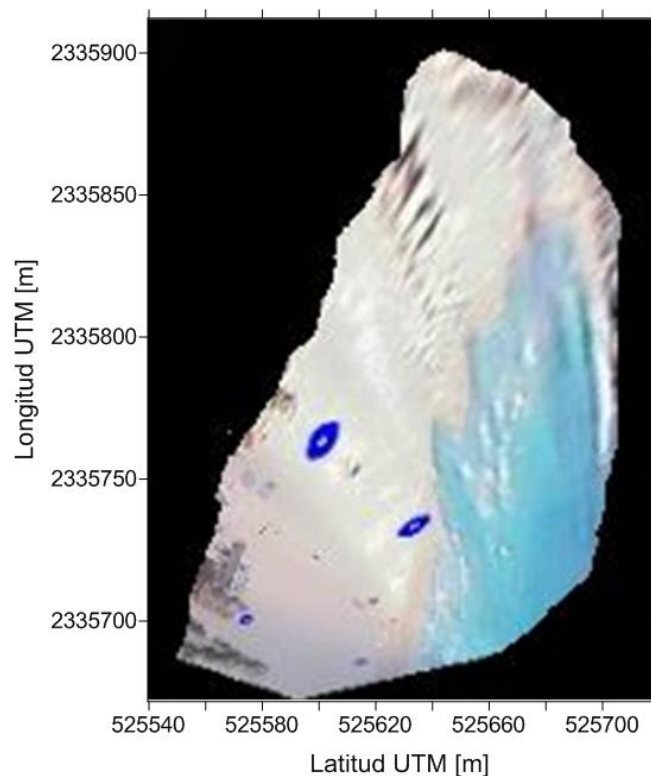


Figura 3.35 Proyección Perspectiva.

Removiendo los hoteles de la imagen y realizando un acercamiento se obtiene la Figura 3.36.



**Figura 3.36** Proyección perspectiva en coordenadas UTM.

La deformación que se aprecia en los círculos de la imagen es clara, los más cercanos a la cámara se redujeron y los más lejanos se alargaron. Cerca de los puntos de control es alta la confiabilidad de que las coordenadas sean correctas. Una vez rectificadas y georeferenciadas esta imagen es procesada con el método Bayesiano para extraer la línea de costa en coordenadas UTM. Los resultados se presentan en la Figura 3.37, en la que solo se grafican los puntos cercanos a los de control (hasta la longitud 2335800).



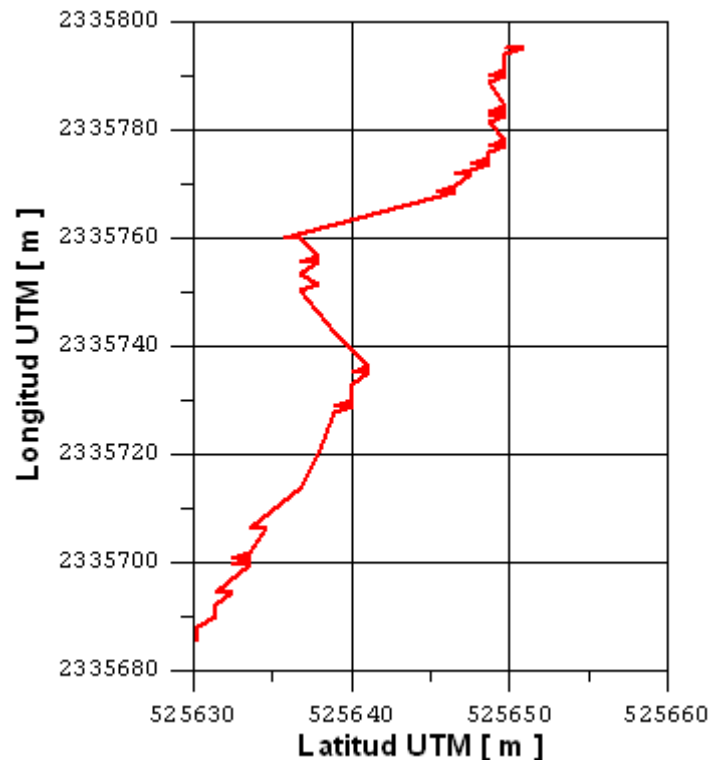


Figura 3.37 Línea de costa de la figura 1.36, localizada con el clasificador bayesiano.

### 3.7 Conclusiones

El color tiene un papel importante para identificar correctamente la línea de costa, sobre todo las diferencias presentadas entre el agua y la tierra, siendo diferente bajo distintas condiciones climáticas:

- Estados energéticos de viento y oleaje incrementan la disipación y el mar se torna más blanco.
- El cielo, al nublarse, vuelve a las imágenes más oscuras.
- La lluvia al caer sobre la carcasa de las cámaras impide la correcta visualización de la playa, una vez que se seca la condensación permitió detectar un problema con la configuración de las cámaras. Al estar activado el autoenfoco la cámara se centra en las gotas de agua o en suciedad que pueda existir sobre la carcasa, además que para enfoques diferentes se requiere de parámetros diferentes para la rectificación de las imágenes.
- Dependiendo de la hora, la luz del sol genera reflejos en la cámara.
- Gran cantidad de bañistas pueden afectar en la calidad de la imagen promedio.
- Otros aspectos que podrían cambiar la información de color son los rellenos y su procedencia o pastos marinos.

El clasificador bayesiano tiene resultados coherentes con los puntos tomados manualmente, sin embargo, la línea de costa que encuentra es, consistentemente, desplazada mar adentro, además que tiene variaciones de alta frecuencia espacial lo cual se atribuye al hecho de que solo se usa un canal para realizar la discriminación de los píxeles.

El modelo SDM usa 2 canales de información para la discriminación, se puede incluir al clasificador el uso de 2 canales de dos maneras, la primera es la unión de la información en un canal a partir de una operación entre 2 canales y la segunda es incrementar al clasificador a 2D, 3D o n-D, de esta manera se pueden incluir no solo 2 canales sino más y poder explorar la discriminación con mayor información.

Una ventaja de trabajar con estas imágenes es que permite cambiar las coordenadas de píxeles a las del mundo real. Se mostró la factibilidad de hacerlo, pero es necesario realizar una campaña de campo con 2 objetivos: el primero sería tomar varios puntos de control en una imagen para poder corroborar a partir de 4 la exactitud de los demás y el segundo tomar más imágenes para calibrar la cámara con un tablero más grande y con más cuadros.

## 4 MEDICIONES DEL ASCENSO MÁXIMO

### 4.1 Introducción

Cuando las olas se aproximan a la costa, la mayor parte de la energía del oleaje se disipa a lo largo de la zona de surf por rotura y fricción, mientras que la porción restante de la energía se pierde durante el proceso de ascenso hasta alcanzar su máxima cota (ascenso máximo).

El ascenso máximo es de interés para el estudio de la dinámica costera, ya que estos movimientos determinan en buena medida la energía que se disipa en forma de lámina de agua que asciende y desciende determinando la zona hasta donde llega la acción del oleaje, y en clima extremo la condiciones de erosión y formación y destrucción de dunas (Sallenger, 2000). Es importante medir el ascenso máximo de manera confiable, tanto en laboratorio como en campo, para poder predecir los alcances de dicho proceso en distintas condiciones de playa (pendiente, tamaño de grano, etc.) y de clima de oleaje (periodo, altura de ola, dirección, etc.).

El ascenso máximo,  $R(y, t_i)$ , es la elevación máxima del nivel de agua medida en vertical respecto al nivel medio del mar (que ocurre en la ausencia de oleaje), la variación del ascenso máximo depende de la localización a lo largo de la costa,  $y$ , y el tiempo,  $t$  (Stockdon *et al.*, 2006).

El ascenso máximo resulta de la combinación de dos diferentes procesos dinámicos: el primero es la sobreelevación de la ola,  $\langle \eta \rangle$ , la elevación promedio del nivel de agua en la línea de costa y el segundo es la zona de lavado,  $S(y, t_i)$ , que representa las fluctuaciones verticales respecto de la media temporal.

La sobreelevación de la ola se ha entendido como un proceso desde los años sesenta con la introducción al concepto de tensor de radiación (Longuet-Higgins y Stewart, 1962). La sobreelevación de la ola es una sobre elevación del nivel medio de agua por arriba del nivel medio (MSL), que es provocada por el gradiente del tensor de radiación resultante de la rotura del oleaje. En la Figura 4.1 se muestra la sobre elevación del nivel medio de agua, es importante notar que la línea de agua y la línea de costa no coinciden debido al ascenso máximo que ocurre en la zona de lavado.

Ecuación 4.1

$$B(x) = \bar{\eta}(x) - SWL$$

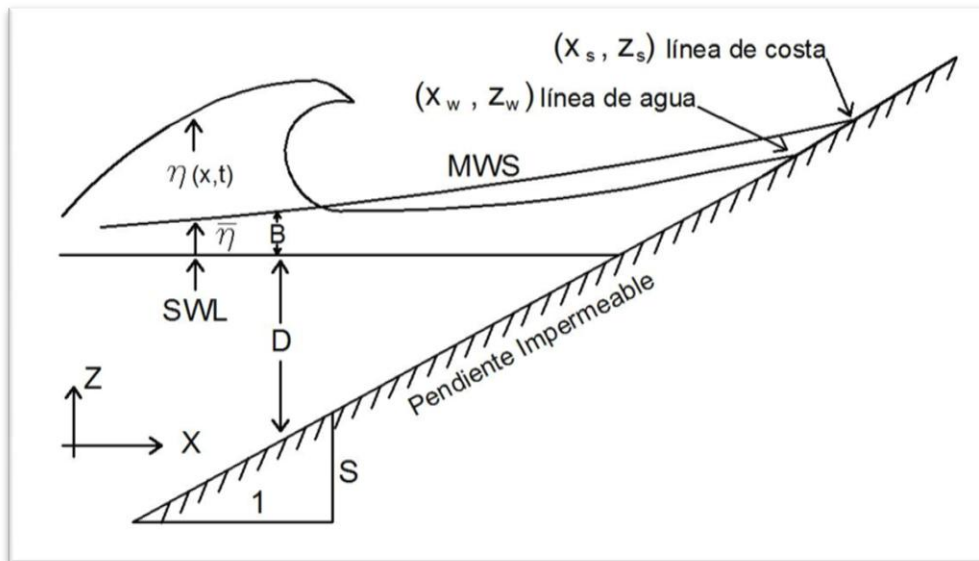


Figura 4.1 Esquema de la sobreelevación del nivel del mar.

Una de las consecuencias importantes de la sobreelevación de la ola es la generación de corrientes de retorno, que resulta de la diferencia del nivel del agua a lo largo de la pendiente de la playa. Se han realizado varios estudios para calcular la sobreelevación de la ola; Bowen *et al.* (1968) escribieron una expresión simplificada para su calcular asumiendo oleaje normal a la costa cuya altura dentro de la zona de surf es limitada por una constante fraccional de la profundidad local,  $\gamma = H/h$ .

Ecuación 4.2

$$\frac{\langle \bar{\eta} \rangle}{H_b} = 0.38\gamma$$

Donde  $H_b$  es la altura a la que rompe la ola.

El segundo proceso que influencia al ascenso máximo ocurre en la zona de lavado, S, y se define como el lugar, que varía en el tiempo, donde se encuentra la intersección entre el mar y la playa. Las olas monocromáticas se componen de dos partes: una componente progresiva cuya energía es disipada durante la rompiente y una que se refleja, teniendo su máximo en la línea de costa, la zona de lavado representa la última componente (Stockdon *et al.*, 2006).

La hidrodinámica y el transporte de sedimentos en la zona de lavado es de interés por varias razones: primero, una parte importante del transporte litoral ocurre en la zona de lavado, la concentración de sedimentos en suspensión es generalmente más alta en la zona de lavado que en la zona interna de transformación (Osborne y Rooker, 1999); segundo, el ascenso máximo en la zona de lavado influencia el diseño ingenieril, produce erosión de dunas y daños a infraestructura cercana a la playa (Xiao y Wenrui, 2008); tercero, la zona de lavado juega un papel importante en la

ecología de la zona inter-mareal (Nordstrom *et al.*, 2006); y cuarto, ayuda a definir límites territoriales legales como en Inglaterra (McGlashan, Duck, y Reid, 2005). En la Figura 4.2 se representan las distintas zonas cercanas a la costa.

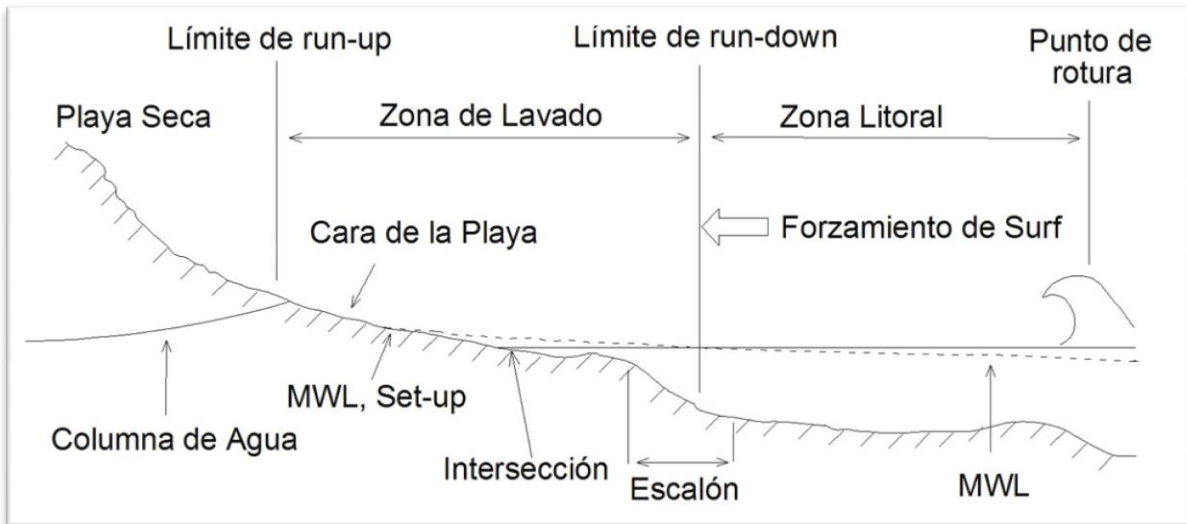


Figura 4.2 Esquema de las zonas cercanas a la costa.

Entre los primeros estudios que se realizaron para comprender y predecir el ascenso máximo destaca el realizado por Hunt (1959), quien señaló que hay una pendiente de onda crítica bajo la cual la energía del oleaje es completamente reflejada por una playa con pendiente plana. Para la situación más común de las olas rompiendo, la energía es disipada a lo largo de la zona de surf. Hunt propuso una fórmula empírica para calcular el ascenso máximo en términos del número de Iribarren,  $\xi$ .

Ecuación 4.3

$$R/H = K\xi$$

En la que  $H=H_0$ . Sin embargo, esta ecuación no toma en cuenta directamente la sobreelevación de la ola. Hedges y Mase (2004) modificaron la ecuación de Hunt para agregar el efecto de la sobreelevación de la ola.

Las mediciones de ascenso máximo, en condiciones tanto de laboratorio como de campo, son importantes para el ajuste de los parámetros para las distintas ecuaciones que tratan de predecir el ascenso máximo, éstas retoman los trabajos realizados por Guza y Thornton (1982) con circuitería diseñada por el laboratorio de procesos costeros del instituto de oceanografía de Scripps (Flick, Lowe, Freilich, y Boylls, 1979). El sistema de medición consistía de cables de doble resistencia que se soportaban 3cm arriba de la cama de arena, se requería de una calibración previa a la toma de datos y cuando había marea alta se requería de una calibración posterior que consistía de

acortamiento de los cables en distintos lugares y la toma del voltaje. Existían diversos problemas operacionales como: presencia de algas; el cambio en la cama de arena requería la elevación y la bajada de los soportes; distintos animales, bañistas, surfistas y turistas que tropezaban con los cables. Estas inconveniencias resultaron en la pérdida de datos y en la confiabilidad limitada de las mediciones.

Holman y Guza (1984) compararon este sistema con uno basado en fotografía, este consistía en colocar la cámara en terreno elevado, posicionar 2 marcadores visibles espaciados una distancia conocida perpendicular a la playa para proveer un rango y escala, mediciones del perfil (igual que en el sistema electrónico) que permitan la transformación de coordenadas horizontales a verticales y una digitalización manual laboriosa. Concluyeron que los dos sistemas son buenas técnicas para la medición del ascenso máximo. Cada una tiene sus ventajas; para el sistema de cables, fácil digitalización, objetividad e insensible a la visibilidad (meteorología); para el sistema fotográfico, bajo costo, facilidad logística, el potencial de tomar varios datos con la misma película y habilidad para “ver” el fenómeno. Ambos tienen sus desventajas; para el sistema de cables, una variedad de problemas logísticos, variabilidad en la ganancia del voltaje y sensibilidad a la altura sobre la cama de arena; para el fotográfico, la tediosa naturaleza y subjetividad de la digitalización.

Entre otros trabajos que utilizaron herramientas de video resaltan los realizados por Holman *et al.* (1990). El proceso era en tiempo quasi real usando un paquete de procesamiento de imágenes que buscaba transectos individuales en intervalos de tiempo regulares.

Las técnicas de medición de ascenso máximo siguen en desarrollo, un ejemplo de ello es el trabajo de (Salmon *et al.*, 2007) donde combina las imágenes promedio con las imágenes apiladas temporalmente “time-stack”, Figura 4.3.

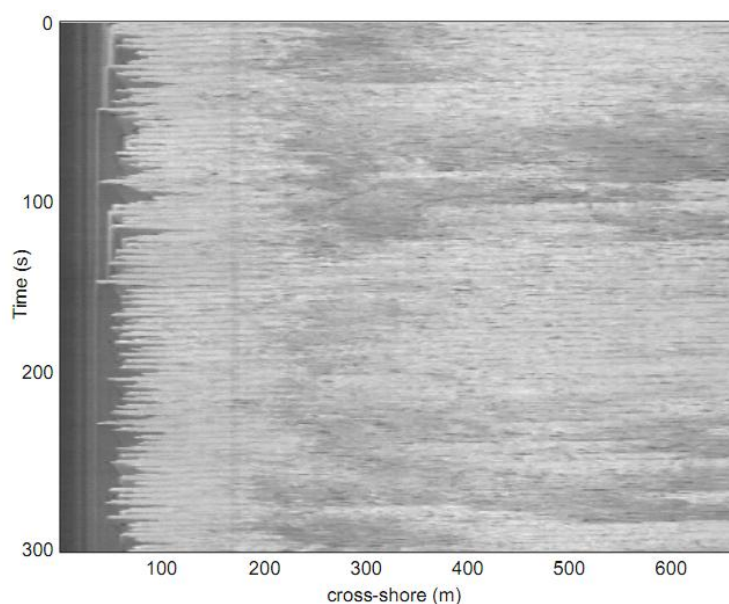


Figura 4.3 Imagen tipo apilado temporal “time-stack” tomada de (Salmon *et al.*, 2007).

## 4.2 Secuencias de imágenes

Se puede pensar en una secuencia de imágenes (o cuadros de un video) como un espacio tridimensional  $x$ - $y$ - $t$ , teniendo 2 dimensiones longitudinales ( $x,y$ ) paralelas al plano de la imagen y una perpendicular al mismo, dimensión temporal ( $t$ ). De este pensamiento abstracto se puede utilizar una técnica simple y poderosa con la cual videos y secuencias de imágenes pueden ser analizados. En inglés se conoce como “time slicing” o secuencias temporales de imágenes, y resulta en imágenes apiladas temporalmente “time-stack”, traducidas como apiladas en el tiempo. Éstas son las resultantes de las metodologías usadas por los autores previamente mencionados (Holman; Guza; Salmon).

La técnica necesita 2 condiciones para su funcionamiento óptimo, una espacial y una temporal: la espacial se refiere a que las coordenadas reales, o absolutas, de las imágenes no cambien, esta condición se cumple automáticamente cuando se toman las imágenes con una cámara estacionaria y la temporal; se refiere a que el intervalo de tiempo entre las imágenes sea constante, esta condición se cumple si se trata de un video (número fijo de cuadros por segundo).

Los cortes de las imágenes son verticales, o paralelos a la dirección del tiempo y pueden tener cualquier dirección y longitud en el plano  $x$ - $y$ , a este corte se le conoce como transecto. La Figura 4.4 presenta una esquematización del proceso necesario para construir la secuencia temporal de un transecto en una serie de imágenes.

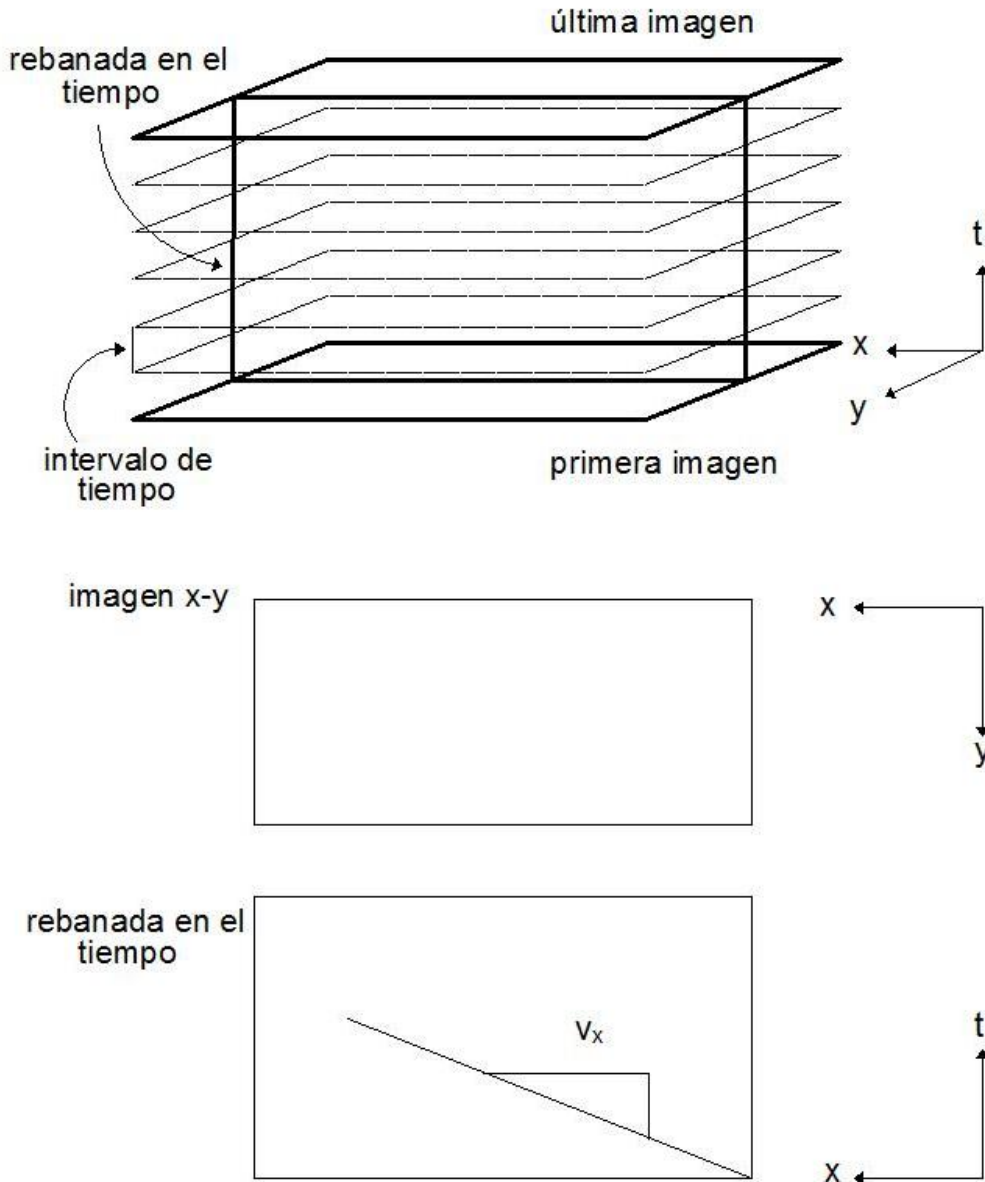


Figura 4.4 Esquemmatización de la técnica de rebanar en el tiempo.

Con el objetivo de revisar la funcionalidad del método se utilizaron videos de ascenso máximo en laboratorio tomados con el mismo modelo de cámara empleado en el capítulo de extracción de línea de costa, se configuró la cámara para tomar los videos con una duración de 2 minutos a una tasa de 30 cuadros/segundo y una resolución de 480\*640 pixeles, sin ningún tipo de iluminación especial.

El tomar mediciones a través de video ha sido un proceso de aprendizaje, lo cual generó ineficiencias en la toma de los videos, por lo que se requirió aceptar algunas hipótesis para resolver los siguientes problemas: no se tomaron puntos de control sobre la pendiente de la playa para poder realizar una transformación afín o una homografía, por lo que se supone que la vista es en planta; ya que la vista se asume en planta se usa la pendiente de la playa (conocida) para



transformar las medidas horizontales a verticales. Por otra parte en los videos no se alcanzaron a grabar las aguas quietas por lo que se tienen que referenciar a un objeto del que se conoce su posición respecto al nivel de agua; debido a que no se pusieron referencias para transformar la equivalencia de pixeles a cm se usa el ancho del canal cuya medida es de 80cm; no se tomaron las imágenes tablero para obtener los parámetros intrínsecos de la cámara por lo que se asume que la cámara no tiene distorsiones radiales y tangenciales.

Las suposiciones que se realizan, estrictamente llevan a mediciones imprecisas y su efecto sobre el resultado real se discute a continuación.

Con respecto a la suposición de que la vista es en planta, el problema yace en que hay cierto ángulo de la cámara, significando que entre más lejos de la cámara se encuentren los pixeles mayor distancia representarán. Los pixeles en la orilla izquierda de la imagen apilada "time-stack" miden lo mismo que los pixeles de la orilla derecha en coordenadas de la cámara, sin embargo, en coordenadas reales los pixeles en la orilla derecha son más grandes. Si se suma el cambio de la profundidad, provocado por la pendiente de la playa, se tiene que entre más abajo estén los pixeles más distancia representaran, surgiendo la incógnita de en qué punto sería más conveniente escalar el ancho del canal.

En la Figura 4.5 el oleaje proviene de la parte inferior de la figura y la altura de la arena es mayor en la parte superior. Por la profundidad de la playa se esperaría que se requieran menos pixeles para representar el ancho del canal en el inferior de la imagen sin embargo esto no ocurre debido a que es mayor el efecto del ángulo de la cámara. Si se escala la línea verde con los 80 cm de ancho del canal y las líneas azules, se tienen valores de 78.22cm, 77.108cm, 76.322cm, 75.337cm y 74.822cm. Donde la mayor diferencia es de 5%. Considerando las escalas, se decide usar la que se encuentra a la mitad del transecto de interés.

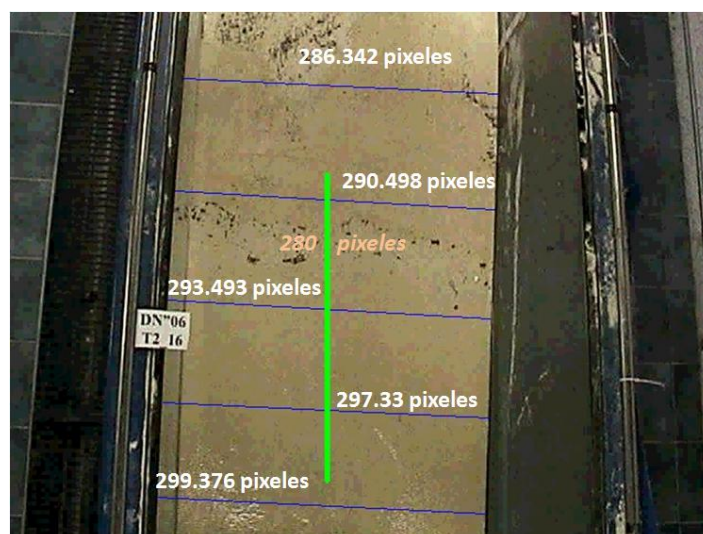


Figura 4.5 Diferencia en pixeles de la anchura del canal.

El usar el perfil inicial (una pendiente sin cambios) para transformar las coordenadas horizontales a verticales es una práctica común (Guza y Thornton, 1982; Bailey y Shand, 1994; Holland *et al.*, 1995; Salmon *et al.*, 2007), aunque las mediciones podrían ser imprecisas si el perfil cambia abruptamente, no siendo éste el caso en el ejemplo.

A pesar de que no entra en el campo de visión del video el nivel de aguas quietas, se conoce la distancia de la imperfección del vidrio al nivel, por lo que se referenciara el ascenso máximo con esta imperfección, Figura 4.7, y a la vez con el ascenso máximo.

Al no tener los parámetros intrínsecos no se puede calibrar la cámara sin embargo el ascenso máximo generalmente se encuentra en el centro de la imagen, donde las distorsiones radiales y tangenciales son despreciables (ver Figura 4.6).

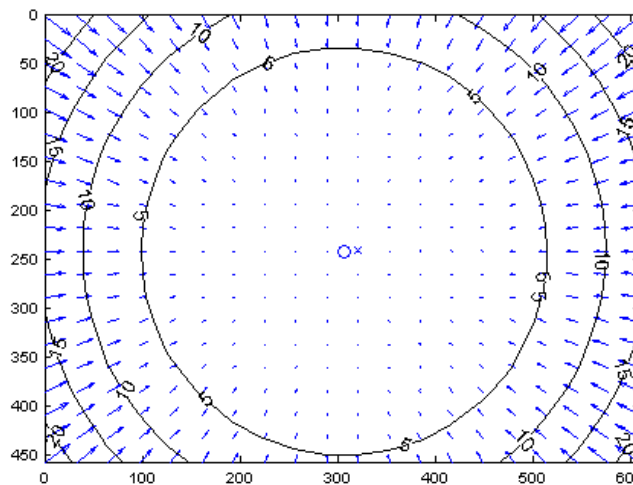


Figura 4.6 Estimación del error en píxeles de la cámara robótica.

Finalmente la iluminación para dichos videos es la “natural” que existe en el laboratorio de costas y puertos, al igual que la configuración de la cámara donde se dejó activado el auto aumento "zoom" y la auto compensación de luz. Estos parámetros tienden a disminuir la consistencia de las imágenes convirtiendo al proceso de automatización en una tarea retardadora y frustrante.

Tomando en cuenta el panorama general de los videos, se decidió realizar una herramienta semi-automática en la cual la decisión de la posición final del ascenso máximo por ola está sujeta a la observación y criterio del usuario.

El primer paso es la elección de los parámetros iniciales: la línea verde define el transecto, la roja la posición de la referencia y la azul es la escala que se usa para pasar de píxeles a cm. Las líneas las define el usuario a través de una interacción con el ratón de la PC.

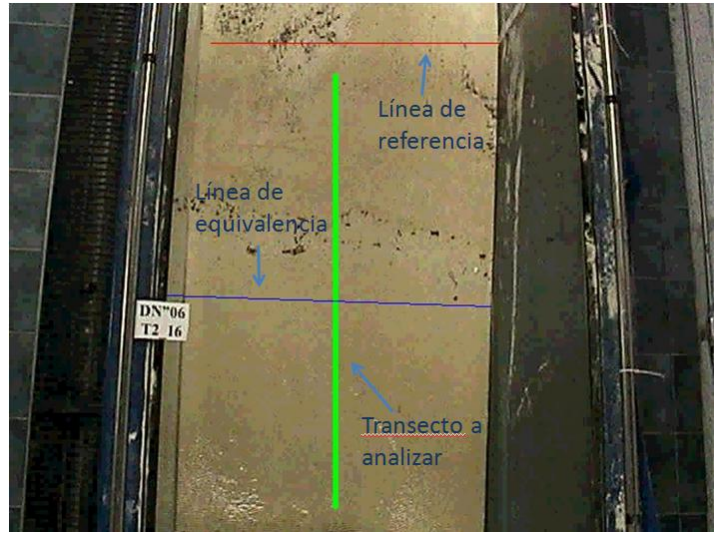


Figura 4.7 Líneas definidas por el usuario.

Una vez elegidos los parámetros empieza el proceso de crear la imagen apilada "time-stack". Para poder apreciar de mejor manera la imagen, esta se divide en varias partes, ya que dura 2 minutos el video y a 30 cuadros por segundo se tiene una imagen con un largo de 3600 pixeles y un ancho, en promedio, de 350 pixeles. En esta división se vuelve a interactuar con el ratón, con cada click se escribe en un archivo la posición y el tiempo, la posición se obtiene por la coordenada x y el tiempo se obtiene por la coordenada y (Ver Figura 4.8).

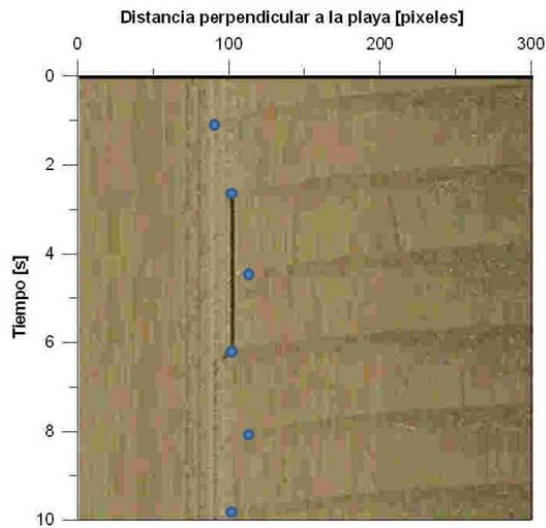


Figura 4.8 Imagen apilada "time-stack"

La Figura 4.9 presenta en centímetros contra tiempo, los puntos encontrados en 3 corridas realizadas por distintos usuarios y una realizada con suma cautela para poder tener una medida de error por la subjetividad, cada persona tardó un minuto en analizar el video y se encontró el error cuadrático promedio para cada serie de datos siendo de 1.61mm, 2.02mm y 2.36mm.

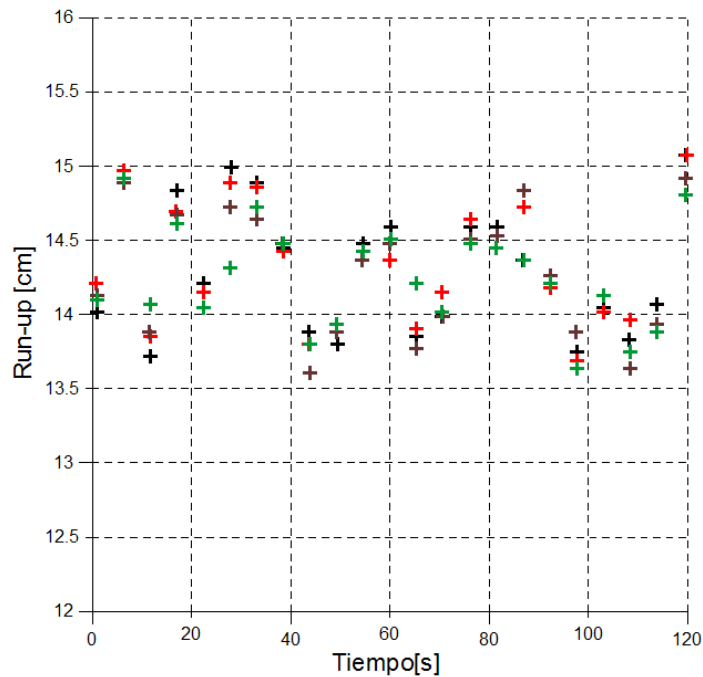


Figura 4.9 Ascenso máximo encontrado por 3 diferentes usuarios y uno de comparación (rojo).

### 4.3 Segmentación por Movimiento

La forma anterior de extraer el ascenso máximo es útil en una amplia variedad de casos, la inconveniencia principal es el hecho de que el proceso no es completamente automático. Es posible extraer de las imágenes apiladas en el tiempo el ascenso máximo automáticamente, por ejemplo a través de una diferenciación de los colores o de la pureza de los mismos, sin embargo, las limitaciones de la cámara usada no permite una segmentación confiable. Una forma que se puede explorar para la segmentación es el movimiento.

La detección de movimiento es una forma poderosa a través de la cual los seres vivos extraen objetos de interés a partir de un fondo irrelevante. El movimiento es fácilmente definido e identificado en aplicaciones donde la cámara permanece estacionaria, de esta manera el movimiento resulta a partir de desplazamientos relativos entre objetos de una serie de imágenes.

#### 4.3.1 Procedimiento Básico

Una de las formas más sencillas para detectar cambios entre 2 imágenes,  $f(x,y,t_i)$  y  $f(x,y,t_j)$  tomadas en los tiempos  $i$  y  $j$ , es comparar las imágenes pixel por pixel, es decir se evalúan las diferencias entre imágenes.

Suponiendo que se tiene una imagen de referencia conteniendo solamente componentes estacionarios y se compara ésta contra una imagen de la misma escena y se incluye un objeto en

movimiento, al restar una de la otra se cancelaran los elementos estacionarios, haciendo a los pixeles diferentes a cero pertenecer a los componentes no estacionarios.

En teoría si se restan 2 imágenes de una misma escena tomadas en 2 tiempos distintos (sin agregar o quitar elementos) la imagen resultante sería completamente negra (ya que el cero representa el negro en el modelo RGB) pero en la práctica esto es difícil de lograr (a menos que las imágenes hayan sido tomadas con una cámara de alta velocidad), esto se debe a que la luz que recibe el lente no es constante. Este efecto se incrementa (en los pixeles estacionarios) si se agrega un elemento, ya que éste puede tener el efecto de reflejar mayor luz al lente o a los elementos estacionarios, por lo que se vuelve necesario el uso de un umbral que ayude a robustecer la detección de movimiento.

Se puede hacer una imagen binaria que defina movimiento (siendo los pixeles blancos el objeto el movimiento) con la siguiente definición:

**Ecuación 4.4**

$$d_{i,j} = \begin{cases} 255 & \text{si } |f(x, y, t_i) - f(x, y, t_j)| > T \\ 0 & \text{de otra manera} \end{cases}$$

Donde T es un umbral definido,  $f(x, y, t_i)$  la intensidad de cada pixel de un cuadro,  $f(x, y, t_j)$  la intensidad de cada pixel del cuadro subsecuente y  $d_{i,j}$  el valor resultante de cada pixel.

Se puede realizar este proceso para un video entero y el producto final sería un video de movimiento. Sin embargo, la única parte donde interesa registrar cualquier cambio es sobre la pendiente de la playa, esta área se puede definir al comparar cada cuadro con una imagen "máscara", Figura 4.10 (b), la cual consiste en una imagen negra de las mismas dimensiones que las del video y se genera una área blanca (ya sea con una serie de líneas, círculos o figuras), la cual define las coordenadas de los pixeles donde se aplica la Ecuación 4.4. El primer pixel (color negro) que se encuentra de izquierda a derecha define el ascenso máximo en cada cuadro, Figura 4.10 (c).

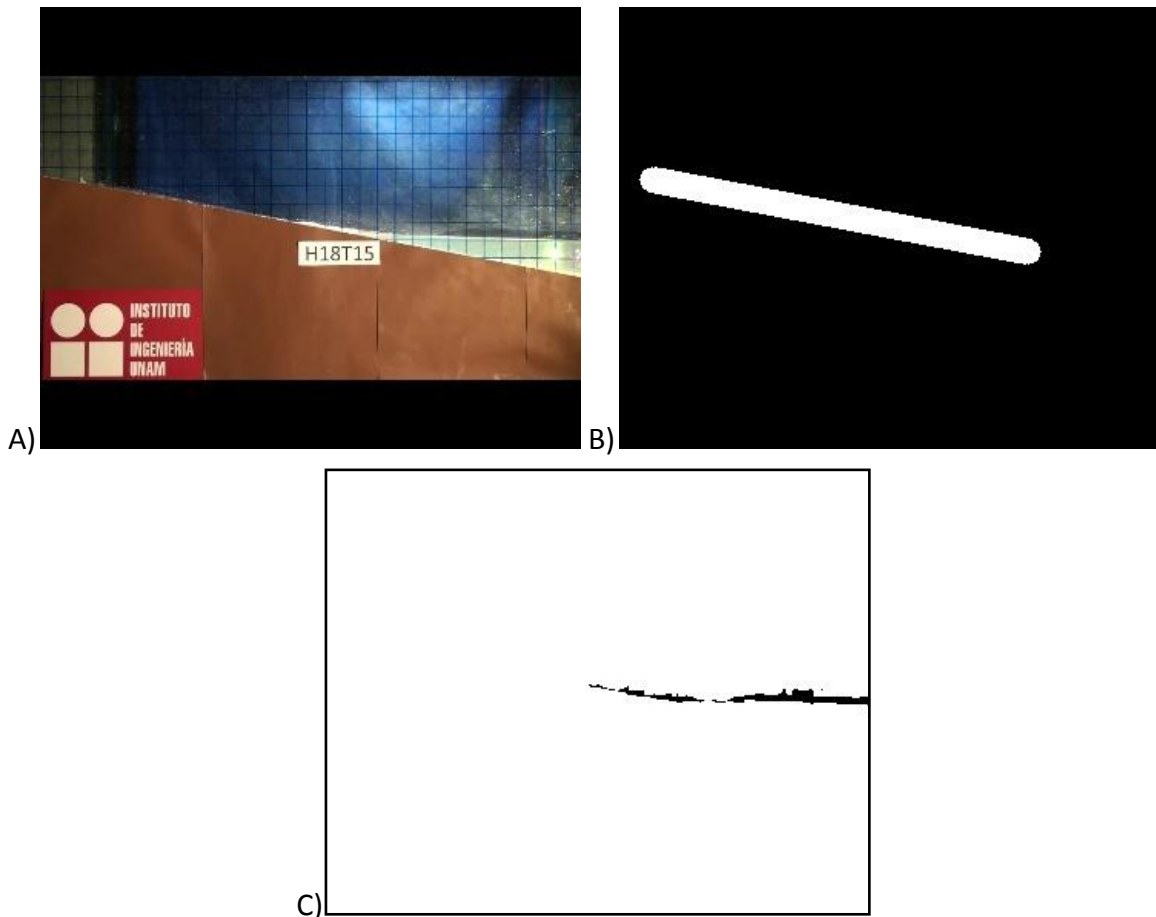


Figura 4.10 a) Cuadro # 105; b) Imagen máscara con las coordenadas de la pendiente; c) Píxeles detectados en movimiento.

Cuando se compara al cuadro con la imagen máscara parece notarse que la pendiente en la máscara se encuentra desplazada, sin embargo, es un efecto óptico de perspectiva.

Para revisar el comportamiento de la técnica, se utilizó un video que toma el perfil del canal de oleaje de la UNAM a una tasa de 30 cuadros por segundo. Se tomaron las coordenadas en píxeles del ascenso máximo observado, donde solo se consideraron los cuadros donde fuese notoria la punta de agua, a éste se le define como ascenso máximo real.

Se ejecutaron 2 casos para obtener las imágenes binarias: la primera comparar los cuadros contra el primer cuadro (donde todo está estático) y la segunda a cada cuadro contra el anterior. Se varió la intensidad del umbral para cada caso 2 unidades de 10 a 80, y se comparó con el ascenso máximo real encontrando el error cuadrático medio para cada una de las 72 pruebas, Figura 4.11.

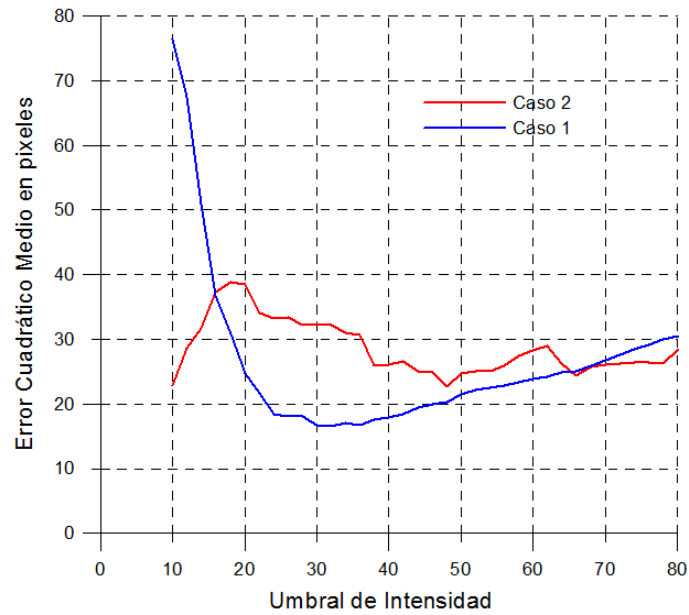


Figura 4.11 Errores cuadráticos con diferentes umbrales para los casos 1 y 2.

El umbral que mejor ajustó fue de 32 para el caso 1, donde todos los cuadros se compararon contra el primero. A continuación se muestra la gráfica donde se compara el ascenso máximo (medido sobre el eje x) encontrado manualmente (línea negra) contra el encontrado automáticamente (línea azul).

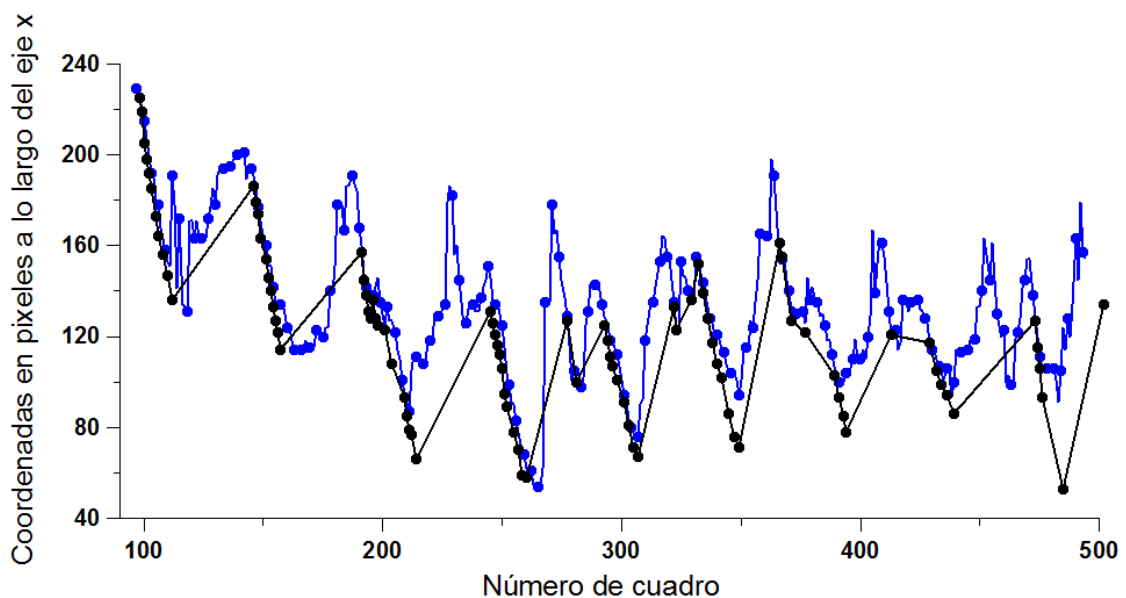


Figura 4.12 La línea negra corresponde a las coordenadas en x del ascenso máximo observado, la azul representa al encontrado a través de la segmentación para los primeros 500 cuadros equivalentes a los primeros 17 segundos.

Es necesario mencionar que la punta de agua era visible solo durante el ascenso, al momento de bajar por la pendiente era confuso por lo cual la mayoría de los puntos observados son en el

momento de subir. Debido a que el análisis se hace en coordenadas de píxeles los puntos negros que corresponden al real se ven en bajada en la gráfica, ya que el agua sube de derecha a izquierda y la orientación de las coordenadas es contraria. El error medio cuadrático en píxeles fue de 17.1, aunque si se transforma a coordenadas en el eje [y] el error es de 3.45 píxeles, en centímetros representa 1.5 cm.

#### 4.3.2 Promedio Corrido

El problema principal que se tiene al aplicar la técnica anterior al video son los grandes cambios de iluminación que se generan al romper la ola sobre la pendiente de la playa, en un punto cercano a la pendiente se obtuvo la información de iluminación de un píxel a lo largo de los primeros 600 cuadros, en la siguiente gráfica se puede notar que la amplitud de los cambios de iluminación son de aproximadamente 20 unidades; en la Figura 4.13 se observa la variación de la intensidad donde las crestas coinciden con la rotura del oleaje.

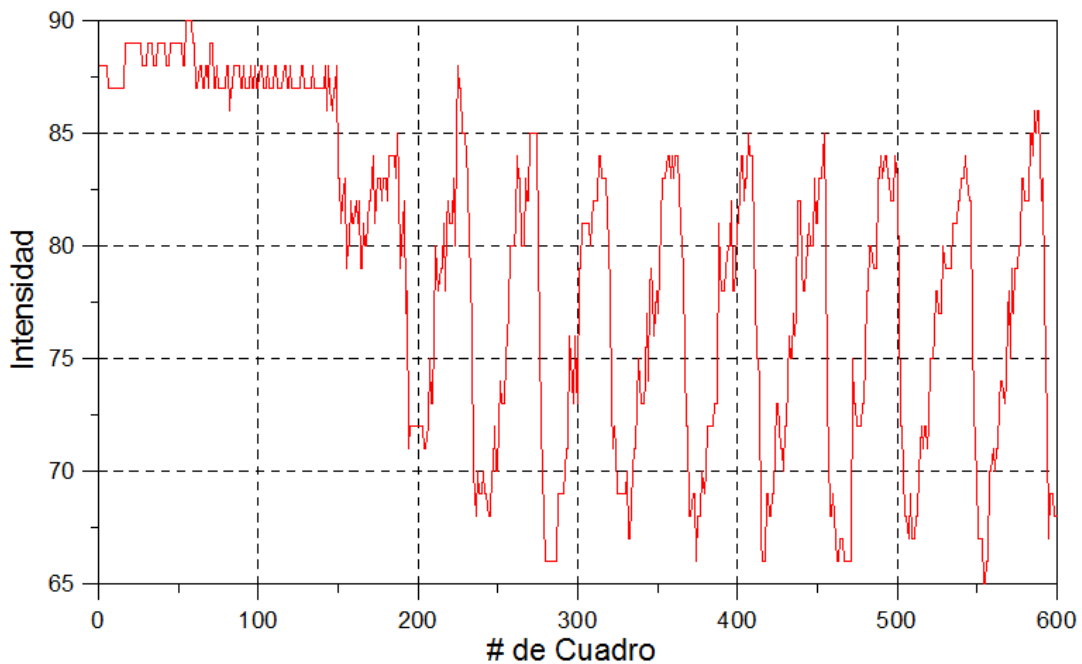


Figura 4.13 Variación de la iluminación en un punto debajo de la pendiente de la playa.

Este cambio contribuye a que el umbral que dio mejores resultados en la técnica anterior fuese de 32, que se considera un valor alto ya que equivale al 12.5% de la gama de iluminación, una posible manera de corregir esto es con el promedio corrido.

Ecuación 4.5

$$a_n(x, y) = (1 - \alpha) * a_{n-1}(x, y) + \alpha * imagen(x, y)$$



Con la ecuación anterior se generan imágenes de comparación, como las del caso 1 y 2 de la técnica pasada, con las que se espera reducir estos cambios de intensidad. Si se toma al coeficiente alfa como 1 se genera el caso 1 (donde la imagen permanece constante y es el primer cuadro) y si se toma como 0 se genera el caso 2 (la imagen es el cuadro anterior). El parámetro alfa esencialmente define la cantidad de tiempo necesaria para que la influencia de los cuadros previos se desvanezca.

Ya que la funcionalidad de la técnica depende de 2 parámetros, el umbral y alfa, es necesario encontrar la combinación que genere los mejores resultados para lo cual se varia el umbral de 6 a 34 unidades cada 2 y alfa de .01 a 1 cada .01, generando 1500 resultados, cada resultado se comparó contra el ascenso máximo observado calculando el error medio cuadrático.

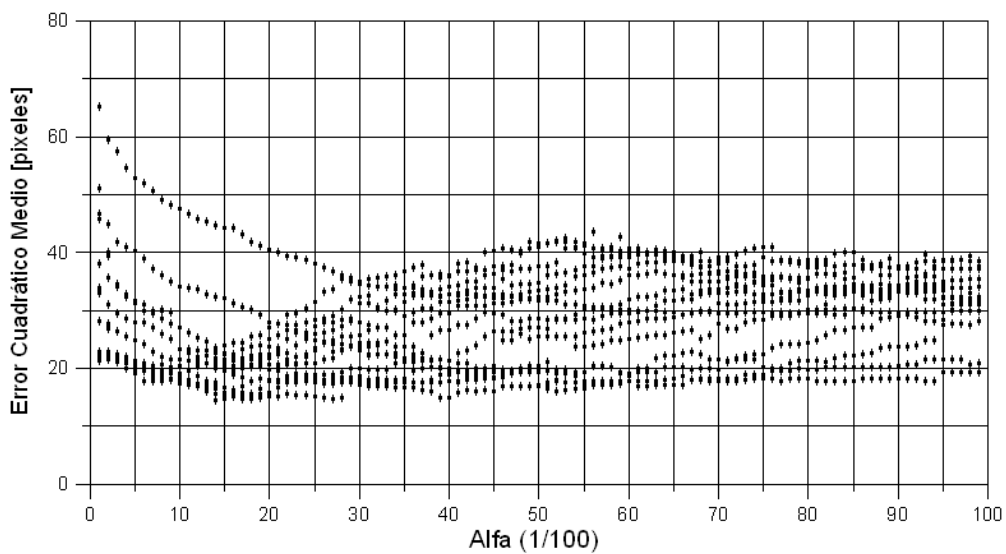


Figura 4.14 Error Cuadrático Medio para cada alfa con 15 umbrales.

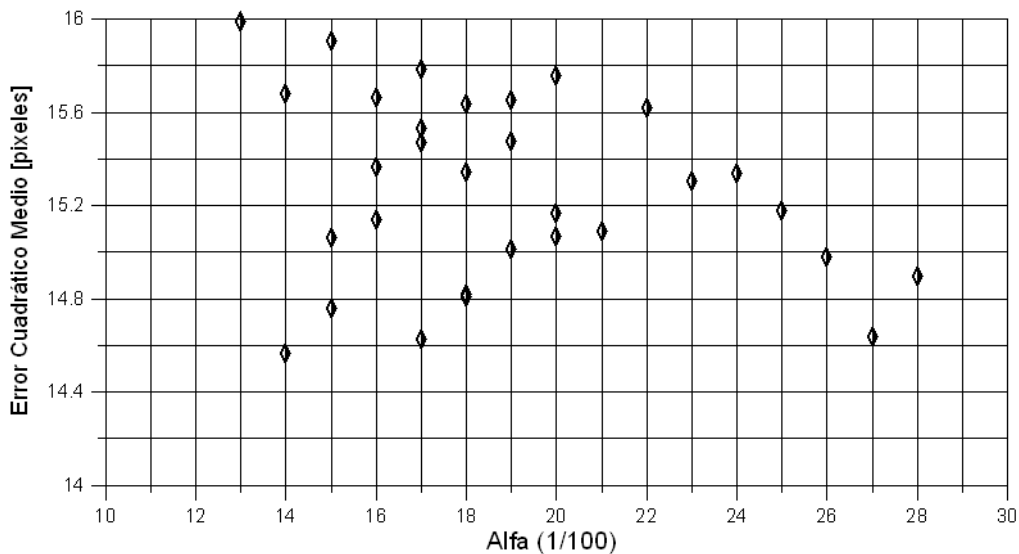


Figura 4.15 Acercamiento de la figura anterior en la zona con menores errores.

Las gráficas anteriores representan los errores cuadráticos medios encontrado en cada alfa, los que tienen los errores más bajos son los alfas de 0.14, 0.17 y 0.27. En la siguiente gráfica se muestra el comportamiento del umbral de cada alfa contra el error cuadrático medio encontrado.

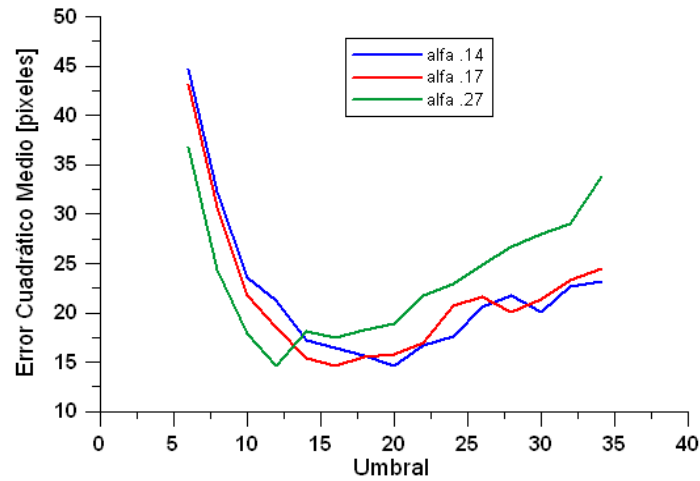


Figura 4.16 La variación del error para los 3 alfas con mejor desempeño.

Aparentemente de los 3 alfas el más inestable es el de 0.27 por lo que se descarta, los parámetros 0.14 y 0.17 parecen tener un comportamiento similar y el que tiene el error menor es el de 0.14 con 14.56 píxeles y un umbral de 20, sin embargo, se decide escoger un alfa de 0.17 ya que el umbral que usa es de 16 y un error de 14.62 píxeles, al usar un menor umbral, permite una mayor fineza al momento de encontrar la punta del ascenso máximo. Los errores que se calcularon fueron a lo largo del eje x, el ascenso máximo se reporta generalmente a lo largo del eje y, al tener una pendiente de 1 a 5 el error se reduce a 2.92 píxeles y si se escala a centímetros se habla de un error cuadrático medio de 1cm.

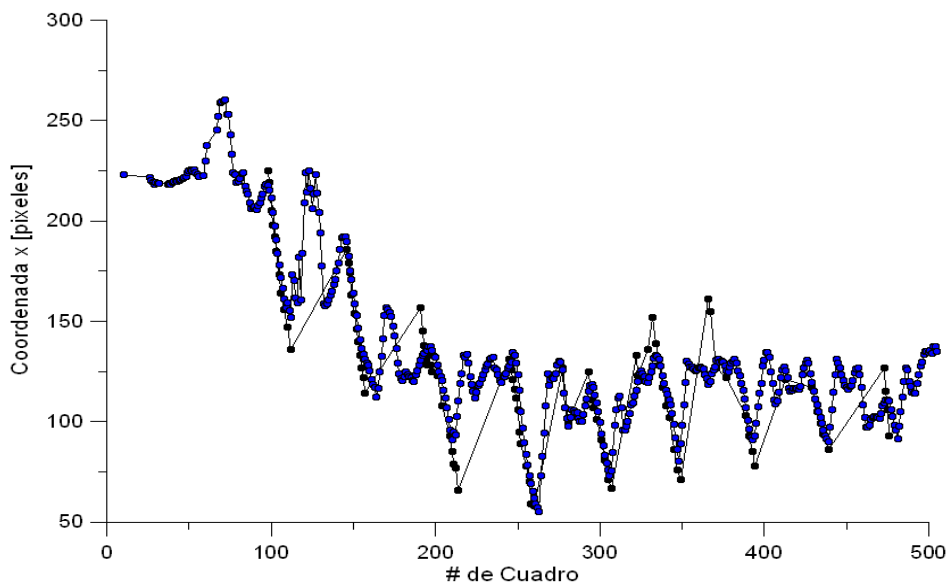


Figura 4.17 Ascenso máximo observado (negro) contra ascenso máximo encontrado (Azul), con un alfa de 0.17 y umbral 16.

## 4.4 Conclusiones

Se mostraron 2 formas distintas de realizar mediciones del ascenso máximo: la primera a través de imágenes apiladas en el tiempo y la segunda a través de la segmentación por movimiento.

Para mediciones en campo, la primera forma es más apta ya que la segunda depende en gran medida de la iluminación y en campo no hay control de las condiciones ambientales. Sin embargo, la segmentación por movimiento tiene un gran potencial para realizar mediciones en laboratorio.

Para disminuir los errores de segmentación por movimiento en laboratorio es necesario realizar ajustes a la cámara, las opciones automáticas deben ser apagadas, la resolución se debe aumentar y la iluminación del ambiente se debe hacer lo más constante posible con la ayuda de lámparas y hojas de materiales, como el polietileno, que ayuden a difuminar la luz. Adicionalmente es posible tapar el canal de forma que solo sea visible el área de la pendiente de la playa, de esta manera la luz reflejada por la rompiente de la ola no afectará a la cámara.

Para usar a las imágenes apiladas en el tiempo en campo o en laboratorio de manera confiable es necesario realizar homografías para proyectar a la imagen en vista de planta y con los puntos usados para realizar las mismas, escalar las unidades de pixeles a métricas. Una vez que se tengan las imágenes apiladas en el tiempo escaladas se buscaría una forma de segmentación basada en los distintos modelos de color para tratar de automatizar el proceso.



## 5 VELOCÍMETRO DE IMÁGENES

### 5.1 Historia

La forma más simple para inferir la velocidad de un flujo probablemente se remonta a la primera vez que alguien observó un puñado de hojas arrastrado por un río. A pesar de la sencillez del concepto, se han desarrollado técnicas de gran complejidad a lo largo de los últimos 30 años. El presente sub capítulo se basa en un artículo que escribió R.J. Adrian por la conmemoración de 20 años del PIV (Adrian R. J., 2004).

No se puede hablar de un velocímetro de imágenes sin hablar del PIV (Particle Image Velocimetry o velocímetro de imágenes de partículas), esto se debe a que la técnica del PIV ha estado en constante desarrollo y validación desde que se vislumbró su capacidad en 1977 y es considerada actualmente como la técnica no intrusiva más robusta para mediciones de velocidades de flujos.

El mismo PIV proviene de una rama alejada a los fluidos, la mecánica de sólidos. Específicamente, en 1977 diferentes grupos de investigación utilizaron la técnica LSV (Laser Speckle Velocimetry) para medir el perfil parabólico de un flujo laminar (Barker y Fourney, 1977). Los principales elementos del LSV eran el uso de fotografías de doble exposición, la iluminación era a través de un láser y el algoritmo se basaba en la interferencia de Young.

La persona en continuar con el uso del método fue Meynart (1983), Roland Meynart (Meynart, 1983) quien demostró que no solo era factible el uso de la técnica en flujos laminares sino también en turbulentos, lo que incrementó el interés de la comunidad interesada en la mecánica de fluidos.

El desarrollo de la técnica continuó con R.J. Adrian, que en 1984 argumentó que las imágenes que se obtenían con el LSV no eran dispersiones de luz sino partículas individuales (Adrian R. J., 1984), acuñando así el nombre PIV. El modo de operación dependía del número de partículas en la imagen, que se distinguía con un número adimensional llamado “densidad de la fuente”.

El incremento en el interés de parte de la comunidad científica y su uso en flujos turbulentos influenció en las decisiones tomadas para el desarrollo del método. Tomando en cuenta que las escalas físicas de la turbulencia se manejan en un rango amplio, la técnica debía poder medir un rango dinámico amplio de distancias y velocidades. La naturaleza aleatoria de la turbulencia hacía imposible conocer, a priori, la dirección del flujo por lo que era necesario que pudiera reconocer flujo en todas las direcciones. Debido a las grandes aceleraciones de los flujos turbulentos, las partículas debían ser lo suficientemente pequeñas para seguir al flujo en estas condiciones extremas.

Las pequeñas partículas requerían de láseres con mayor intensidad de iluminación, abriendo la posibilidad de diferentes tipos de iluminación: láseres C W, láseres de pulsación, iluminación C W con flash de xenón, entre otras. Finalmente surgieron láseres de pulsación doble de estado sólido, que se convirtió en el caballo de fuerza del PIV.

Quedaban una serie de cabos sueltos para definir la técnica, a mediados de los ochentas se encontraban docenas de artículos describiendo distintos sistemas para medición de flujos, cada uno difiriendo del otro en iluminación, código, partículas, modo de grabación y modo de interrogación. Es común, en el desarrollo tecnológico, empezar con una amplia variedad de opciones y confusión antes de terminar con pocas buenas opciones y bien entendidas, para el caso del PIV: láseres de doble pulsación de estado sólido, cámaras de video de transferencia interlineal y la interrogación por análisis de correlación (al menos en ese momento).

El uso del análisis de auto correlación en imágenes de doble exposición, 2 imágenes montadas en una, se propuso por primera vez en 1983 (Adrian y Yao, 1983). El método de auto-correlación llevó a la definición de un segundo número adimensional llamado "densidad de la imagen" que es igual al número promedio de rayos encontrados en un volumen, es decir, el área de interrogación por el grosor de la hoja de iluminación. Imágenes de baja densidad corresponden al seguimiento de partículas; cuando se trata de imágenes de alta densidad corresponde a la correlación de partículas múltiples. El mayor problema de las imágenes de doble exposición era la ambigüedad de la dirección del flujo, este problema se resolvió al aplicar una traslación a la segunda imagen en cierta dirección y así el traslape de partículas también se eliminó.

En la primera década del PIV, el reto mayor era la interrogación de las imágenes, esta interrogación se define en el presente capítulo como la forma matemática en la que se identifica el desplazamiento de una partícula, en parte por las limitaciones computacionales. En 1985 se utilizaban computadoras DEC PDP 11/23 que tenían 128 Kbytes de memoria RAM y 30 MB de disco duro; mantener el sistema operativo y los datos exigía gran esfuerzo y ni pensar en otras tareas simultáneas como escuchar música que es tan común hoy día. La forma de superar estas limitantes, en esa época, fue el uso de la correlación bidimensional por medios ópticos analógicos, y dependiendo de la imagen, podía haber 2 problemas: si había pocas partículas (imágenes de baja densidad) eran pocos los vectores de velocidad que se encontraban, por lo que no era posible resolver campos turbulentos completos, por otro lado, si las imágenes era de alta densidad era probable que se encontrara más de un desplazamiento, en cuyo caso se tenía que asumir que el desplazamiento menor era el correcto.

Uno de los cambios más importantes en el PIV fue pasar de sistemas fotográficos a videos digitales, lo cual interfiere desde la definición de la resolución de los sistemas; en películas fotográficas se trataba de líneas por milímetro mientras que para video se definen por pixeles. La resolución al principio era mucho mayor en las películas fotográficas, se hablaba de 300 líneas por mm, sin embargo, la aceleración en la mejora de la resolución en cámaras digitales era y es impresionante,

pasando de 500x500 píxeles a 1000x1000 píxeles y aún más de 2500x3000 píxeles que equivale a una película de 35mm.

Otro de los grandes pasos tecnológicos tomados en el PIV, fue la invención de cámaras de transferencia interlineal que pueden grabar dos imágenes en sucesión rápida al transferir la primera imagen a un chip y grabar la segunda imagen, de esta manera la toma entre 2 imágenes es extremadamente veloz, permitiendo medir velocidades mayores en rangos espaciales menores. Louis Lourenco fue el que convenció a la empresa Kodak para crear estas cámaras (Lourenco *et al.*, 1994). Este avance representó 3 ventajas: teóricamente la correlación cruzada de dos imágenes es superior a la auto correlación de una imagen, la correlación de imágenes separadas elimina la necesidad de hacer una traslación, en la toma de la segunda imagen, de la doble exposición determinando la dirección por el orden de las imágenes y, tercero, al eliminar la sobreposición de pequeños desplazamientos, el rango de velocidades aumenta.

En el PIV la súper resolución se refiere a formas de interrogación que mejoran la resolución espacial otorgada por la correlación básica. Donde investigadores como Yamamoto, Kobayashi y Okamoto han avanzado en estos procesos (Proceedings of the International Simposia on Particle Image Velocimetry, 1999, 2001, 2003), desarrollando distintas técnicas, como la correlación jerárquica donde se empieza en un amplio espacio de la imagen y se va disminuyendo según los niveles de correlación, viendo al final el proceso como una malla quadtree. Otra técnica es el uso de ventanas variables con 6 grados de libertad, 2 para rotar la matriz, 2 para trasladarla y finalmente 2 de estiramiento, parecido al proceso que se hace en una transformación afín.

Finalmente otro avance importante que se ha dado es el PIV 3D cuya ventaja es el hecho de que se conoce el campo completo de las velocidades y, con imágenes holográficas, se elimina el problema de que una partícula tape a otra ya que se puede aislar un plano a la vez. Entre las desventajas se encuentra el costo, que la persona que lo maneje debe ser experta y que si se buscan mediciones en flujos turbulentos, es difícil grabar suficientes hologramas para obtener estadísticas turbulentas estables.

## 5.2 Correlación

Como se puede apreciar, el desarrollo de la técnica más robusta, no intrusiva, para medir velocidades de flujos turbulentos para poder llegar a ser lo que es, hoy en día, requirió de gran esfuerzo, distintos grupos de investigación, muchos años, cambios y un proceso de aprendizaje exhaustivo.

En este capítulo se busca medir velocidades de flujos con el mismo principio que el PIV, pero con la diferencia del tipo de imágenes, para lo cual se entenderá el proceso partiendo de la simplicidad.

La correlación, en su forma más simple, entre 2 funciones,  $f$  y  $w$ , es:

Ecuación 5.1

$$c(x, y) = \sum_s \sum_t f(x, y)w(x + s, y + t)$$

Para  $x=0,1,2,\dots,M-1$ ,  $y=0,1,2,\dots,N-1$ , y la operación se realiza sobre la región en la que se traslapan las funciones,  $w$  y  $f$ . En la siguiente figura se ilustra el procedimiento, se asume que el origen de  $f$  se encuentra en la esquina superior izquierda y el origen de  $w$  en el centro. Para un valor de  $(x,y)$  por ejemplo  $(x_0, y_0)$  se aplica la ecuación 1 otorgando un solo valor de  $c$ . Al variar  $x$  y  $y$ ,  $w$  se mueve a lo largo del área de la imagen. El valor máximo de  $c$  indica la posición en la que  $w$  tiene una mayor correspondencia con  $f$ . Es importante notar que la ecuación no tiene sentido cerca de las orillas ya que parte de la función  $w$  estaría afuera de la función  $f$ , es posible eliminar este problema al duplicar filas y columnas de la función  $f$ , sin embargo, los cálculos perderían precisión con el error siendo proporcional al tamaño de  $w$ .

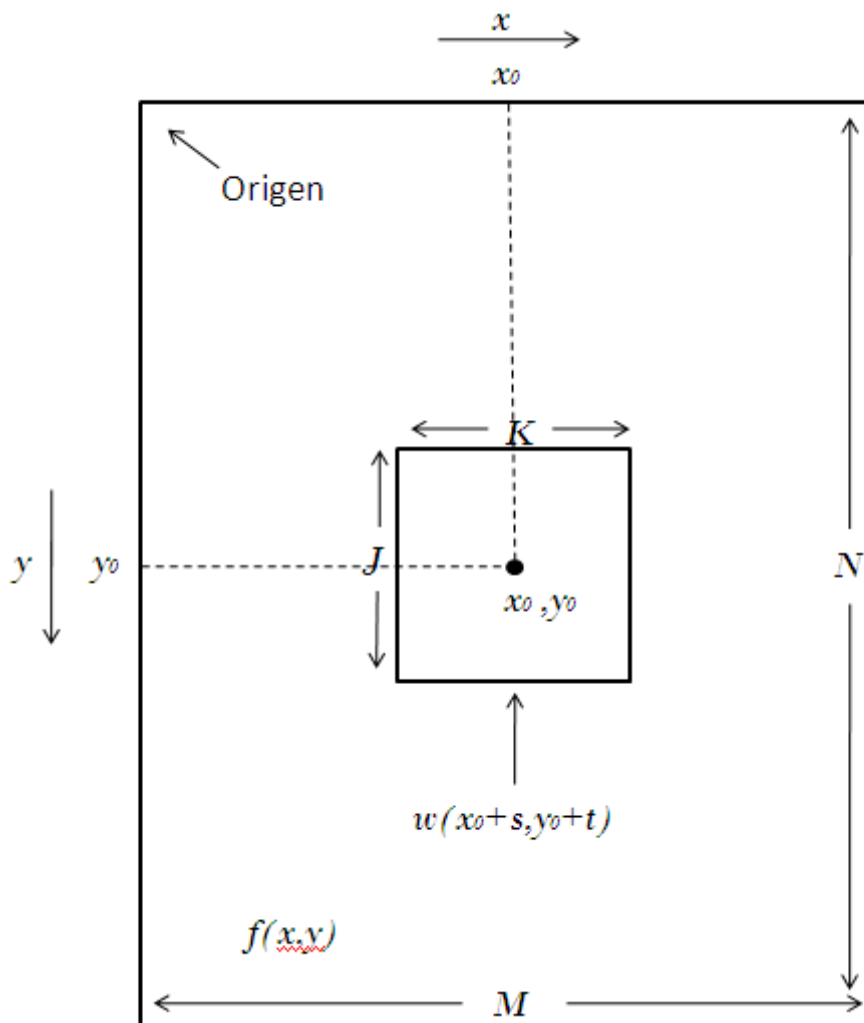


Figura 5.1 Esquema de la Correlación entre una ventana menor y una mayor.



La función de correlación pasada tiene la desventaja de ser sensible a los cambios de amplitud de  $f$  y  $w$ . Por ejemplo, si se duplican los valores de  $f$  se duplican los valores de  $c$ , para superar esta dificultad se puede normalizar la ecuación y así obtener lo que se conoce normalmente como coeficiente de correlación:

Ecuación 5.2

$$\gamma(x, y) = \frac{\sum_s \sum_t [f(s, t) - \bar{f}(s, t)][w(x + s, y + t) - \bar{w}]}{\left\{ \sum_s \sum_t [f(s, t) - \bar{f}(s, t)]^2 \sum_s \sum_t [w(x + s, y + t) - \bar{w}]^2 \right\}^{\frac{1}{2}}}$$

Donde  $x=0,1,2,\dots,M-1, y=0,1,2,\dots,N-1$ ,  $\bar{w}$  es el valor promedio de los píxeles en  $w$  (se calcula una sola vez),  $\bar{f}$  es el promedio del valor  $f$  en la región coincidente con la ubicación de  $w$ . El coeficiente de correlación,  $\gamma$ , va de -1 a 1, independiente de los cambios en la amplitud de  $f$  y  $w$ . Un ejemplo simple puede ser el de encontrar la localización de una letra en el abecedario sin conocimiento previo, Figura 5.2.

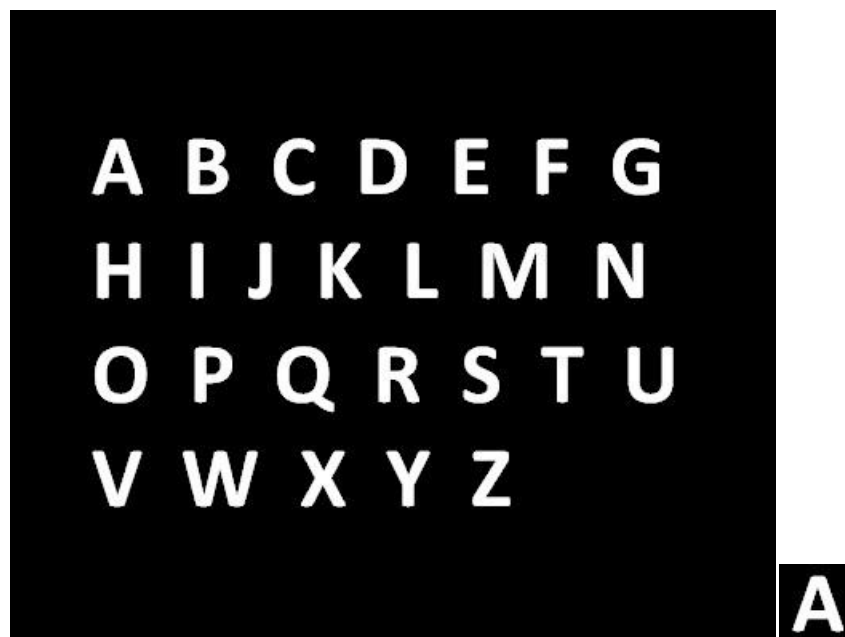


Figura 5.2 (I) El abecedario representado en una imagen (D) La sub-ventana que representa la letra A

0	0	0	0	0	0	0	0	0	3	195	255	252	249	255	255	252	255	255	253	255	250
0	0	0	0	0	0	0	0	2	0	248	255	255	254	255	253	255	254	254	255	251	255
0	0	0	0	0	0	0	0	0	0	150	255	255	247	255	255	246	245	255	253	255	253
0	0	0	0	0	0	0	0	1	213	246	255	254	254	253	209	209	255	255	254	255	253
0	0	3	1	0	2	4	0	3	241	255	255	250	254	255	151	155	255	253	254	254	255
3	0	0	0	9	0	3	0	154	255	252	253	255	248	245	8	0	247	255	255	251	255
0	5	0	1	0	0	6	1	202	254	255	252	253	255	211	0	1	209	255	252	255	254
0	1	0	2	0	5	0	0	255	255	255	249	255	251	156	9	0	159	255	252	249	255
5	0	7	1	0	0	4	171	255	252	254	255	254	253	0	0	5	0	245	255	255	250
0	6	0	5	0	0	4	208	251	251	255	255	251	208	2	5	3	0	203	255	254	255
4	1	0	0	5	3	0	255	255	255	253	254	255	145	4	0	0	2	154	255	248	254
0	0	1	5	0	0	170	252	254	255	251	255	247	3	1	0	4	0	5	241	254	255
0	0	6	0	0	0	224	255	252	255	251	255	208	0	0	0	0	0	0	208	255	251
5	0	0	8	0	107	255	255	255	247	255	252	152	0	1	0	0	1	0	152	252	255
0	0	3	0	8	172	247	255	253	255	255	252	1	4	0	2	2	0	4	1	252	255
0	5	0	0	2	234	255	254	254	253	255	199	0	0	0	4	4	0	0	0	199	255
0	2	3	0	99	255	255	255	253	254	254	154	1	2	7	0	0	7	2	1	154	254
1	0	0	2	183	252	254	251	255	255	255	252	255	254	244	255	255	244	254	255	252	255
0	7	0	0	238	255	255	252	255	255	253	255	254	255	255	255	255	255	255	254	255	253
1	0	3	129	250	255	247	255	255	251	255	254	255	253	255	254	254	255	253	255	254	255
1	0	2	200	253	255	255	254	255	252	255	254	254	255	250	254	253	255	250	253	254	255
2	0	0	237	255	254	255	255	255	255	254	250	255	252	255	255	255	250	255	255	255	250
1	5	126	255	251	255	255	251	252	206	137	142	139	139	138	139	142	144	145	135	144	139
0	0	195	255	255	255	250	255	255	131	0	0	8	0	1	3	0	1	0	5	0	0
11	0	249	254	253	253	255	250	243	0	3	0	0	6	3	0	4	0	11	0	0	3
0	153	249	255	250	250	255	255	206	0	0	0	4	0	2	3	0	0	0	9	0	0
0	221	251	255	255	254	252	255	111	10	0	5	0	1	0	0	0	4	0	0	2	0
0	245	255	255	252	255	254	251	3	1	0	0	0	1	1	2	2	0	4	0	0	0
136	255	255	252	255	255	255	208	0	0	0	0	0	0	0	0	0	0	0	0	0	0
127	255	253	253	255	251	249	125	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	103	139	146	3	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figura 5.3 Representación numérica de la Figura 1.2 (letra A)

Si se desea localizar la letra A en la imagen binaria anterior, lo primero que se debe hacer es representar la letra A en un arreglo de números, al igual que la imagen, después se aplica la ecuación a lo largo de la imagen original, para la ubicación de cada pixel se obtiene un coeficiente de correlación, lo que se puede representar de varias maneras, graficando la superficie de la correlación o aprovechando una manera visual. Resultando el pico de correlación más alto, Figura 5.4, o el pixel más brillante, Figura 5.5, como la posición en la que se encuentra la letra.

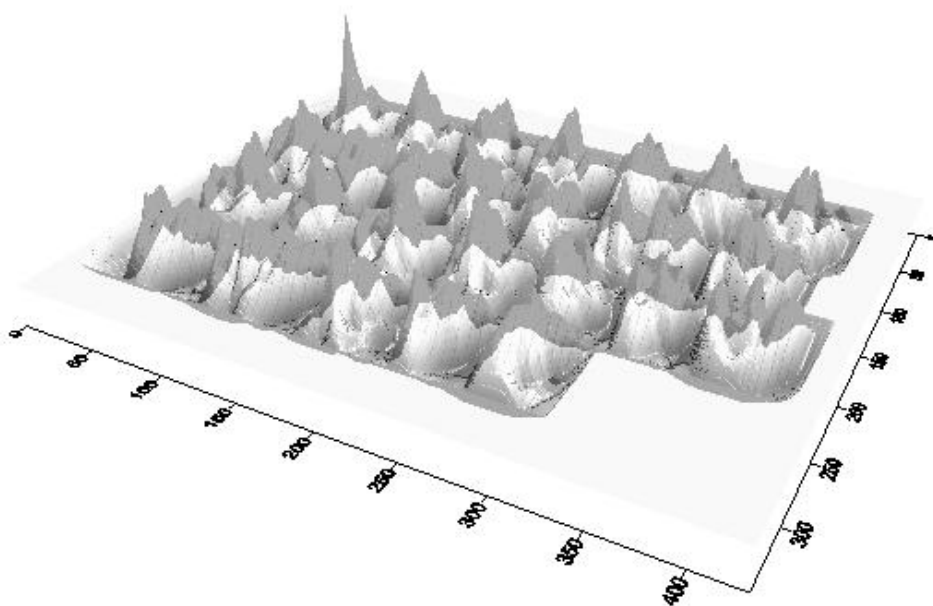


Figura 5.4 Superficie de Correlación.

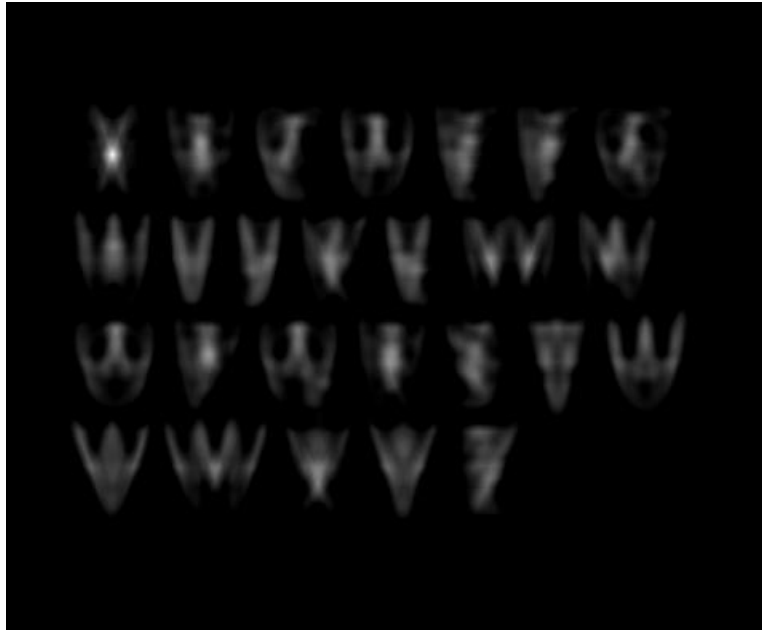


Figura 5.5 Representación visual del campo de correlación.

### 5.3 Aplicación de la técnica a casos reales

Conociendo el funcionamiento básico de la correlación se pueden aprovechar sus bondades al combinarlo con imágenes de alta velocidad. Este proceso lo siguió Ryu, con una iluminación tipo “shadowgraphy”, de sombras chinescas. En su tesis doctoral (RYU, 2006) y en artículos como (Ryu *et al.*, 2007) donde las mediciones se realizaban en flujos con burbujas, característico de la turbulencia.

Las imágenes con las que se trabaja para medir velocidades fueron tomadas con una cámara de alta velocidad que tiene la capacidad de tomar 8000 cuadros por segundo, en los fenómenos en los que se trabajó fueron 2: una ola con rotura tipo voluta y un chorro provocado por una ola entrando a una tuba invertida.

El algoritmo trabaja de la siguiente forma: lee el video cuadro por cuadro, en cada cuadro que lee guarda la información del cuadro anterior, a este se le llamará el cuadro A y al siguiente cuadro B; divide al cuadro A en sub-imágenes cuyo tamaño de ventana puede ser cuadrado o rectangular; a cada sub-imagen del cuadro A se realiza el proceso visto en el apartado anterior con la diferencia que no realiza todo el proceso en todo el cuadro B sino hasta una distancia definida, esto es, se define un nivel de correlación aceptable. Finalmente se realiza este proceso para el video entero encontrando la velocidad al dividir el desplazamiento entre el número de cuadros por segundo para el cual se programa la cámara. Cada uno de los pasos que se han descrito tienen una razón de ser, que se explicará a continuación.

El tamaño de la ventana es de extrema importancia y depende del fenómeno que se quiera medir. El tamaño de la ventana no solo influye considerablemente en los resultados que se obtienen sino también en el número de vectores por imagen que se encontrarán, entre más grande sea la ventana, menor es el número de vectores; es práctica común manejar números impares para que el centro de la imagen sea el centro de un pixel.

Se define una distancia para la búsqueda de la sub-imagen del cuadro A en el cuadro B por tres razones: para reducir el gasto computacional, por probabilidad y por el conocimiento "a priori" del problema. El gasto computacional, como es lógico pensar, a mayor distancia de búsqueda aumenta el número de operaciones a realizar, si se busca el pico de correlación hasta a 10 pixeles de distancia, se necesita realizar la operación 441 veces, si se busca hasta 15 pixeles 961 veces. Entre más lejos se busque es más probable que exista una configuración de partículas que se parezca, sin ser la real. El conocimiento "a priori" del fenómeno da una idea de las velocidades máximas que se pueden encontrar, por ejemplo, si 100 pixeles representan 5 cm y se grabó a 3000 cuadros por segundo, la velocidad del fluido, si hay un desplazamiento de 50 pixeles, sería de 83 m/s. Entonces, si el fenómeno que se analiza tiene velocidades máximas de 10 m/s no tiene caso buscar más allá de 8 pixeles.

Se define un nivel de correlación aceptable ya que existe la posibilidad de que el pico de correlación mayor no indique el verdadero desplazamiento, esto sucede generalmente cuando las correlaciones que se obtienen son bajas.

Es importante grabar a una tasa de cuadros por segundo que concuerde con el fenómeno, por ejemplo, si se graba a 8000 cuadros por segundo y 50 pixeles representan 5cm; un desplazamiento de 1 pixel representa una velocidad mínima medible de 7.77 m/s.

Tomando en cuenta todos estos detalles, se procedió a encontrar los vectores de desplazamiento en un video de alta velocidad de una ola rompiendo, no con el objetivo de tomar velocidades con precisión sino de observar si el proceso realizado arroja resultados coherentes.

Se observa que en la cresta de la ola las velocidades son superiores, Figura 5.6, que los vectores de velocidad en la zona de turbulencia parecen formar un remolino, Figura 5.7, que las velocidades en la superficie son falsas y que las gotas de agua que salen disparadas, Figura 5.8, son seguidas de buena manera por el algoritmo. Sin embargo, al no tener una referencia en donde escalar los pixeles a cm, solo se obtienen velocidades en pixeles.

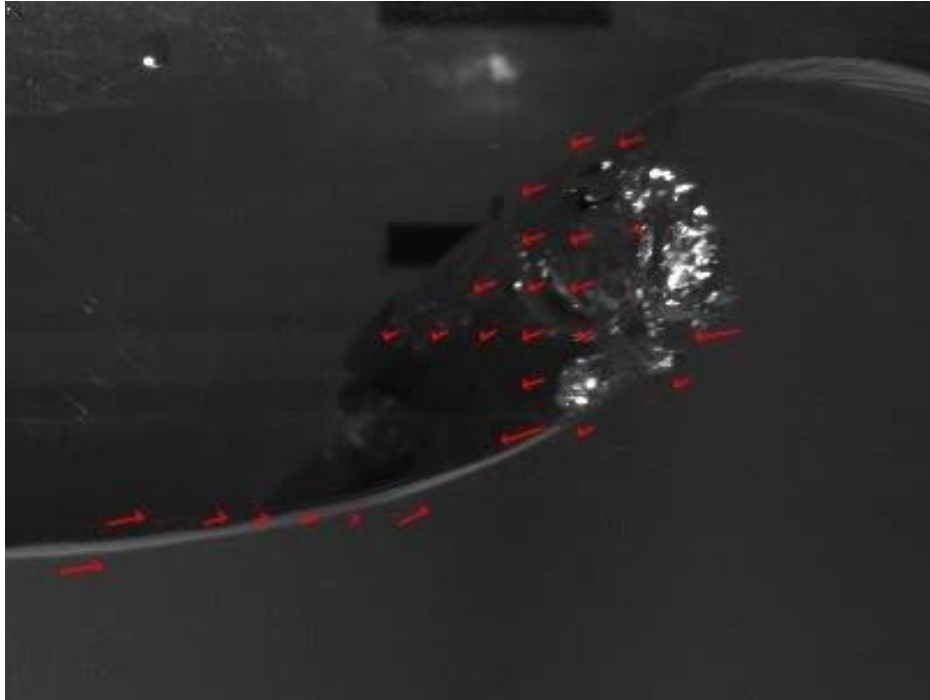


Figura 5.6 Velocidades encontradas en una ola antes de romper.

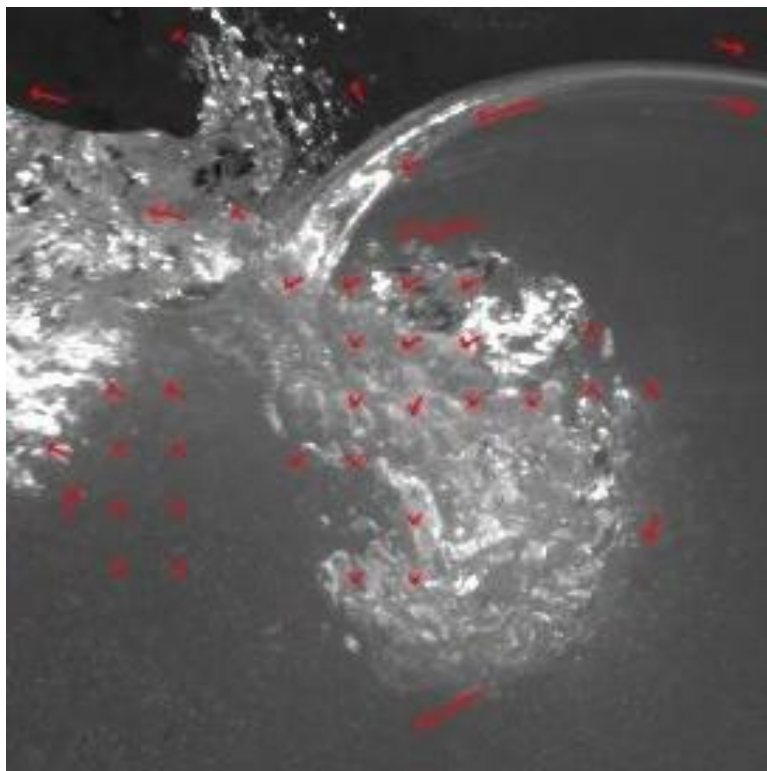


Figura 5.7 Velocidades encontradas después de que la rotura ocurrió.

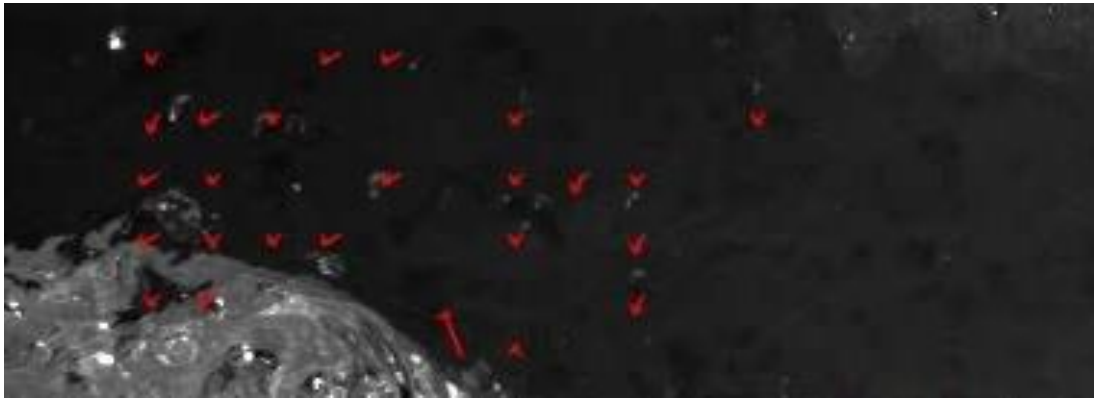


Figura 5.8 Velocidades de gotas cayendo provocadas por la salpicadura de la ola.

Para medir las velocidades en el chorro de agua, el procedimiento consiste en definir un área de procesamiento a la salida del chorro y se obtiene el promedio de velocidades para cada cuadro la escala de pixeles a cm se realiza con ayuda del diámetro externo del tubo que es conocido. La equivalencia de este caso específico es de 68.5 pixeles a 2.54 cm. Cabe mencionar que el video fue grabado originalmente a 8603 Hz, la velocidad mínima era de 3.20 m/s, lo que indica que la frecuencia de grabación es muy alta para el fenómeno, por lo que se tuvieron que considerar solo uno de cada 10 cuadros, resultando en un nuevo video con una frecuencia de 860 Hz.



Figura 5.9 Definición del área de cálculo.



Figura 5.10 Velocidades encontradas en un instante dado.

Al ser variables el tamaño de la ventana y la correlación mínima aceptable, se tiene una gran variedad de combinaciones por lo que es necesario probar valores hasta encontrar el funcionamiento óptimo para cada caso específico. Existen 2 grandes problemas con la forma en que se llevó a cabo este experimento: el primero es que no hay información contra la cual comparar de manera confiable los resultados y el segundo la iluminación. El criterio para considerar óptimo al método es teniendo una evolución de las velocidades concordante a las obtenidas por el PIV y sin tener vectores extraños.

En el video se observan 2 chorros de agua, por lo que se esperan 2 picos de velocidad en las gráficas. Se analizaron ventanas de tamaño 9 pixeles hasta 33 pixeles cada 2 pixeles. Se observó que de 9 a 15 pixeles se encuentran muchos desplazamientos no reales, ya que entre chorro y chorro encuentra velocidades a pesar de que no hay movimiento.

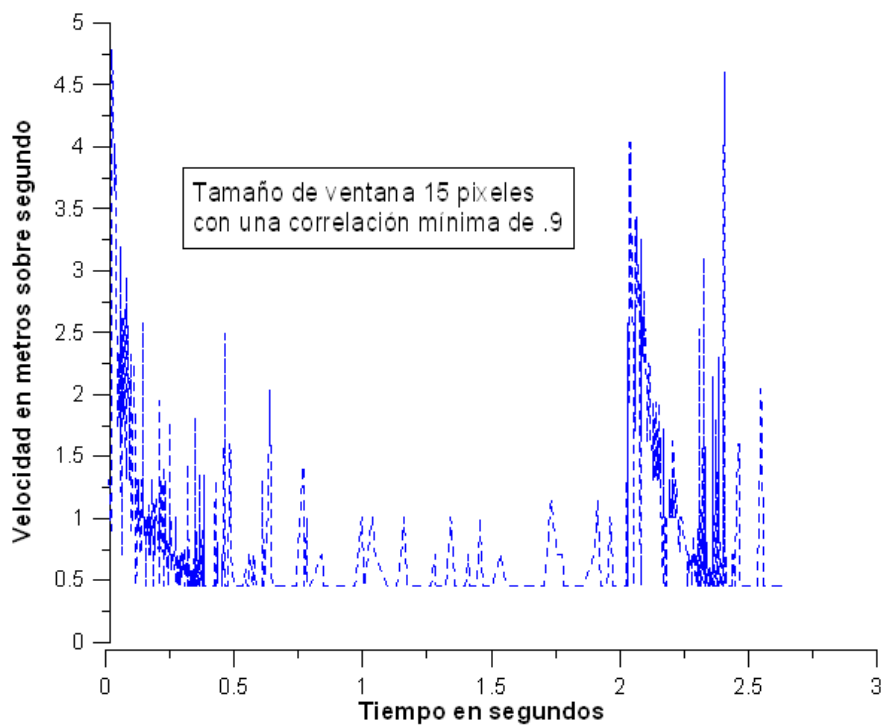


Figura 5.11 Velocidades promedio encontradas con un tamaño de ventana de 15 pixeles.

El siguiente análisis se realizó a partir de una correlación mínima de .9 y tamaño de ventana de 17 a 33 pixeles, observando que cuando la ventana es grande se encuentran pocos vectores.





## 5.4 Conclusiones

El análisis de correlación para encontrar velocidades en flujos turbulentos con una cámara de alta velocidad es factible. Sin embargo, existen ineficiencias tanto en el algoritmo que se desarrolló como en el experimento realizado.

Entre las ineficiencias detectadas en el experimento se cuentan la iluminación, se encontraron reflejos de luz hacia la cámara que no permiten distinguir con claridad las burbujas, que son a fin de cuentas, los trazadores. Con una iluminación proveniente de la parte trasera del chorro de agua se hubiesen encontrado mayor número de vectores de velocidad, esta iluminación es la de sombras chinescas (shadowgraphy).

No se puede decir con confianza que las velocidades medidas son reales ya que no se comparó contra datos reales, sin embargo, la evolución de las velocidades es consistente con la reportada por (Chávez, 2010).

El algoritmo tiene varias áreas de mejora: la primera es la revisión de la consistencia espacio-temporal de los vectores, de esta manera la mayor parte de los vectores no reales serían automáticamente identificados; la segunda es el gasto computacional, se podría reducir si se realiza en el espacio de la frecuencia; la tercera es dada por las bondades de la cámara de alta velocidad, al tener tantas imágenes a tan alta frecuencia cada cuadro se podría comparar contra 10 o más, obteniendo así un mayor número de mediciones y al usar métodos estadísticos, encontrar velocidades con mayor confiabilidad.

Finalmente se considera que esta técnica, por si misma, puede ser aplicable no solo en mediciones de flujos turbulentos sino también en mediciones de velocidades de partículas como lo podría ser la velocidad de caída de sedimentos.



## 6 CONCLUSIONES GENERALES

El propósito general de esta tesis fue demostrar la factibilidad de usar sensores capaces de adquirir información cuantitativa, en forma de matriz, para describir información tanto cuantitativa como cualitativa del mundo físico. Para lo cual se requirió revisar el mundo de la óptica, la matemática, el procesamiento de imágenes digitales, el reconocimiento de patrones y el de visión computacional, entre otros.

De una manera más específica se usó una imagen o conjunto de imágenes para realizar mediciones en distintos rubros de la hidráulica marítima, tanto en laboratorio como en campo, esto fue posible gracias a la naturaleza multi-disciplinaria de la visión computacional y el desarrollo tecnológico que ha existido a lo largo de los últimos años. Así que se podría resumir el presente trabajo, con el objetivo de tener una idea más clara del trabajo realizado, en 2 enfoques: la hidráulica marítima y la visión computacional.

Desde el enfoque de la hidráulica marítima: se realizaron mediciones no intrusivas de la línea de costa de Cancún; del ascenso máximo en pendientes impermeables estáticas; y de velocidades de flujos turbulentos.

Desde el enfoque de visión computacional se usaron varias técnicas, conceptos y algoritmos: transformaciones de un modelo de color a otro; el cálculo de imágenes con significado estadístico; algoritmos para discriminar clases de píxeles; transformaciones geométricas que permitieran el mapeo de coordenadas de píxeles a coordenadas del mundo real; calibraciones de cámaras para disminuir el error introducido por la naturaleza física de los lentes; la creación de imágenes a partir de videos para obtener una distinta interpretación del mismo; algoritmos basados en umbrales y en el movimiento para la segmentación; y finalmente, el uso de la correlación con el objetivo de reconocer objetos de una imagen en otra.

El valor de la unión de ambas ciencias o el uso de una sobre la otra es el potencial de las mediciones que se pueden realizar, ya que son no intrusivas y son capaces de manejar un rango amplio de escalas espacio-temporales, características de amplia importancia en el mundo de la marítima.

Los algoritmos usados para la extracción de la línea de costa permiten un análisis casi inmediato de las imágenes que ha recolectado el sistema IBIS por los últimos 5 años y abren el camino para mediciones de otro tipo, como lo es la medición indirecta de la altura de ola, la dirección del oleaje, etc. Haciendo una extensión del clasificador bayesiano de 1-D a n-D permitiría el uso de información de más de 2 canales de información para la discriminación de los píxeles playa-mar, teniendo mayor confiabilidad de la discriminación y más posibilidades de mediciones reales en situaciones energéticas que impidan la obtención de imágenes de buena calidad.

Se pueden realizar mediciones de ascenso máximo, tanto en laboratorio como en campo, con los programas escritos para el capítulo 4.

Lo descrito en el capítulo 5 se puede usar como una herramienta numérica de la cámara de alta velocidad para medir las velocidades de flujos turbulentos, o de partículas, como lo es el sedimento marino.

## 6.1 Recomendaciones

Hay varias áreas de oportunidad en esta tesis, tanto de algoritmos como detalles de los experimentos sobre los que trabajan.

En los experimentos resalta el uso de la iluminación adecuada, la calibración previa de las cámaras, el uso de referencias para realizar la transformación de píxeles a cm y la frecuencia de la toma de datos pensada para el fenómeno a estudiar.

Es necesario realizar la comprobación adecuada de los algoritmos, ya que no se tienen datos reales contra los cuales comparar las mediciones: para la extracción de la línea de costa se podrían tomar mediciones in situ y compararlas contra la transformación geométrica y calcular el error, que se esperaría crezca conforme se alejen del centro de la imagen; para las mediciones del ascenso máximo se podría instrumentar una rampa que tenga un arreglo de sensores de presión lo suficientemente sensibles para detectar la punta del ascenso máximo; para el velocímetro de imágenes se podría comparar las velocidades contra una herramienta robusta como el PIV en el chorro de agua o sedimentando un flujo constante.

## 6.2 Futuras Líneas de Investigación

Es posible combinar lo visto en esta tesis con el fin de tener un software que trabaje con las imágenes del sistema IBIS, que sea capaz de obtener distintos parámetros en la zona cercana a la costa, como lo es la línea de costa, mediciones de ascenso máximo, imágenes panorámicas, alturas de ola, direcciones, periodos, etc.(Davidson, y otros, 2007)

Haciendo una extensión del clasificador bayesiano de 1-D a n-D permitiría el uso de información de más de 2 canales de información para la discriminación de los píxeles playa-mar, teniendo mayor confiabilidad de la discriminación y más posibilidades de mediciones reales en situaciones energéticas que impidan la obtención de imágenes de buena calidad.

El velocímetro de imágenes tiene una amplia zona de mejora, como lo es la revisión posterior de la consistencia espacio-temporal de los vectores encontrados, además de programar otros métodos de búsqueda como lo es el MQD.

## 7 Bibliografía

- Aarninkhof, S. (2003). Nearshore bathymetry derived from Video Imagery, PhD Thesis. *Delft University of Technology, The Netherlands* .
- Aarninkhof, S.G.J., Turner I.L., Dronkers, T.D.T., Caljouw, M., & Nipius, L. (2003). A video-based technique for mapping intertidal beach bathymetry. *Coastal Engineering* , 275-289.
- Adrian, R. J. (1984). Scattering particle characteristics and their effect on pulsed laser measurements of fluid flow: speckle velocimetry vs particle image velocimetry. *Appl. Opt.* 23 .
- Adrian, R. J. (2004). Twenty Years of Particle Image Velocimetry. *International Symposium of Laser Techniques to Fluid Mechanics* .
- Adrian, R., & Yao, C. (1983). Development of Pulsed Laser Velocimetry of Measurement of Fluid Flow. *Proceedings 8th Symposium on Turbulence* , 171-186.
- Anquez, J., Angelini, E., Bloch, I., Merzoug V., Bellaiche-Millischer, A.E., & Adamsbaum, C. (2007). Interest of the Steady State Free Precession (SSFP) sequence for 3D modeling of the whole fetus. *eProceedings of the 29th Annual International Conference of the IEEE EMBS* , ThP2B2.7.
- Bailey, D., & Shand, R. (1994). Determining Wave Run-up using Automated Video Analysis. *Proceedings 2nd NZ Conference on Image and Vision Computing* .
- Barker, D., & Fourney, M. (1977). Measuring fluid velocities with speckle patterns. *Optics Letters* , 135-137.
- Bechle, A. J., & Wu, C. H. (2011). Virtual wave gauges based upon stereo imaging for measuring surface wave characteristics. *Coastal Engineering* , 305-316.
- Bildstein, K. L. (1993). *White Ibis: Wetland Wanderer*. Washington D.C.: Smithsonian Institution.
- Bowen, A.J., Inman, D.L., & Simmons, V.P. (1968) Wave 'Set-Down' and Set-Up. *Journal of Geophysical Research* , 2569-2577.
- Caljouw, M. (2000). Video-based monitoring of the Egmond beach and shoreface nourishments. *MSc Thesis* .
- Chávez Cárdenas, X. (2010). Blow-jet, tecnología oceanomotriz de la conversión de energía del oleaje (WEC), en sus etapas: concepto de diseño y caracterización del prototipo a escala en un canal de olas. *Tesis de Maestría, UNAM* .
- Conner, N. (2010). *The everything classical mythology book*. Massachusetts: Adams Media.
- Crowson, R., Birkeimer, W., Klein, H., & Miller, H. (1988). SUPERDUCK Nearshore Processes Experiment: Summary of Studies CERC Field Research Facility. *Coastal Engineering Research Center, Washington, D.C.*
- Davidson, M., Van Koningsveld, M., de Kruif, A., Rawson, J., Holman, R., Lamberti, A., y otros. (2007). The CoastView project: Developing video-derived Coastal State Indicators in support of Coastal Zone Management. *Coastal Engineering* , 463-475.

- De Vriend, H. J. (1997). Prediction of aggregated-scale coastal evolution. *Proc. of Coastal Dynamics* .
- Flick, R., Lowe, R., Freilich, M., & Boylls, J. (1979). Coastal and Laboratory Wavestaff System. *IEEE Oceans*
- Geissler, P., & Jahne, B. (1995). Measurements of Bubble Size Distributions with an Optical Technique Based on Depth from Focus. *Air-Water Gas Transfer* , 351-362.
- González Leija, M. B. (2009). *Análisis de las Variaciones Morfodinámicas en la playa de Cancún, Quitana Roo: Herramientas para el Manejo Costero*. Tesis de Maestría.
- Gonzalez, R. C., & Woods, R. E. (2002). *Digital Image Processing*. New Jersey: Prentice-Hall.
- Greated, C. A., & Emarat, N. (2000). Optical studies of wave kinematics. *Advances in Coastal and Ocean Engineering* .
- Guza, R., & Thornton, E. (1982). Swash Oscillations on a Natural Beach. *Journal of Geophysical Research* , 483-491.
- Hamm, L., Capobianco, M., Dette, H.H., Lechuga, A., Spanhoff, R., & Stive, M.J.F. (2002). A summary of european experience with shore nourishment. *Coastal Engineering* , 237-264.
- Hassan, Y. A., Schmidl, W. D., & Ortiz, J. (1998). Investigation of three-dimensional two-phase flow structure in a bubbly pipe. *Measurement Science and Technology* , 633-645.
- Hedges, T.S., & Mase, H.(2004). Modified Hunt's Equation Incorporating Wave Setup. *Journal of Waterway, Port, Coastal and Ocean Engineering* , 109-113.
- Holland, K.T. & Holman, R.A. (1993) Statistical Distribution of Swash Maxima on Natural Beaches. *Journal of Geophysical Research* , 10271-10278.
- Holland, K., Raubenheimer, B., Guza, R., & Holman, R. (1995). Runup kinematics on a natural beach. *Journal of Geophysical Research* , 4985-4993.
- Holland, K.T. & Holman, R.A. (1997) Video estimation of foreshore topography using trinocular stereo. *Journal of Coastal Research* , 81-87.
- Holman, R.A. & Guza, R.T. (1984) Measuring Run-Up on a Natural Beach. *Coastal Engineering* , 129-140.
- Holman, R.A., Howd, P.A., Oltman-Shay, J., & Komar, P.D. (1990) Observations of the swash expression of far infragravity wave motions. *International Proceedings of the 22nd Conference of Coastal Engineering*, 1242-1253.
- Holman, R. A., & Stanley, J. (2007). The history and technical capabilities of Argus. *Coastal Engineering* , 477-491.
- Hunt, I.A. (1959). Design of seawalls and breakwaters. *Journal of Waterways Harbors Division* , 123-152.
- Lindken, R. & Merzkirch, W. (2001). A novel PIV technique for measurements in multi-phase flows and its application to two-phase bubbly flows. *4th International Symposium on Particle Image Velocimetry*. Paper A231.

- Lippmann, T., & Holman, R. A. (1989). Quantification of sand bar morphology: a video technique based on wave dissipation. *Journal of Geophysical Research* , 995-1011.
- Lippmann, T., & Holman, R. (1990). The spatial and temporal variability of sand bar morphology. *Journal of Geophysical Research* , 575-590.
- Longuet-Higgins, M., & Stewart, R. (1962). Radiation stress and mass transport in gravity waves with applications to surf beats. *Journal of Fluid Mechanics* , 481-504.
- Lourenco, L. M., Gogineni, S. P., & Lasalle, R. T. (1994). On-line particle image velocimeter: an integrated approach. *Applied Optics* , 2465-2470.
- Maerz, N. H., Palangio, T. C., & Franklin, J. A. (1996). WipFrag Image Based Granulometry System. *Proceedings of the FRAGBLAST 5 workshop on Measurement of Blast Fragmentation* , 91-99.
- Marr, D. (1982). *Vision*. W.H. Freeman and Company, New York .
- McGlashan, D., Duck, R., & Reid, C. (2005). Defining the foreshore: coastal geomorphology and British laws. *Estuarine, Coastal and Shelf Science* , 183-192.
- Melen, T. (1994). *Geometrical modelling and calibration of video cameras for underwater navigation*. Norges tekniske hogskole: Institutt for teknisk kybernetikk.
- Meynart, R. (1983). Instantaneous velocity field measurements in unsteady gas flow by speckle velocimetry. *Applied Optics* , 535-540.
- Molder, C., Serban, F., Vizitiu, I., & Stanciu, M. (2008). Image processing for sediment classification. *Conference on Computational Intelligence, man-machine and cybernetics* .
- Morck, T., Andeersen, P.E., & Westergaard, C.H. (1993) Processing speed of photorefractive optical correlators in PIV processing. In: Adrian, R.J., Durao, D.F.G., Durst, F., & Heitor, M.V. (eds) *Laser techniques and applications in fluid mechanics*. Springer, Berlin Heidelberg New York, 227-242.
- Nordstrom, K., Jackson, N., Smith, D., & Weber, R. (2006). Transport of horeshoe crab eggs by waves and swash on an estuarine beach: Implications for foraging shorebirds. *Estuarine, Coastal and Shelf Science* , 438-448.
- Osborne, P., & Rooker, G. (1999). Sand re-suspension events in a high energy infragravity swash zone. *Journal of Coastal Research* , 74-86.
- Osorio, A. (2005). Desarrollo de técnicas y metodologías basadas en sistemas de vídeo para la gestión de la costa. *Tesis Doctoral, Universidad de Cantabria, España* .
- Pedrozo, A., Ruiz, A., Torres, A., Mendoza, E., & Silva, R. (2011). Laboratory investigation of pressure gradients induced by plunging breakers. *Coastal Engineering* , 222-238.
- Plant, N., & Holman, R. (1997). Intertidal beach profile estimation using video images. *Marine Geology* .
- Rivillas, G., Pedrozo, A., Silva, R., Torres, A., & Gutierrez, C. (2011). Estimation of the velocity field induced by plunging breakers in the surf and swash zones. *Experiment in fluids* .

- RYU, Y. U. (2006). Extreme Wave Impinging and Overtopping. *PhD Tesis, Texas A&M University* .
- Ryu, Y., & Chang, K. (2008). Green water void fraction due to breaking wave impinging and overtopping. *Experiments in Fluids* , 883-888.
- Ryu, Y., Chang, K., & Mercier, R. (2007). Runup and green water velocities due to breaking wave impinging and overtopping. *Experiments in Fluids* .
- Sallenger, A.H.(2000) Storm Impact Scale for Barrier Islands. *Journal of Coastal Research* , 890-895.
- Salmon, S., Bryan, K., & Coco, G. (2007). The use of video systems to measure run-up on beaches. *Journal of Coastal Research* , 211-215.
- Slama, C. C. (1980). *Manual of Photogrammetry, 4th Edition, American Society of Photogrammetry*.
- Southgate ,H.N., Wijnberg, K.M., Larson, M., Capobianco, M., & Jansen, H. (2003). SA nalysis of field data of coastal morphological evolution over yearly and decadal time scales. Part 2: Non-linear techniques. *Journal of Coastal Research* , 776-789.
- Stockdon, H.F., Holman R.A., Howd, P.A., & Sallenger Jr., A.H. (2006) Empirical parameterization of setup, swash, and runup. *Coastal Engineering* , 573-588.
- Xiao, H., & Wenrui, H. (2008). Numerical modeling of wave runup and forces on an idealized beachfront house. *Ocean Engineering* , 106-116.
- Zhang, X. (2003). Surface image velocimetry for measuring short wind wave kinematics. *Experiments in Fluids* , 653-665.