



UNIVERSIDAD CATÓLICA DEL NORTE

FACULTAD DE INGENIERÍA Y CIENCIAS GEOLÓGICAS

DEPARTAMENTO DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN

**IDENTIFICACIÓN AUTOMÁTICA DE
FRACTURAS EN ROCAS SOBRE IMÁGENES DE
POZOS DE SONDAJES**

Propuesta de Tesis para optar al grado de
Magister en Ingeniería Informática

Diego P. Urrutia Astorga

Profesor Guía:

Dr. Claudio Meneses Villegas

Antofagasta, Chile

Diciembre 2014

Resumen

La geofísica de pozos es un método utilizado para obtener los rasgos estructurales de un sondaje. Consiste en una serie de mediciones obtenidas por un instrumento compuesto de uno o más sensores, el cual es introducido a lo largo del pozo (o perforación de barreno) y va registrando información a medida que avanza. Entre la información que puede ser extraída se encuentra la litología, alteración, mineralización, estructuras, etc. De los datos obtenidos, uno importante es identificar las fracturas presentes, las cuales pueden ser de origen natural o inducidas en el proceso de perforación. La importancia radica en que son una buena fuente de estimación de la orientación del *stress* que posee la roca y partir de esta información, interpretar el comportamiento geofísico del pozo.

Una forma de extraer las fracturas es analizar las imágenes obtenidas mediante técnicas de televisión acústica. Este proceso es desarrollado por geólogos, pero de forma manual con ayuda de un *software* de visualización.

Se presenta un método automático de identificación de líneas base y parametrización de fracturas, el cual se encuentra basado en técnicas de procesamiento de imágenes y un proceso de votación implementado por medio de la transformada de Hough. Además, con el fin de evaluar la calidad de identificación y parametrización, se calculan algunas métricas sobre los resultados obtenidos y se comparan con la clasificación de fracturas desarrolladas por un experto.

El método propuesto es aplicado sobre dos conjuntos de imágenes: uno generados de manera sintética a partir del análisis de imágenes reales del pozo; y otro obtenido a partir de imágenes de pozos de la II región de Antofagasta. A pesar que la calidad de la identificación y parametrización obtenida, no es comparable a la realizada por un humano, la aplicación del método propuesto muestra ser una opción válida como pre-proceso de clasificación de fracturas sobre imágenes de televisión acústica.

Índice general

1. Introducción	5
1.1. Motivación de la Tesis	6
1.2. Objetivos	7
1.2.1. Objetivo General	7
1.2.2. Objetivos Específicos	7
1.3. Contribución y visión general de la Tesis	8
1.3.1. Contribución	8
1.3.2. Visión General	8
2. Geofísica de Pozos	9
2.1. Televisión Acústica	10
2.1.1. Tiempo de Tránsito	12
2.1.2. Amplitud	13
2.1.3. Factores que afectan la adquisición de los datos e interpretación	13
2.2. Andes Geofísica	14
3. Problema de Investigación	18
3.1. Detección y Parametrización de Estructuras	18
3.1.1. <i>Breakouts</i> (Deformaciones)	18
3.1.2. Discontinuidades (Fracturas)	20
3.1.2.1. Fracturas de procedencia natural	20
3.1.2.2. Fracturas inducidas por perforación	21
3.1.2.3. Fracturas abiertas o cerradas	21
3.2. Identificación de Fracturas	23

3.2.1. Proceso de identificación manual de fracturas	24
4. Estado del Arte	27
4.1. Discusión de Trabajos	28
5. Metodología	32
5.1. Procesamiento de las imágenes	33
5.1.1. Obtención de las imágenes	34
5.1.2. Procesamiento	35
5.1.3. Mejora propuestas sobre la calidad de las imágenes	36
5.1.3.1. Realce de contraste y brillo	37
5.1.3.2. Binarización y detección de bordes	38
5.1.3.3. Eliminación de ruido	40
5.2. Identificación de estructuras	40
5.2.1. Algoritmo de localización e identificación	40
5.2.1.1. Descripción de una curva sinusoidal	40
5.2.1.2. Identificación de línea base	41
5.2.1.3. Determinación de amplitud y fase	42
5.2.1.4. Implementación del Algoritmo	43
5.2.2. Modificación propuesta sobre los parámetros del algoritmo	44
5.2.2.1. Alto de la ventana de análisis	44
5.2.2.2. Votación en identificación de línea base	45
5.2.2.3. Votación en amplitud y fase	45
5.2.3. Modificación propuesta al algoritmo	46
5.2.3.1. Semejanza por vecindad	47
5.2.3.1.1. Algoritmo de Detección de Líneas Base	48
5.3. Métricas de evaluación	49
5.3.1. Medida de rendimiento de un algoritmo de localización e identificación	49

5.3.1.1.	Línea base	50
5.3.1.1.1.	Métrica	50
5.3.1.1.2.	Precisión	51
5.3.1.2.	Amplitud y fase	51
5.3.1.2.1.	Métrica	52
5.4.	Conjunto de imágenes de evaluación	52
5.4.1.	Construcción de imágenes sintéticas	52
5.4.1.1.	Ruido en las imágenes	53
5.4.2.	Imágenes clasificadas por parte del experto	55
6.	Resultados	57
6.1.	Clasificación sobre Imágenes Sintéticas	57
6.1.1.	Clasificación de Fracturas Sintéticas	57
6.1.1.1.	Espacio de Votación 2-D (amplitud y fase)	61
6.1.1.2.	Comparación con el algoritmo original	62
6.1.2.	Clasificación con Ruido	63
6.1.2.1.	Clasificación Sin Limpieza	63
6.1.2.2.	Clasificación con Limpieza	64
6.1.2.3.	Sensibilidad al Ruido	67
6.1.2.4.	Resultados	67
6.2.	Identificación de Fracturas en Imágenes Reales	70
6.2.1.	Clasificación Sobre Imágenes Específicas	70
6.2.1.1.	Resultados	73
6.2.2.	Identificación Sobre Imágenes Pre-Clasificadas	77
6.3.	Problemas Encontrados	78
6.3.1.	Semejanza por Vecindad	78
6.3.1.1.	Implementación de la ventana móvil	79
6.3.1.2.	Implementación de LBP	79

6.3.2. Sensibilidad de <i>peakdetect</i>	80
7. Conclusiones y Trabajo Futuro	81
7.1. Resultados Obtenidos	81
7.1.1. Limitaciones	81
7.2. Algoritmo	81
7.2.1. Mejora del Algoritmo	82
7.2.2. Limitaciones	83
7.3. Proceso de Limpieza y Binarizado de la Imagen	83
7.3.1. Limitaciones	83
7.4. Utilización de Información Local	83
7.4.1. Semejanza por Vecindad	84
7.4.2. Limitaciones	85
7.5. Métricas	85
7.6. Materiales Utilizados	85
7.6.1. Conjunto de Imágenes	85
7.6.2. Lenguaje de Programación	86
7.6.3. WellCAD	86
Apéndice A. Formato LGX	92

Capítulo 1

Introducción

La geofísica de pozos es un método utilizado para obtener los rasgos estructurales de un sondaje. Consiste en una serie de mediciones obtenidas por un instrumento compuesto de uno o más sensores, el cual es introducido a lo largo del pozo o perforación de sondaje y va registrando información a medida que avanza.

Existen distintas técnicas y dispositivos para registrar la medición, cada uno de ellos depende del tipo de mineral existente en el sondaje y de los parámetros que se deseen identificar, por ejemplo: litología de la zona, daño estructural, orientación de estructuras, resistividad, densidad volumétrica, geometría, porosidad, etc.



Figura 1.1: Empresa realizando la perforación de un sondaje.

Un trabajo importante realizado luego de la captura de datos de un sondaje (ver figura 1.1), es analizar la información tales como la resistencia del roca, variaciones de temperatura, trayectoria del pozo y geometría del mismo, con el fin de identificar y clasificar las distintas estructuras presentes en el pozo. Este proceso es realizado con ayuda de herramientas de *software*, sin embargo, sólo como ayuda para desplegar los datos.

Entre la información que se puede obtener de un pozo, una importante es la identificación de fracturas, las cuales pueden ser de origen natural o inducidas en el proceso de perforación. La importancia radica en que son una buena fuente de estimación de la orientación del *stress* que posee la roca (ver figura 1.2) y por lo tanto es posible interpretar la geofísica del pozo.



Figura 1.2: *Fracturas inducidas por stress (El Salvador, foto: Chuck DeMets).*

Una técnica importante utilizada para obtener información sobre las fracturas presentes en el pozo, es la utilización de técnicas de televisión acústica[27]. Al utilizarla, se obtiene una imagen de las paredes del pozo que describen sus características. La utilización de este tipo de tecnología, genera grandes volúmenes de datos y son una fuente rica en información que debe ser procesada y cuantificada como parte del proceso de interpretación.

1.1. Motivación de la Tesis

La interpretación de las imágenes del pozo, comienzan con la detección y clasificación de las fracturas presentes, un proceso manual que consume bastante tiempo y su calidad se encuentra sujeta a variaciones de acuerdo a la experiencia del profesional. Esta tarea es muy repetitiva debido a que en una imagen es común encontrar cientos y a veces miles de estas fracturas. La posibilidad de automatizar

este proceso es evidente y tiene el potencial de entregar los siguientes beneficios:

1. Identificación y parametrización automática de estructuras presentes en una imagen.
2. Mejora en la velocidad de análisis de los datos de los pozos, debido a la pre-identificación de estructuras.

1.2. Objetivos

A raíz de la importancia de la interpretación de imágenes de un pozo de sondaje, se propone el siguiente objetivo general y específicos.

1.2.1. Objetivo General

Detectar automáticamente fracturas de rocas presentes en imágenes de pozos de sondajes, con una certeza igual o superior al experto humano.

1.2.2. Objetivos Específicos

Los objetivos específicos son los siguientes:

1. Capturar el procedimiento que utiliza el equipo de expertos para clasificar fracturas, formalizando la secuencia de pasos y actividades necesarias para caracterizar las curvas de identificación de fracturas.
2. Utilizar un filtro de imagen aplicados a la información gráfica de un sondaje, con el fin de mejorar el contraste entre las distintas estructuras geológicas.
3. Identificar automáticamente las fracturas presentes en un sondaje por medio de una combinación de técnicas computacionales de visión por computador y reconocimiento de patrones.
4. Proponer una métrica que permita realizar una medida de las mejoras introducidas al algoritmo de identificación y parametrización.

1.3. Contribución y visión general de la Tesis

Este trabajo presenta las mejoras a un proceso automático para identificar fracturas de origen natural por medio de técnicas de procesamiento de imágenes. El proceso incluye una fase de obtención de las imágenes a partir de los archivos generados por el captador de datos, identificación de líneas base de las fracturas y en forma posterior parametrización de las mismas.

1.3.1. Contribución

La investigación realizada posee el valor de proponer mejoras al algoritmo de clasificación de fracturas propuesto por Changchun [7], proponer un proceso de detección y parametrización de fracturas, identificar las variables que influyen en la precisión del algoritmo y proponer métricas para medir el rendimiento de las mejoras al algoritmo.

1.3.2. Visión General

El capítulo II presenta una revisión del concepto, herramientas y descripción de la geofísica de pozos. El capítulo III describe el problema de búsqueda y parametrización de fracturas. El capítulo IV presenta el estado del arte. El capítulo V se enfoca en la metodología propuesta. El capítulo VI presenta los resultados obtenidos. Finalmente, se presentan las conclusiones del trabajo realizado, se identifican y discuten las limitaciones y trabajo futuro.

Capítulo 2

Geofísica de Pozos

Los registros de geofísica de pozos, consisten en una serie de mediciones obtenidas por una sonda con varios sensores o antenas transmisoras y receptoras que se introducen en una perforación de sondaje para determinar las características de cada zona que se desee conocer [2].

Las técnicas a utilizar dependerán de las profundidades, de los tipos de minerales involucrados o los parámetros que se deseen identificar. Con estas técnicas es posible determinar los diferentes tipos de rocas, daños estructurales, orientación de las estructuras, resistividad, densidad volumétrica, geometría, porosidad, permeabilidad y otros parámetros.

Al insertar y realizar las mediciones repetidamente sobre el mismo pozo se pueden causar perturbaciones en las paredes del mismo, es por ello que dependiendo de las características del suelo a analizar, se debe seleccionar el más adecuado, esto con el fin de evitar el posible colapso del sondaje, dejando atrapado el instrumento de medición [26].

Debido a que las condiciones del pozo afectan directamente la calidad y tipo de sensor a utilizar, es que se realizan varias consideraciones prácticas, por ejemplo el diámetro del pozo, si se encuentra inundado con agua/lodo, etc. La técnica que mejor se aplica al tipo de pozos estudiados, es la de televisión. Esto debido a que permite obtener una imagen sin perturbaciones [13].

2.1. Televisión Acústica

La técnica del televisor acústico consiste en una o más imágenes construidas a partir de los datos extraídos del sondaje. Éstas son obtenidas por medio de un conjunto de mediciones realizadas a lo largo del pozo, con la ayuda de un instrumento que recorre la paredes del mismo, generando vibraciones ultrasónicas [34].



Figura 2.1: *Probeta de televisión acústica, transmite y registra vibraciones ultra sónicas.*

La probeta contiene un transductor cerámico piezo-eléctrico de alta frecuencia (ver figura 2.1), el cual transmite vibraciones ultrasónicas que son direccionadas por medio de un espejo acústico rotatorio en 360 grados hacia las paredes del pozo. La señal acústica es reflejada de vuelta al transductor y es enviada por medio de un cable de comunicaciones, a un dispositivo ubicado en la boca el pozo que registra y almacena la amplitud y tiempo de tránsito para cada punto. Debido al tipo de frecuencias utilizadas, se requiere de un fluido al interior del pozo para

transmitir la onda acústica. Los fluidos pueden ser ópticamente opacos (aguas turbias o lodos de bentonita utilizado en el proceso de perforación).

El muestreo es realizado muchas veces horizontalmente y a alta velocidad (20 rpm), con ello se obtiene una matriz densa de datos. Los puntos de muestreo son definidos por medio de la división de las paredes del pozo en unos intervalos regulares, vertical y horizontalmente, con ello se obtiene una matriz de pixels. Cada pixel es un punto de muestreo que representa la amplitud y tiempo de viaje. Dependiendo del diámetro de pozo, el tamaño del pixel puede variar. La resolución en la vertical es menor que dos milímetros y la resolución horizontal depende del número de muestreos realizados en cada grado [12].

La imagen de las paredes del pozo, es obtenida a partir de la información registrada de manera posterior a la extracción de los datos. La imagen que representa los valores extraídos no es la medida exacta, si no que una representación visual del tiempo de tránsito y amplitud de la onda acústica. La imagen es creada definiendo un color particular para cierto rango de valores. La selección de colores y filtrado de la imagen pueden ser modificados de manera posterior en el *software* de procesamiento de datos (ver figura 2.2)

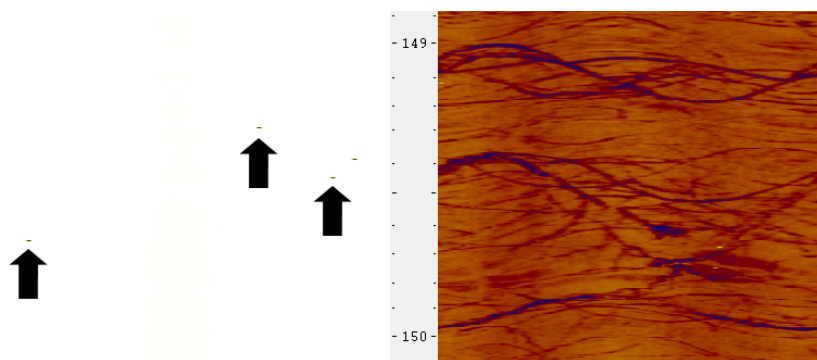


Figura 2.2: Representación visual de los datos de tiempo de tránsito (izq.) y amplitud (der.) de la señal acústica. Se ha indicado con una flecha los lugares donde el tiempo de tránsito posee alguna anomalía.

La orientación de la imagen es realizada por medio de tres acelerómetros y un magnetómetro de tres ejes, es por ello, que la orientación con respecto al norte magnético puede ser obtenida con precisión y ser mantenida en la fase de análisis de las estructuras. Las imágenes pueden ser obtenidas en pozos perforados con diamantina o aire reverso, aunque las paredes más suaves de un sondaje perforado con diamantina producirán una imagen con mejor detalle y nitidez.

La matriz de datos obtenida se compone de dos elementos: un valor que representa el tiempo de tránsito de la onda y otro que representa la amplitud de la misma.

2.1.1. Tiempo de Tránsito

El tiempo de tránsito es el tiempo requerido por la onda para la siguiente secuencia de pasos: emisión, reflexión y retorno al transductor. Este parámetro describe la forma y diámetro del pozo. Cambios en las estructuras pueden ser fácilmente reconocibles por tiempos de tránsito altos, además se pueden identificar fracturas abiertas debido a que este tipo no registra regreso de la señal. En la izquierda de la figura 2.3 se puede identificar una fractura a partir de los colores más oscuros.

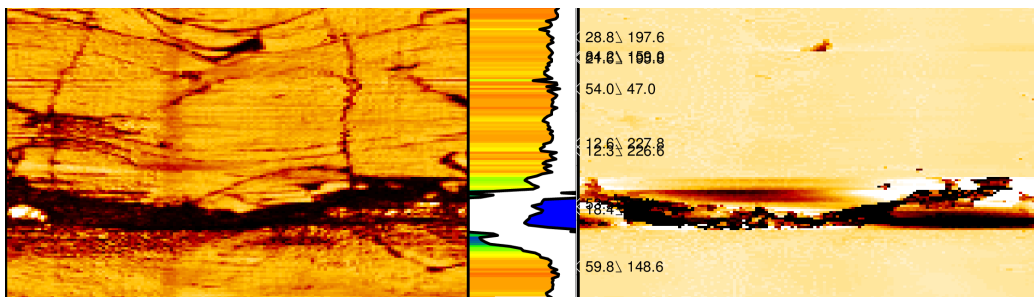


Figura 2.3: Amplitud de la señal (izq.), gráfico de valores de fracturas identificadas (amplitud y dip) y tiempo de tránsito (der.): Se identifica una estructura abierta debido al no retorno de señal (zona oscura en el tiempo de tránsito).

2.1.2. Amplitud

La amplitud de la señal representa la reflexión de la respuesta acumulada de la señal sónica sobre un tiempo predefinido. Las fracturas, visualizadas como áreas oscuras, es la estructura que típicamente es identificada por esta técnica [34].

La reflexión (y por lo tanto la amplitud de la onda registrada) varía con la diferencia de impedancia entre el fluido del pozo y la pared del mismo. Debido a que la impedancia del fluido es constante a través del plano horizontal muestreado, ésta puede ser expresada como el cambio relativo en la litología y propiedades físicas de la pared del pozo.

Formaciones duras dan origen a reflexiones con gran amplitud, debido a que gran parte de la señal es reflejada, es decir, la formación es de alta densidad. Debido a ello se puede asociar la amplitud de la señal acústica a la dureza de la roca.

2.1.3. Factores que afectan la adquisición de los datos e interpretación

Durante el proceso de interpretación de los datos es importante recordar la influencia de la calidad de los mismos, la cual se puede ver afectada por distorsiones o artefactos de origen no geológico:

- El televisor acústico genera grandes cantidades de datos asociado a la alta velocidad de muestreo, por lo tanto es necesario optimizar la resolución para obtener datos de calidad. La resolución depende del diámetro del pozo, cantidad de lodo y/o cantidad de agua contaminada y por lo tanto, es necesario regular la velocidad de muestreo para las características particulares de cada pozo. Dependiendo si se utilizan transductores de alta o baja frecuencia se puede determinar la resolución: sistemas con transductores de

baja frecuencia reduce la resolución y diámetro necesario [34].

- El análisis por ultrasonido sólo puede ser realizado si existe fluido en el pozo, sin embargo, el lodo causa atenuación de la señal, debido a que el pulso de energía es absorbido y disperso [26].
- Un punto importante en esta técnica es la centralización de la probeta, es decir, que se encuentre equidistante el transductor con respecto a cualquier punto del plano horizontal del pozo (ver figura 2.4).
- Las características geológicas pueden ser sobre impresas con fracturas inducidas por el proceso de perforación.
- Distorsión asociada a una corrección imperfecta introducida durante la construcción de la imagen por parte del analista.

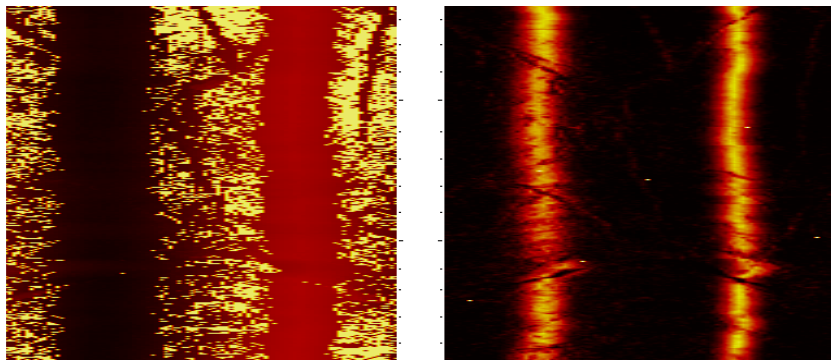


Figura 2.4: *Amplitud de la señal (izq.) vs tiempo de tránsito (der.) obtenidos a partir de una probeta no centrada en el eje del pozo.*

2.2. Andes Geofísica

Andes Geofísica brinda el servicio de orientación de estructuras, en el cual se utilizan elementos y dispositivos de alta tecnología y precisión, principalmente

televisores ópticos y acústicos, en conjunto con un equipo de personas altamente especializado que realizan las siguientes operaciones:

- **Toma de Muestras:** Realizado por un equipo que concurre al lugar del sondaje, cuenta con un vehículo que posee la infraestructura necesaria para la utilización de televisores ópticos y acústicos, los cuales por medio de un sistema de captura de datos (ver figura 2.5), registra la información del sondaje en un computador personal.
- **Análisis de Datos:** Realizado por un equipo de geofísicos y geólogos, los cuales son encargados de la limpieza de datos y de la identificación de las estructuras del sondaje.



Figura 2.5: *Computador de terreno utilizado para el registro de información del sondaje[28].*

El equipo de análisis de datos es el encargado de realizar la identificación de las estructuras a partir de los datos registrados por el equipo de toma de muestras. Este proceso es realizado con la ayuda del software WellCAD[31], él cual es capaz

de desplegar los datos del sondaje de manera gráfica, sin embargo, el proceso es realizado de forma manual, metro a metro a lo largo del sondaje, el cual tiene como promedio entre 200 y 300 metros de longitud.

Del trabajo que realiza el equipo de análisis se pueden desprender los siguientes problemas:

1. **Demora en la recepción de los datos** Debido a las distancias entre el lugar donde se realiza la captura de los datos y el lugar donde se realiza el análisis de estructuras, sumado al tamaño de la información, provoca que la recepción de los datos no sea realizada inmediatamente y transcurran varios días.
2. **Calidad de los datos** Debido a factores operacionales (por ejemplo: polvo en suspensión, condiciones del ambiente, instrumento utilizado, experiencia del equipo de muestreo, etc.) es común que la información contenga ruido, el cual dificulta la tarea de análisis y requiere de un proceso de limpieza manual para mejorar las condiciones de los mismos (ver figura 2.6).
3. **Identificación y clasificación de las estructuras** A raíz que el proceso de análisis de datos es realizado de forma manual, el proceso es muy lento, lo que alarga los tiempos de entrega del informe de análisis. Este trabajo es desarrollado por Geofísicos. La precisión de la identificación y calidad de la clasificación se encuentra directamente relacionada con el *know-how* del experto, es por ello que es importante lograr capturar este conocimiento.
4. **Validación de las estructuras** Una vez que se han identificado las estructuras, deben pasar por un proceso de validación por parte de un geólogo, el cual debido a los retrasos producidos en la entrega inicial y posteriormente en la identificación, provoca tiempos muertos y sobrecarga en distintos tiempos.

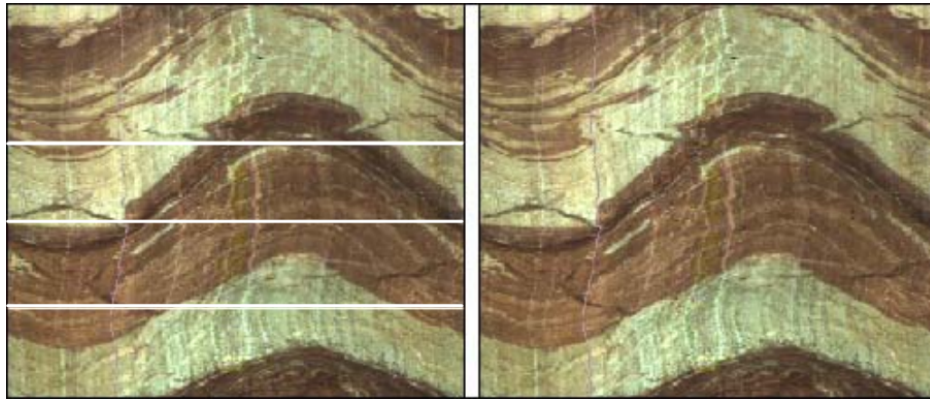


Figura 2.6: *Problema común encontrado en los datos (izquierda) y el resultado luego de la corrección manual (derecha)[32].*

Capítulo 3

Problema de Investigación

3.1. Detección y Parametrización de Estructuras

La visualización de las paredes de un pozo por medio de imágenes, provee mucho más detalle sobre otras técnicas. A partir de ellas es posible extraer información del sondaje, identificar e interpretar las estructuras encontradas. Es importante reconocer que la calidad de los análisis realizados aumenta al utilizar más de una técnica en el mismo pozo.

3.1.1. *Breakouts* (Deformaciones)

Los *Breakout*¹ son deformaciones del pozo producidas por las fuerzas involucradas en la perforación. De manera ideal un pozo posee una sección circular, sin embargo debido a la falla de la roca de las paredes del sondaje, se puede deformar hacia un óvalo. Son importantes de identificar y evaluar debido a que se asocia a las propiedades de la roca y por lo tanto, permite dar soporte a las interpretaciones de dureza, cantidad de *stress* existente y asociar a las estructuras existentes. De manera simplificada se puede decir que a menor dureza de la roca, mayor puede ser el *breakout* debido a la erosión inducida por el método de perforación utilizado, más aún, se puede identificar la dirección principal de la fractura.

Los *breakouts* son identificados por el tiempo de tránsito y por lo tanto se utiliza principalmente la técnica de televisión acústica. El registro de amplitud puede

¹Ensanchado de las paredes del pozo inducido por stress.

ser usado como referencia para identificar capas de rocas suaves susceptibles de producir largos *breakouts*.



Figura 3.1: Imagen 3-D obtenida a partir de los datos del registro de tiempo de tránsito de la onda desde el emisor hasta la pared del pozo[13]. Los *breakouts* existentes se identifican en la imagen como las zonas externas a las paredes del pozo.

Los *breakouts* son visualizados por colores oscuros en el registro de tiempo de tránsito, donde el tamaño de las zonas oscuras se asocia al tamaño del *breakout*. A partir de esta información se puede generar una imagen 3-D (ver figura 3.1), la cual revela una limitante de los registros acústicos: existe una distancia máxima de alcance de la señal acústica (en la 3.1 se visualiza como un corte abrupto del *breakout*). A pesar de esta limitación la imagen puede ser utilizada para evaluar este tipo de estructura y determinar sus características.

Los *breakouts* y paredes del pozo rugosas son en su mayoría creados por el desprendimiento del material en el instante de la perforación. Al desprenderse y ser transportando hacia la superficie, el material duro transportado a lo largo de las paredes del pozo causa erosión.

3.1.2. Discontinuidades (Fracturas)

Es importante lograr identificar las discontinuidades, esto con el fin de construir el modelo de la masa rocosa más apropiado. Al momento de construir el modelo, es importante lograr distinguir si las discontinuidades observadas en la imagen son naturales o artificialmente producidas por la técnica de perforación utilizada.

Las discontinuidades o fracturas pueden ser identificadas en cualquiera de los dos registros de imagen: acústico-amplitud y acústico-tiempo de tránsito. La posibilidad de registrar discontinuidades en el registro acústico nace de las variaciones que se pueden encontrar en la falla. La determinación de la orientación de la discontinuidad puede proveer información sobre el sistema total de discontinuidades.

3.1.2.1. Fracturas de procedencia natural

Las fracturas son fácilmente reconocibles en el registro acústico debido a que son representadas por líneas delgadas de color oscuro. Las fracturas verticales (cortan el eje del pozo en algún ángulo) son más fáciles de reconocer que las horizontales ya que no es posible confundirlas con otras características, por ejemplo, con una capa delgada de material de baja dureza.

Cuando la imagen es abierta y es desplegada en forma de plano, una fractura que corta en un ángulo el eje del pozo se transforma en una **curva sinusoidal**.

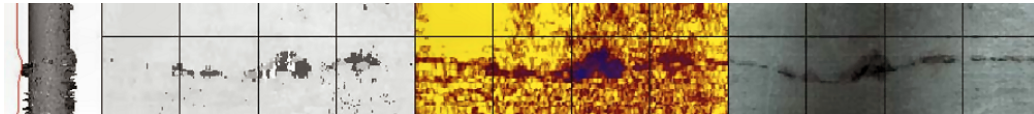


Figura 3.2: Fractura vertical que puede ser identificada en los tres registros: ópticos, acústicos (amplitud y tiempo de tránsito)[30]. La distancia entre el punto más alto y bajo de la fractura es aproximadamente 3 cm.

En la figura 3.2 se puede reconocer una senoide, la amplitud de la fractura es de 3 centímetros y utilizando la ecuación 3.1 se puede determinar que el ángulo con respecto a la horizontal es de 12 grados.

$$\text{ÁnguloFractura}[\text{°}] = \frac{\text{amplitud}}{\text{diámetro}} \quad (3.1)$$

A pesar que en el registro acústico es fácil identificar las ondas sinusoidales, puede ser complejo distinguir entre una fractura y una capa que separa distintos tipos de material. En estos casos las dos imágenes acústicas pueden colaborar en la interpretación, además, si se agrega una imagen 3-D construida a partir de los datos acústicos de amplitud, esta nueva dimensión permite una mejor vista de las fracturas y su caracterización.

3.1.2.2. Fracturas inducidas por perforación

La mayoría de las fracturas inducidas por la técnica de perforación son horizontales, debido a ello en los Testigos geológicos² se puede evidenciar una típica distribución en forma de discos. En el caso del registro de imagen de televisión acústica, se identifican solamente por una línea horizontal, debido a ello, son descartadas.

3.1.2.3. Fracturas abiertas o cerradas

A partir de los registros obtenidos, se puede identificar si una fractura es abierta o cerrada, este elemento es importante con el fin de evaluar el estado de *stress* de la roca. Las fracturas cerradas son indicativas que la roca se encuentra sujeta a fuerzas de compresión. Las fracturas abiertas son producidas por fuerzas de extensión y por lo tanto si éstas son encontradas, es crítico debido a que podría ser índice de un movimiento de deslizamiento a lo largo de la falla.

²Cilindro de material del pozo obtenido por medio de la perforación mediante diamantina.

Las fracturas abiertas puede ser identificadas en imágenes de televisión acústica [26] examinado los registros de amplitud y de tiempo de tránsito: Cuando la señal ultrasónica generada por el instrumento alcanza una fractura, ésta es refractada en la roca y sólo una pequeña cantidad es reflejada de vuelta. Si la fractura es abierta, mayor cantidad de señal es refractada en la roca, por lo tanto, una fractura abierta puede ser visualizada claramente en el registro de amplitud de señal y al mismo tiempo en el registro de tiempo de tránsito (ver figura 3.3).

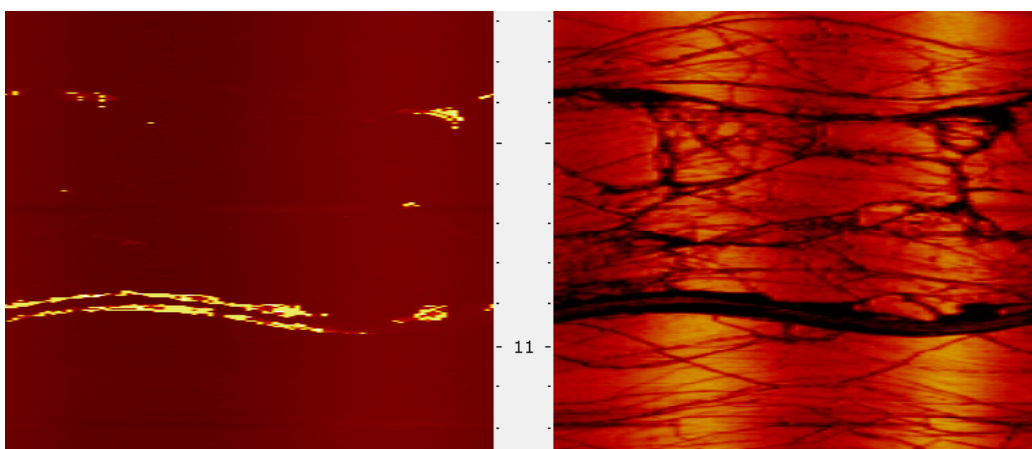


Figura 3.3: *Fractura abierta identificada por medio del registro de amplitud de señal reflejada (izq.). Se pueden identificar varias fracturas cerradas en el registro de tiempo de tránsito (der.) que no aparecen en el registro de amplitud (izq.).*

Las fracturas cerradas sólo pueden ser identificadas examinando el registro de tiempo de tránsito, ésto debido a que la misma cantidad de señal es reflejada, sin embargo, no existe una diferencia apreciable en los tiempos de tránsito entre la fractura y la roca de la pared del pozo.

En las imágenes de televisión acústica son identificables por líneas que aparecen en el registro de tiempo de tránsito, pero no en el registro de amplitud (ver figura 3.3).

3.2. Identificación de Fracturas

Las fracturas a identificar son caracterizadas por medio del siguiente conjunto de datos: profundidad, **dip**³ y **azimut**⁴ (ver figura 3.4).

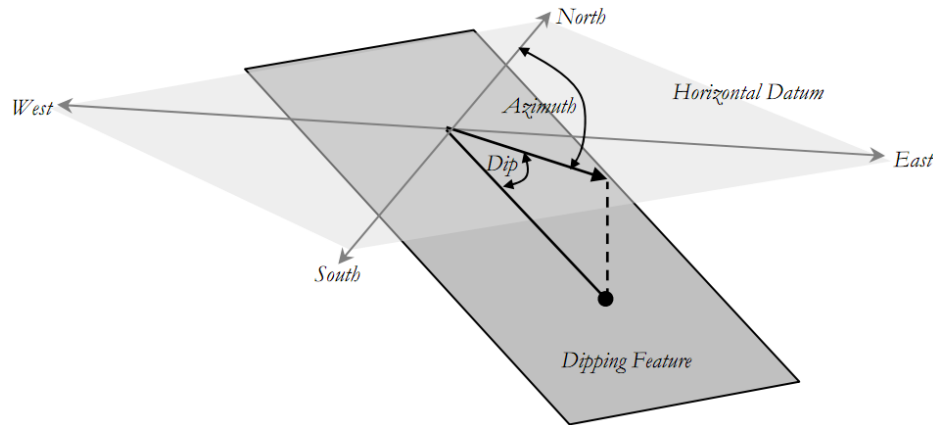


Figura 3.4: Relación entre el ángulo de azimut y el ángulo de dip [32].

Las fracturas o planos que atraviesan el sondaje son representados por medio de una onda sinusoidal. La razón de ello es que si se considera un plano no perpendicular al eje central del pozo, la intersección de las paredes del pozo con el plano dibuja una elipse. Al momento de abrir el plano 2-D donde se encuentra ésta, se genera una onda sinusoidal (figura 3.5).

En la onda generada en el plano 2-D que representa a una fractura, es importante identificar:

- **Dip:** identificado por la amplitud de la onda.
- **Azimut:** identificado por la posición donde se encuentra el mínimo de la onda.

³Inclinación de un plano con respecto a una referencia vertical.

⁴Ángulo con respecto al norte.

- **Largo:** identificado por el espesor de la onda.
- **Profundidad:** identificado por la localización de la onda.

Por lo tanto, una fractura puede ser identificada por la suma de los siguientes parámetros: profundidad, **azimut**, **dip**, largo, categoría y descripción. Éstos dos últimos son asignados por el experto, indicando la categoría asociada al tipo de característica observada y asignando una descripción de manera optativa.

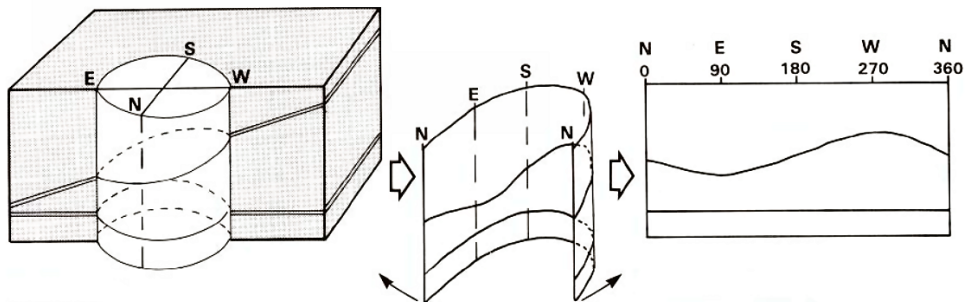


Figura 3.5: Apertura de la imagen obtenida de un pozo representada por un plano. Se puede verificar que el ángulo del plano con respecto al horizontal será representado por la amplitud de la sinusoide[13].

3.2.1. Proceso de identificación manual de fracturas

Una vez que los datos son recogidos por el equipo de análisis, éstos son importados al software WellCAD[31], donde la información recogida es desplegada de manera gráfica, mostrando secciones transversales del sondaje (ver figura 3.6). Si la calidad de los datos no es suficiente, se informa al equipo de toma de muestras, el cual debe realizar un segundo muestreo del sondaje.

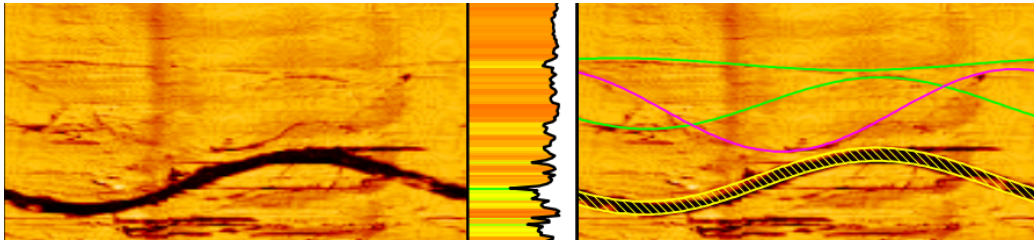


Figura 3.6: *Identificación de estructuras de manera gráfica con ayuda de WellCAD. Se puede apreciar los distintos tipos de estructuras identificadas.*

El proceso de identificación de estructuras es realizado de manera manual con ayuda del *mouse*: el analista selecciona distintos puntos de la posible onda y realiza ajustes de amplitud/desfase (ver figura 3.7)

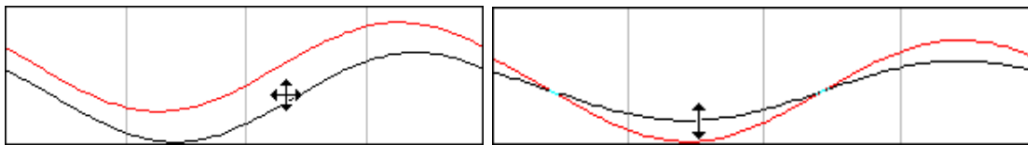


Figura 3.7: *Corrección de la estructura identificada. En el caso de la izquierda se requiere corrección de desfase y en el de la derecha corregir amplitud.*

Una vez identificada la onda, realiza el proceso de clasificación, es decir, asignar la categoría y descripción de una estructura. En caso de existir alguna duda sobre la estructura, se puede indicar dejando una nota.

Cada una de las ondas identificadas, representan una estructura que se encuentra orientada, es decir, posee una profundidad, azimuth y dip (ver figura 3.4). A partir de esta información, la clasificación y debido a que los datos se encuentran geo-referenciados, WellCAD despliega un gráfico de orientación de estructuras (figura 3.8).

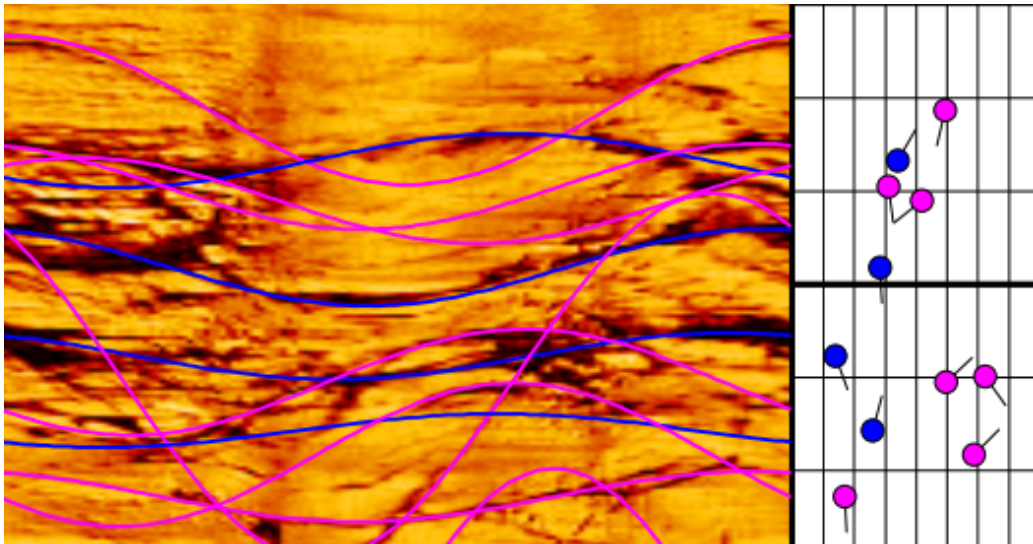


Figura 3.8: *Distintos tipos de estructuras identificadas. Se indica en el gráfico de la derecha el tipo y azimut de la estructura.*

Capítulo 4

Estado del Arte

El proceso de identificación de estructuras ha sido abordado por varios autores y por medio de distintas técnicas. La tabla 4.1 expone un resumen de la bibliografía consultada.

Tabla 4.1: *Resumen del estado del arte consultado.*

Autores	Año	Método	Resultados
Luthi, Souhai [21]	1990	Técnicas simples de estadística sobre medidas de micro-conductividad.	Identificación de fracturas de hasta 1 cm. de separación.
Changchun [7]	2002	Identificación y parametrización de curvas por medio de la transformada de Hough	Algoritmo aplicado sobre curvas sintéticas con y sin ruido logrando alta precisión y robustez al ruido.
Raghupathy [25]	2004	Identificación de fracturas por medio de la transformada de Radon.	Algoritmo propuesto aplicado a imágenes con ruido, logrando alta precisión y robustez al ruido. Requiere parametrización manual.
Jarosinski [17]	2005	Descripción de fracturas sobre medidas de caliper.	Caracterización de una fractura por medio del análisis de los valores de espesor del pozo.
Hubbard, Robertson [16, 15, 27]	2010	Caracterización de fracturas en pozos en glaciares por medio del filtro Canny de detección de bordes y luego ajuste de curva por el método de los mínimos cuadrados.	Caracterización de fracturas en las paredes del pozo de un glacial.
Leta, Clua, Biondi, Pacheco, Socorro [18]	2013	Sistema semi-automático de caracterización de fracturas por medio de la transformada de Hough y un filtro de medianas.	Clasificación semi-automática, pero con baja precisión en caso de curvas sobrepuestas.
Assous, Elkington, Clark, Whetton [1]	2014	Detección de bordes por medio de Lindeberg, método de validación por medio de congruencia de fase y luego aplicación de transformada de Hough.	Menos de 5% de falsos positivos en imágenes con ruido y menos del 2% en caso de imágenes de buena calidad.

4.1. Discusión de Trabajos

Luthi y Souhai [21] proponen un método de tres pasos para detectar, trazar y cuantificar las fracturas presentes en datos obtenidos por medio de un micro-scanner de alta resolución que registra la conductividad de la roca. Las fracturas son identificadas por medio de técnicas simples de estadística, donde se considera las variaciones naturales de conductividad de la roca: Se observa si un valor es estadísticamente superior la matriz de conductividad de una zona, si es así, ese valor puede indicar la presencia de una fractura. La técnica propuesta para caracterizar fracturas logra identificar hasta aquellas de 1 cm. de separación, además incluye información del azimuth de la misma.

Una fractura puede ser representada por una curva sinusoidal de período fijo [11], a partir de ello Changchun [7], propone un algoritmo para identificar curvas sinusoidales de período fijo y por lo tanto fracturas. El algoritmo se compone de dos pasos: inicialmente identificar la línea base de la curva por medio de un proceso de votación en una dimensión; y en forma posterior, para cada línea base, identificar la amplitud y fase original de la curva por medio de un proceso de votación en dos dimensiones.

Para el segundo paso, se utiliza una variante de la transformada estándar de Hough (SHT), esto porque la aplicación directa de la transformada de Hough involucra un alto costo computacional.

El método propuesto por Changchun [7] descompone el problema de 3-D en dos: uno de 1-D y el otro de 2-D. Con esto realiza una disminución importante en requerimientos de espacio y cómputo. El trabajo experimental con el algoritmo propuesto es aplicado sobre imágenes sintéticas, inicialmente sin ruido y en forma posterior incluyendo ruido con densidad igual a 0,02. Obteniendo como conclusión que el método parametriza las curvas de forma muy precisa y posee robustez al ruido.

Raghupathy [25] presenta en su trabajo de tesis la aplicación de la transformada de Hough y su versión más general: la transformada General de Radon (GRT) [33, 9] para la detección de líneas en imágenes con ruido. En su trabajo aborda tres problemas que sufre la utilización de GRT: detección de múltiples curvas debido al espesor de la misma, identificación falsos-positivos debido a utilización de información global de la curva y no local; y por último, cuando el modelo propuesto es insuficiente para extraer las curvas de las imágenes. Los problemas de las mejoras que propone, es que requiere de selección manual de parámetros, incluyendo: espesor de la fractura (para usar en la transformada de Radon), selección del tamaño de la ventana y forma de la misma, además de algunas condiciones de borde para eliminar falsos positivos.

Jarosinski [17] propone el *software* SPIDER para procesar y caracterizar las propiedades de una fractura, siendo necesario ingresar una serie de parámetros de filtrado: coeficiente de forma, coeficiente de asimetría, el azimut y la desviación con respecto a la horizontal. El trabajo principal es describir las características de la fractura, sin embargo, esto lo realiza una vez identificada la posición y parámetros de la misma. El programa fue diseñado para procesar información de pozos obtenidas mediante la técnica del Caliper, a raíz de ello al aplicarlo sobre imágenes de televisión acústica, remueve más de la mitad del largo de las fracturas.

Hubbard y Roberson [16, 15, 27] en sus investigaciones sobre la aplicación de técnicas de televisión óptica sobre pozos en glaciares, realizan las anotaciones sobre las características de los mismos, con la ayuda del *software* BIFAT [22] (*Borehole and Ice Feature Annotation Tool*), el cual posee un módulo de detección de capas de manera automática.

El método utilizado por BIFAT se basa en el utilización del filtro Canny detector de bordes [5, 4] en la imagen del pozo. El procedimiento planteado primero realiza la convolución de la imagen con una función de Gauss con el fin de

remover ruido. Debido a que los bordes buscados pueden apuntar en cualquier dirección, se utilizan uno o más operadores direccionales para calcular la intensidad del gradiente de cada pixel en la convolución de la imagen. Luego, se realiza un proceso de descarte de bordes, lo que involucra la comparación del valor del borde en la dirección del pixel y seleccionando sólo los con valores mayores. Finalmente se realiza un proceso de histéresis, removiendo cualquier valor que sea inferior o superior a un valor umbral fijado. Una vez que los bordes han sido identificados, se procede a realizar un proceso de ajuste por medio de los mínimos cuadrados, de una onda sinusoidal sobre los bordes encontrados.

Leta, Clua, Biondi, Pacheco y Socorro [18] desarrollaron un sistema semi-automático de detección de fracturas por medio de las técnicas de procesamiento de imágenes. Inicialmente plantean mejorar la calidad de las imágenes del pozo por medio de la aplicación de algún filtro para mejorar la búsqueda de bordes: Sobel, Prewitt, Roberts, Laplacian, Zero-Crossing y Canny. A pesar de ello, seleccionan usar un filtro simple de medianas de 9×9 para descartar el ruido. En esta fase, la imagen original es transformada desde sus valores originales a escala de grises de 256 niveles y luego reducida a 32 niveles de gris, ésto con el fin de facilitar y aumentar la velocidad de proceso. Por último, la imagen es binarizada, considerando el histograma de la imagen. En particular, se selecciona el umbral donde el 3 % de los pixel representan un color. El paso siguiente es aplicar la transformada de Hough de tal manera que todos los puntos que pertenecen a la misma curva, son mapeados a un sólo punto en el espacio creado por los parámetros buscados. Los resultados obtenidos indican que la técnica es viable, sin embargo, es necesario algunas mejoras con el fin de identificar fracturas intercaladas.

Assous, Elkington, Clark y Whetton [1] describen un método para la detección de discontinuidades prominentes y el cálculo de sus ángulos de dip. Éste combina un proceso de detección de bordes basado en el algoritmo de Lindeberg [19], con un método de validación basado en congruencia de fase, seguido de un método

robusto de detección de sinusoides basado en el propuesto por Changchun [7], descomponiendo el problema en identificar las líneas base de las posibles ondas sinusoidales y en forma posterior, parametrizar las mismas.

El algoritmo se basa en el supuesto de la relación simétrica que poseen dos puntos pertenecientes a una curva sinusoidal distanciados $w/2$ entre sí, y en general no asociados a otras características de la imagen. Ésto sumado al hecho que, dado dos puntos simétricos pertenecientes a la curva, sus pendientes deberían ser igual en magnitud, pero con sentido opuesto. Con esta información, en cada punto se identifican los puntos de la vecindad conectados y luego a partir de ellos, mediante el método de los mínimos cuadrados, identificar la pendiente de ese punto. Este proceso es realizado con el fin de disminuir los falsos positivos en la identificación de líneas base. El método propuesto es aplicado sobre un conjunto de imágenes de pozos, obtenidos por medio de la técnica de micro-resistividad. Las métricas utilizadas para evaluar el desempeño del método propuesto se basa en el porcentaje de falsos positivos encontrados, obteniendo una tasa de falsos positivos mejor que 2 % en imágenes de buena calidad y mejor que 5 % en el caso general.

Capítulo 5

Metodología

Se propone utilizar el algoritmo planteado por Changchun [7] para identificación y parametrización de ondas sinusoidales de período fijo, en imágenes reales de fracturas cerradas, obtenidas por medio de técnicas de televisión acústica.

Para el pre-proceso de las imágenes, se analiza el procedimiento propuesto por Leta [18], con el fin de limpiar las imágenes del ruido existente.

Para evaluar la calidad del algoritmo, se propone aplicarlo sobre imágenes de fracturas generadas sintéticamente y con valores pre-establecidos. A continuación, agregar ruido de fondo y un gran número de fracturas sobrelapándose, ésto para simular las condiciones reales.

Por último, una vez identificada la confianza del algoritmo, aplicarlos sobre imágenes de pozos de perforación reales, los cuales ya fueron procesados para identificar las curvas por parte de los expertos. Ésto con el fin de obtener una comparativa de rendimiento del algoritmo versus el experto. Las imágenes y clasificación de fracturas fueron obtenidos de Andes Geofísica.

En la figura 5.1 se puede ver un resumen del proceso llevado a cabo para identificar y clasificar fracturas.

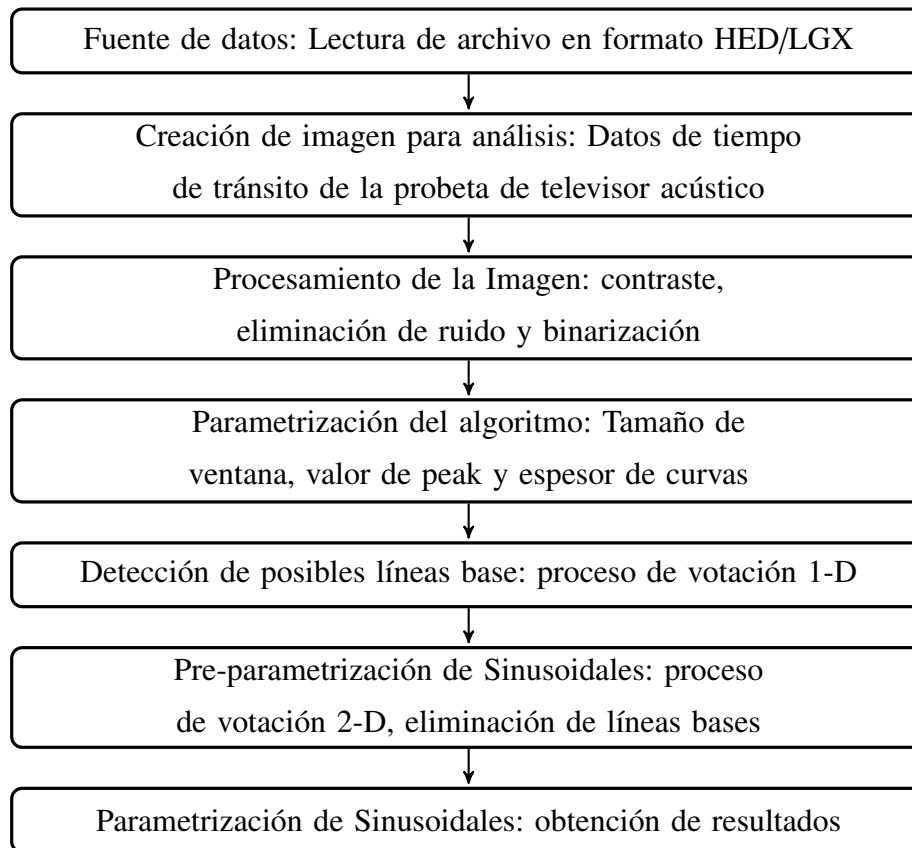


Figura 5.1: Resumen del proceso de detección y parametrización de fracturas.

5.1. Procesamiento de las imágenes

Con el propósito de identificar las fracturas presentes en el pozo, es necesario obtener la imagen a partir de los datos de la probeta, para ello los datos deben ser procesados para obtener la imagen y luego realizar una etapa de limpieza y binarizado para la posterior aplicación del algoritmo de clasificación. Es importante identificar que un factor importante en el éxito de la clasificación es el proceso de limpieza y binarizado, el cual debe ser realizado examinando la calidad de la imagen original, particularmente debido a las distintas características de los pozos.

5.1.1. Obtención de las imágenes

Los datos que representan las imágenes de las paredes del pozo, son obtenidos por medio de un *software* asociado a la probeta utilizada para muestrear. En el caso de los datos de Andes Geofísica, se utiliza la probeta *Acoustic Borehole Televiwer* fabricada por *Electromind* [12]. Los datos obtenidos de esta probeta se encuentran en el formato **LGX** propietario de *Robertson Geologging* [14].

Para la lectura de estos archivos y obtener las imágenes, es necesario comprender el formato interno, el cual se encuentra descrito en el apéndice [A](#).

Los datos del pozo que la probeta captura son almacenados en dos archivos: HED y LGX. El primero contiene la descripción y algunos parámetros del pozo: fecha, profundidad, velocidad de muestreo, resolución, velocidad de los datos, etc. El segundo es un archivo binario separado en canales, donde se encuentran los datos de cada punto muestreado: profundidad, tiempo, ángulo, ganancia de señal, resolución, tamaño del paso, ancho del pulso de comunicación, magnetómetro (x, y, z), aceleración (x, y, z), temperatura del sensor y los datos acústicos.

El canal de datos acústicos es el importante, ya que es el que incluye la información de amplitud y tiempo de tránsito de la señal.

Con la información del formato LGX, es posible leer los archivos binarios y construir una representación matricial de la información de amplitud y tiempo de tránsito. Para la identificación de los parámetros que representan la curva, es común sólo utilizar los datos de amplitud reflejada en las paredes del pozo, esto debido a que es el conjunto de datos que más información entrega.

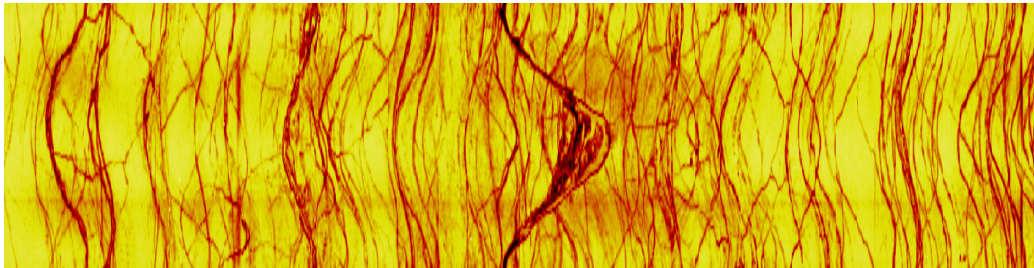


Figura 5.2: Trozo de aproximadamente 4 mt. del canal acústico (amplitud) leído desde el archivo LGX (la imagen se encuentra rotada 90°). Se ha coloreado artificialmente a partir de los valores normalizados entre 0 y 255 utilizando una paleta de colores que representa niveles de calor.

En la figura 5.2 se puede apreciar un trozo de los datos de amplitud obtenidos de un pozo de la segunda región.

5.1.2. Procesamiento

Leta[18] propone una serie de pasos aplicados a las imágenes como preparación para el algoritmo de clasificación:

1. Adquisición de la imagen: a partir de la matriz de datos de la probeta, se transforma a una imagen por medio de un proceso específico de coloreado.
2. Pre-proceso: Incluye la aplicación de una serie de técnicas que permiten disminuir la complejidad computacional y facilitar la clasificación:
 - a) Transformación de la imagen en color a una imagen de escala de grises con 256 niveles.
 - b) Cuantificación: los colores de la imagen son reducidos a 32 niveles de grises.
 - c) Aplicación de un filtro de medianas: Se aplica un filtro con una máscara de 9x9 para simplificar la imagen y reducir el ruido existente.
 - d) Segmentación: Para binarizar la imagen se selecciona un punto en

el histograma considerando como punto de corte el 3 % del total de pixeles presentes.

e) Se obtiene el negativo de la imagen.

Un ejemplo de aplicación de los pasos propuestos por Leta se puede ver en la figura 5.3.

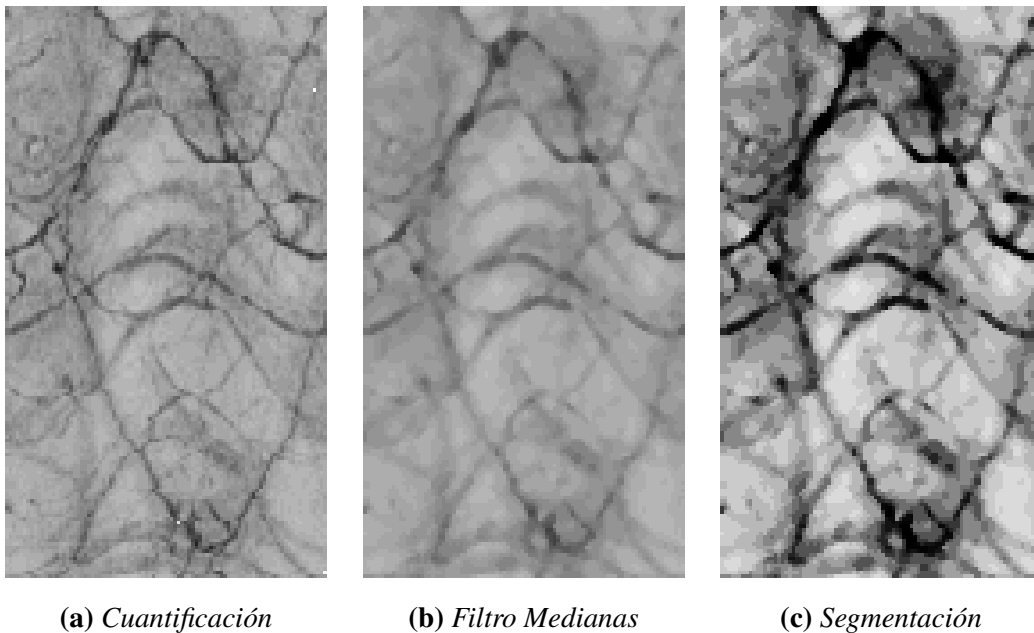


Figura 5.3: *Procesamiento de la imagen propuesto por Leta[18]: Imagen en escala de grises y reducción a 32 niveles de grises (5.3a), aplicación de filtro de medianas (5.3b) y segmentación (5.3c).*

5.1.3. Mejora propuestas sobre la calidad de las imágenes

Debido que el algoritmo de identificación y clasificación depende directamente de la calidad de la imagen binarizada, este paso es clave para su buen funcionamiento. Para ello se propone realizar correcciones sobre la imagen original.

5.1.3.1. Realce de contraste y brillo

Debido a las diferencias de contraste y brillo encontrado en las imágenes, en particular por la diferencias en la calidad del muestreado o por el tipo de terreno analizado, se propone una mejora de los mismos, pero realizado de manera local, es decir, por medio de una ventana deslizante.

En cada paso o ventana deslizante, se realiza ecualización del histograma de los valores encontrados, pero de manera local. Un ejemplo de aplicación se puede ver en la figura 5.4.

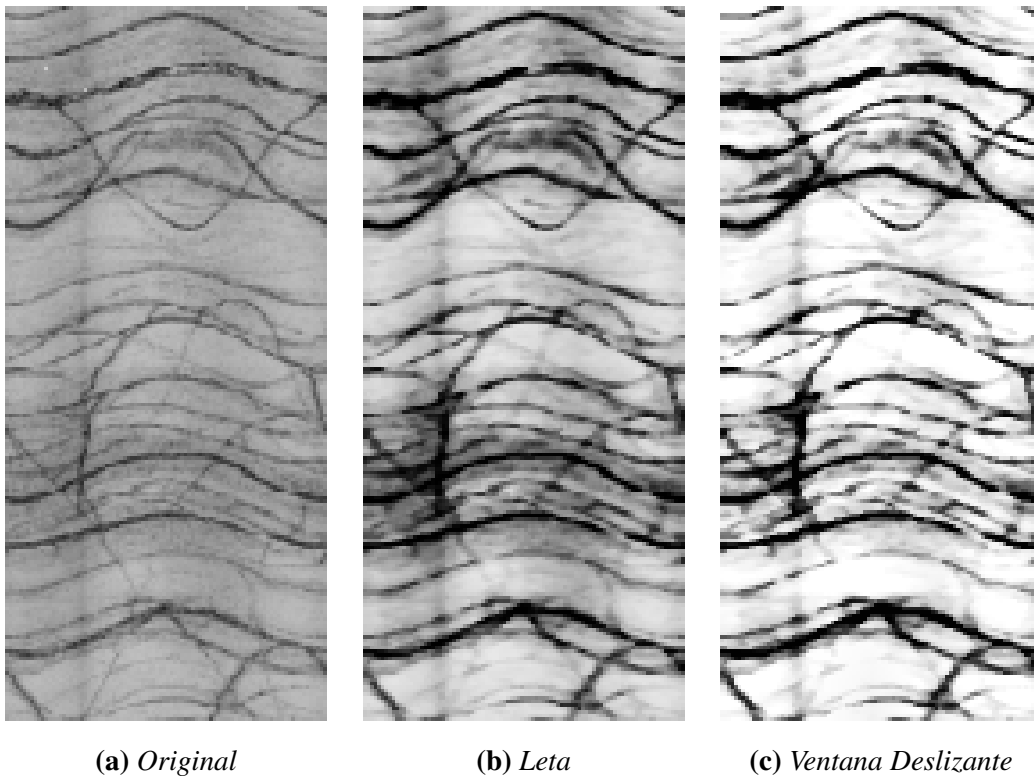
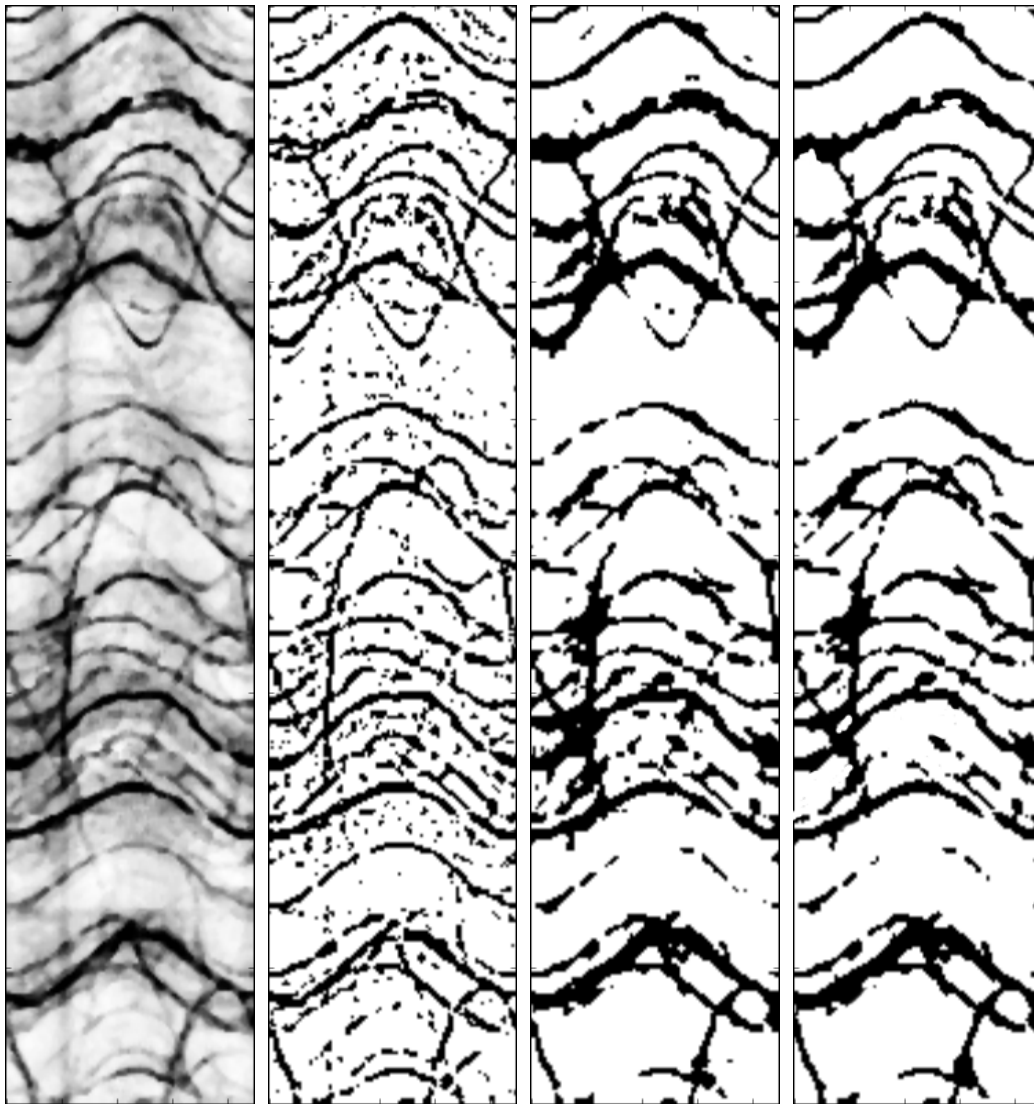


Figura 5.4: Aplicación de realce de contraste y brillo: imagen original (5.4a), procesamiento propuesto por Leta (5.4b) y aplicación de equalización de histograma por medio de una ventana deslizante (5.4c).

5.1.3.2. Binarización y detección de bordes

En el proceso de binarización, se propone utilizar filtros morfológicos y no sólo un valor porcentual propuesto por Leta[18]. Entre los filtros propuestos se propone: OTSU local y luego global[24]. Además utilizar el método adaptativo de Chan [6] y *adaptive thresholding* implementado en la librería scikit-image[29].

Se propone además la utilización del operador de gradiente Sobel en unión con el filtro Canny para detección de bordes, sin embargo, se requiere una evaluación de los parámetros de los mismos para medir su comportamiento.



(a) *Original*

(b) *Adaptativa*

(c) *OTSU Global*

(d) *OTSU Local*

Figura 5.5: *Aplicación de distintos algoritmos de binarización sobre la imagen de un pozo: original, binaria adaptativa, OTSU global y OTSU local. Se puede ver que las diferencias entre OTSU global y local son mínimas, sin embargo esto se explica por la uniformidad de contraste en la imagen original.*

5.1.3.3. Eliminación de ruido

Para la eliminación del ruido se propone el uso de varias técnicas: filtro de medianas, filtros morfológicos (erosión y dilatación) y filtros de ruido gaussiano.

5.2. Identificación de estructuras

Una fractura se puede describir como la intersección de un cilindro con un plano tal como se muestra en la figura 3.5. El resultado de la intersección es una onda sinusoidal, la cual puede ser parametrizada por medio de tres valores: línea base, amplitud y fase original.

5.2.1. Algoritmo de localización e identificación

El algoritmo planteado por Changchun[7], descompone el problema de encontrar los valores de línea base, amplitud y fase; es decir localizar y parametrizar una curva, en dos pasos:

1. Identificar la línea base de la curva por medio de un proceso de votación en un espacio 1-D.
2. Identificar la amplitud y fase original de la curva correspondiente a la línea base por medio del uso de la transformada de Hough en 2-D.

5.2.1.1. Descripción de una curva sinusoidal

La curva sinusoidal buscada S satisface la ecuación 5.1, donde y_0 es la posición de la línea base (profundidad donde se encuentra la curva), A es la amplitud, φ es la fase y ω es la frecuencia angular ($\omega = 2\pi/T$); donde T es el período fijo.

$$y = A \sin(\omega x + \varphi) + y_0 \quad (5.1)$$

5.2.1.2. Identificación de línea base

Changchun[7] propone seleccionar dos puntos en la curva sinusoidal S , con una distancia horizontal entre ellos de $T/2$. La coordenadas de ellos son identificadas como $P(x_p, y_p)$ y $P'(x_{p'}, y_{p'})$, donde se cumple la ecuación 5.2.

$$|x_{p'} - x_p| = T/2 \quad (5.2)$$

$C(x_c, y_c)$ representa el punto medio entre el par de puntos P y P' . La coordenada vertical y_c se puede calcular a través de la ecuación 5.3.

$$y_c = (y_p + y_{p'})/2 \quad (5.3)$$

De acuerdo a la propiedad de una curva sinusoidal, el punto medio de C debe coincidir con la línea base de la onda. En una curva con N puntos, se puede encontrar $N/2$ pares y para cada par de puntos, se puede encontrar su línea base. Si todos los pares de puntos pertenecen a la curva S , su línea base debe ser la misma.

A partir de esta idea, la coordenada vertical de los pares de puntos de la imagen son llevados a un espacio de votación de una dimensión. En este caso, debe existir un *peak* que corresponda la línea base de la curva sinusoidal (Ver figura 5.6).

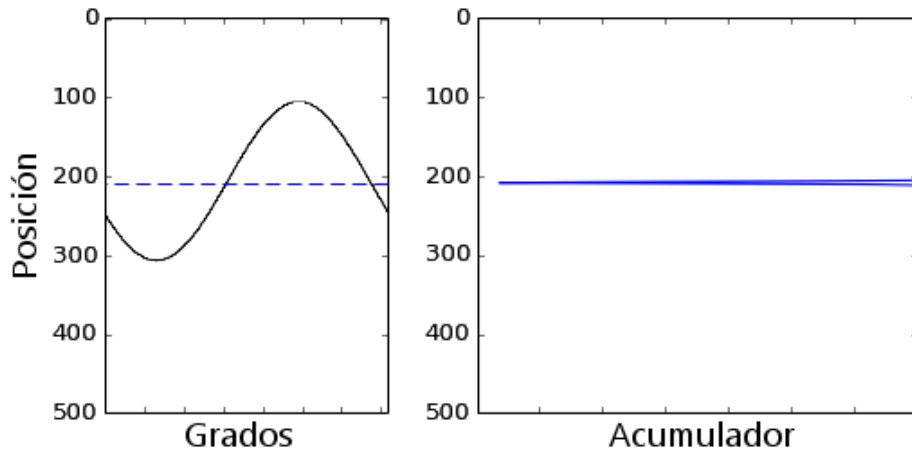


Figura 5.6: Ejemplo de detección de línea base de una onda sinusoidal por medio del espacio de votación. Se identifica un peak en la posición $y_0 = 209$.

5.2.1.3. Determinación de amplitud y fase

Una vez que se ha encontrado la línea base, se deben identificar los otros dos parámetros: Amplitud A y fase original φ . Para ello se utiliza la transformada de Hough.

La ecuación 5.1 puede ser re-escrita como 5.4. En este caso, el espacio de votación 2-D se construye a partir de los valores de A y φ válidos.

$$A = \frac{y - y_0}{\sin(\omega x + \varphi)} \quad (5.4)$$

Para llenar el espacio de votación 2-D, se realiza una iteración de todos los puntos de la imagen $P(x, y)$ que cumplan con la ecuación 5.4, iterando sobre el valor de φ . En este caso, en la posición A, φ del espacio de votación, se incrementa en uno.

Luego de generar el espacio de votación, se busca el máximo general del espacio de votación, en ese máximo se debería encontrar los valores de A y φ asociados a la sinusoidal encontrada en la línea base y_0 (ver figura 5.7).

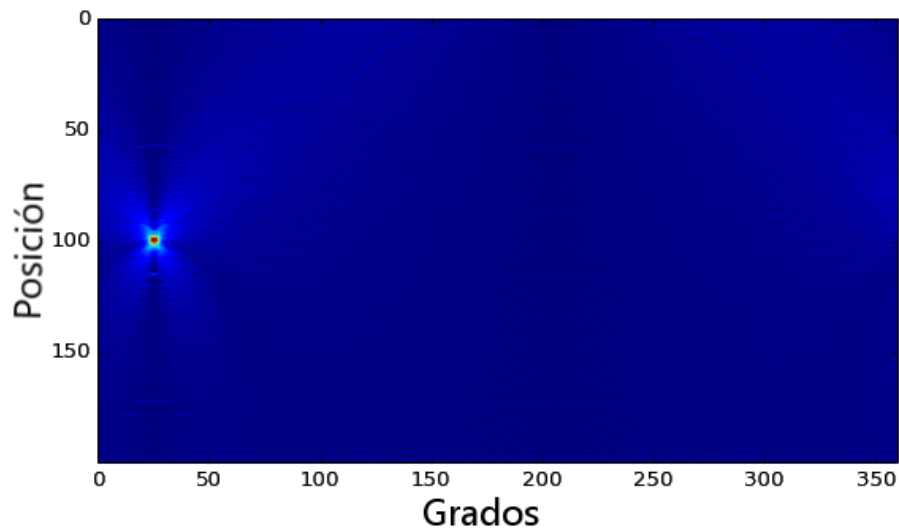


Figura 5.7: Ejemplo de detección de los parámetros A y φ en una onda sinusoidal por medio del espacio de votación 2-D. Se puede apreciar que hay un máximo ubicado en la posición $A = 100$ y $\varphi = 25$, valores con los cuales fue creada la onda de la figura 5.6.

5.2.1.4. Implementación del Algoritmo

El algoritmo fue escrito en el lenguaje de programación Python¹ v2.7.8, utilizando NumPy² v1.9.0 y ejecutado sobre Microsoft Windows 7, instalado sobre una plataforma Intel i7-4700 con 8Gbyte RAM DDR3 3200 Mhz. El IDE de desarrollo fue PyCharm 4³.

¹<https://www.python.org/>

²<http://www.numpy.org/>

³<https://www.jetbrains.com/pycharm/>

5.2.2. Modificación propuesta sobre los parámetros del algoritmo

El algoritmo propuesto por Changchun[7] detecta las líneas base y parametriza las curvas basado en un proceso de votación, sin embargo, no incorpora parámetros asociados a las características específicas de las imágenes: amplitud mínima y máxima de las fracturas buscadas, espesor de las fracturas o calidad de la muestra. Este elemento es clave para mejorar el proceso de clasificación, en particular al contar con la experiencia del experto, el cual puede entregar información sobre la caracterización de las curvas a partir de las características del suelo analizado.

5.2.2.1. Alto de la ventana de análisis

El algoritmo propuesto por Changchun utiliza toda la imagen para construir el espacio de votación 1-D, sin embargo, la línea base de las posibles curvas presentes en la imagen, en *general* no se encuentran correlacionadas entre sí, es decir, una línea base identificada a 5 metros de profundidad no se encuentra relacionada con una encontrada a 20 metros. A raíz de ello, se propone agregar un parámetro que identifica el tamaño máximo de una ventana donde es posible encontrar una línea base.

El valor del alto de la ventana de análisis debe ser lo suficiente para encerrar todos los puntos de una o más ondas en la misma línea base, es por ello, que se decidió seleccionar el doble del alto más grande de una onda que puede ser encontrada en las imágenes, esta información es entregada por parte del experto, en particular debido a que el alto de la onda representa el ángulo con respecto al eje del pozo.

5.2.2.2. Votación en identificación de línea base

La línea base depende de la correcta identificación de los *peaks* en el vector de votación de 1-D. Entre los métodos más utilizados se encuentra buscar la localización donde la derivada de la curva es cero, pero debido al ruido introducido, aparece un a gran cantidad de detecciones en falso. Otro método utilizado es suavizar la curva con algún filtro pasa-bajos, pero los resultados son sub-óptimos debido a que se elimina algunos *peaks* importantes en la señal.

El método propuesto se basa en el ideado por Eli Billauer: **peakdet** [3], el cual se basa en la idea que un *peak* usualmente se encuentra entre dos valles de la señal.

El algoritmo **peakdet** es sensible a dos parámetros que deben ser seleccionados cuidadosamente para una correcta detección:

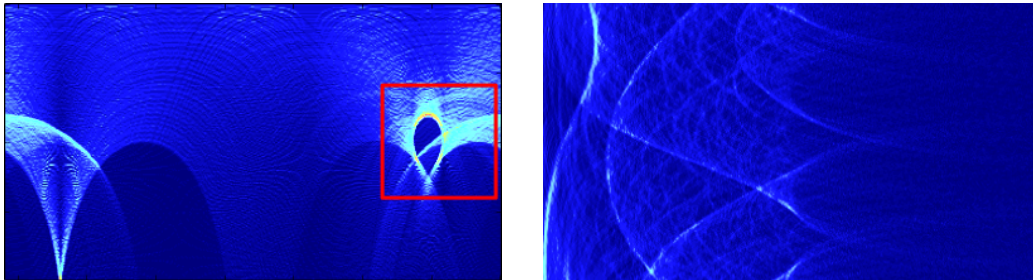
1. **Lookahead**: Distancia a mirar hacia adelante en la señal desde un posible peak. Esto con el fin de determinar si efectivamente es un peak o no. El autor recomienda utilizar el valor: (tamaño de muestra / período buscado) / f , donde $1,25 \leq f \leq 4$.
2. **Delta**: Especifica la diferencia mínima entre un *peak* y los puntos en su vecindad para que ese valor sea considerado un *peak*. El autor recomienda utilizar un valor $\text{delta} \geq \text{RMSnoise} * 5$.

5.2.2.3. Votación en amplitud y fase

La detección de la amplitud y la línea base depende directamente de él o los *peaks* identificados en el espacio de votación 2-D construido. Un *peak* en este espacio de votación es indicativo que para la línea base especificada, existe una amplitud y fase donde la mayoría de los puntos calzan sobre la onda sinusoidal.

En el caso de existir varios *peaks* en este espacio, podría indicar la co-existencia de ondas sobre la misma línea base.

El valor del *peak* además es indicativo de cuantos puntos caen exactamente en una amplitud y fase en particular de una onda. A raíz de este fenómeno, el valor obtenido o su distribución puede ser utilizado para descartar falsos positivos de líneas base (ver figura 5.8).



(a) *Múltiples Máximos*

(b) *Distribución Uniforme*

Figura 5.8: *Espacios de votación 2-D: Múltiples máximos pueden ser identificados, indicando que es posible encontrar sobre la misma línea base más de una sinusoidal (izq.). Distribución uniforme, indicando que no es posible encontrar los valores de amplitud y fase para una línea base (der.).*

5.2.3. Modificación propuesta al algoritmo

Debido a que la precisión del algoritmo identificador de líneas base y parametrización de curvas, se encuentra directamente asociado a la calidad del pre-proceso de la imagen (eliminación de ruido y binarización), se puede concluir que esta fase es la más importante en el proceso completo, sin embargo, es justamente en ésta, donde información relevante sobre la curva puede ser eliminada.

Considerando además que la calidad de las imágenes a identificar es muy variable, y basado en la observación que los puntos que pertenecen a una curva deberían poseer las mismas características, resultado de pertenecer a un mismo plano que intersecta con las paredes del pozo y por lo tanto, con muy alta probabilidad de pertenecer al mismo tipo de roca. Ésto sumado a que el proceso

de binarización es crítico para el buen funcionamiento del algoritmo, se propone modificar el algoritmo para considerar una ventana de trabajo (vecindad) en vez de un valor binario.

5.2.3.1. Semejanza por vecindad

Se propone modificar el proceso de búsqueda de puntos que pertenecen a la onda sinusoidal, por la búsqueda de zonas con cierto nivel de semejanza (textura) que pertenezcan a la onda. Para ello, se propone cambiar la comparación punto a punto, por una realizada vecindad a vecindad.

En la figura 5.9 se ejemplifica la mejora propuesta: revisar la vecindad de un punto en particular e identificar otro punto que posea las mismas características.

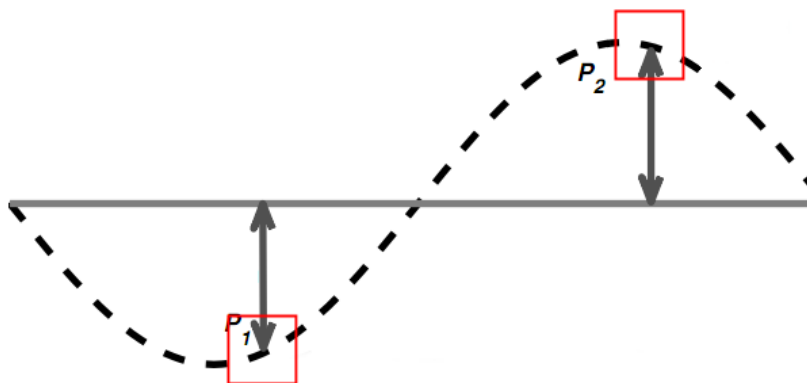


Figura 5.9: El punto P_1 y P_2 deberían poseer las mismas características, esto debido a que pertenecen al mismo plano y por lo tanto con una altísima probabilidad de pertenecer al mismo tipo de material.

La comparación entre dos vecindades asociadas a dos puntos distintos de la curva en particular, debe poseer algunas características específicas: ser inmune a la rotación, ser representativa de los elementos constituyentes de la vecindad y obtener un valor que indique la semejanza entre las dos vecindades.

A partir de las características buscadas, se propone la utilización de **Local Binary Patterns** [23] para representar la matriz de vecindad por medio de un vector de características. La utilización de LBP supone una gran cantidad de ventajas: Puede ser muy tolerante a cambios de iluminación, distorsión debido a cambios de perspectiva, zonas borrosas y además a la oclusión. Justamente, son estos los elementos que abundan en las texturas que constituyen parte de una fractura.

Una vez identificado el vector de características, la semejanza entre las distintas vecindades o *patches* puede ser obtenida vía la comparación de los vectores de características extraídos vía LBP.

La comparación propuesta entre los dos vectores de características se puede ver en la ecuación 5.5: Dados los vectores de características LBP **A** y **B**, la magnitud de un vector **A** definida como $\|\mathbf{A}\|$, el ángulo θ que forman entre ellos se encuentra definido por el producto punto entre ellos.

$$\mathbf{A} \cdot \mathbf{B} = \|\mathbf{A}\| \|\mathbf{B}\| \cos \theta \quad (5.5)$$

5.2.3.1.1. Algoritmo de Detección de Líneas Base El algoritmo con las modificaciones propuestas se puede resumir como:

- 1) Partiendo con la imagen considerando como un arreglo 2-D de pixels, inicializar un acumulador de largo H (altura de la imagen) y avanzar por la imagen fila por fila.
- 2) Para cada fila, considerar una ventana P_1 de pixels de tamaño n arriba y abajo de la fila (vecindad). A partir del centro de esta ventana, identificar una nueva ventana P_2 espejada horizontalmente por $\frac{W}{2}$ (donde W es el ancho de la imagen) con respecto al centro de la imagen y luego deslizar esa ventana en la vertical $\frac{A}{2}$ (donde A es el ancho de la ventana de análisis) tanto hacia abajo y arriba con respecto a la fila actual (figura 5.9).

- 3) Para cada ventana P_2 identificada en la vertical, obtener su vector LBP y realizar la comparación entre ángulos con el vector LBP de P_1 . Si es positiva, incrementar el acumulador H asociado a esa fila.
- 4) Una vez recorrida toda la imagen, identificar los *peaks* presentes en el acumulador H , indicativos de posible localización de línea base.

5.3. Métricas de evaluación

Con el fin de evaluar las mejoras propuestas al procesamiento de las imágenes y al algoritmo de identificación y clasificación; y compararlas con el trabajo realizado por un experto, es necesario definir algunos indicadores.

5.3.1. Medida de rendimiento de un algoritmo de localización e identificación

La evaluación del rendimiento de identificación de fracturas se realiza principalmente basado en un criterio visual con el fin de detectar el porcentaje de fracturas correctamente identificadas [1] [7] [18]: Se proyecta la fractura identificada por el algoritmo sobre la imagen del pozo y se verifica la existencia de la misma y su desviación o error en la amplitud y desfase. Este proceso es subjetivo y afecto a errores a partir de la experiencia de la persona realizando la inspección.

El proceso de inspección manual puede ser una tarea muy tediosa y con alto consumo de tiempo, en particular debido al gran largo que poseen las imágenes de pozos.

Para eliminar la interferencia humana en los resultados de clasificación obtenidos, se propone una técnica de evaluación automática basada en la propuesta por Louloudis[20]. En ella se compara los resultados obtenidos por el algoritmo de clasificación con respecto a la clasificación realizada por los expertos.

La medida de rendimiento del algoritmo puede ser expresada en términos de los parámetros que identifica y clasifica: línea base, amplitud y fase. Es importante diferenciar distintas métricas debido a que la fase de clasificación del algoritmo depende directamente de la identificación de líneas base y a partir de este dato, identificar los parámetros de la onda.

5.3.1.1. Línea base

Debido a que la detección de amplitud y fase es dependiente de una correcta identificación de la línea base, este parámetro es uno de los más importante a evaluar. Concretamente si una línea base es incorrectamente identificada, no representa gran problema para el algoritmo completo, debido a que en la fase de identificación de amplitud y fase, la votación en el espacio 2-D será muy baja y por lo tanto susceptible de descarte.

Se propone una métrica porcentual que identifique el valor del rendimiento del algoritmo con respecto a la clasificación realizada por el experto, esta métrica es semejante a la exhaustividad, pero con la variante de permitir cierta holgura o error en la posición.

5.3.1.1.1. Métrica Sean LE y LR las líneas bases definidas por el experto y las obtenidas como resultado del algoritmo de clasificación respectivamente:

$$LE = \{LE_i\}, i \in [1, LE_{num}] \quad (5.6)$$

$$LR = \{LR_i\}, i \in [1, LR_{num}] \quad (5.7)$$

Es posible definir que una línea base identificada por el experto LE_k ha sido correctamente identificada por el algoritmo, en el conjunto de líneas clasificadas

por el algoritmo, de acuerdo al valor de la función $EXT(LE_k)$.

$$EXT(LE_k) = \begin{cases} 1 & \text{si } \exists k' : \frac{|LE_k - LR_{k'}|}{resolución} < distancia \\ 0 & \text{en caso contrario} \end{cases} \quad (5.8)$$

De acuerdo a la ecuación 5.8, una línea identificada por el experto LE_k ha sido correctamente detectada, sólo si existe una línea $LR_{k'}$ en el conjunto de resultados que se encuentre en la misma posición o a lo más alejado $distancia$, valor definido empíricamente en 0.5. El valor $resolución$ define el error esperado en las imágenes del pozo debido al espesor de las fracturas.

Finalmente, el rendimiento de la detección de la línea base se define por la función $ExLin$ definida en la ecuación 5.9.

$$ExLin = \frac{\sum_{k=1}^{LE_{num}} EXT(LE_k)}{LE_{num}} 100\% \quad (5.9)$$

5.3.1.1.2. Precisión Para evaluar la calidad de la línea base se propone además incorporar la medida de precisión, definida en la ecuación 5.10, donde tp son los verdaderos positivos (líneas identificadas en la posición correcta), fp son los falsos positivos (identificación de una línea donde no existe realmente).

$$Precisión = \frac{tp}{tp + fp} \quad (5.10)$$

5.3.1.2. Amplitud y fase

En este caso, se propone identificar en que porcentaje la curva se ajusta a la real, tanto en error de amplitud y en des-fase.

5.3.1.2.1. Métrica En este caso se propone utilizar el error cuadrático medio como estimador de la calidad de la clasificación de la amplitud y la fase.

$$e_{\varphi} = \frac{\sqrt{\sum(\varphi_e - \varphi_r)^2}}{n} \quad (5.11)$$

$$e_a = \frac{\sqrt{\sum(a_e - a_r)^2}}{n} \quad (5.12)$$

5.4. Conjunto de imágenes de evaluación

Para la evaluación del rendimiento del algoritmo de clasificación, con respecto a la realizada por un experto, es necesario contar con imágenes representativas de los posibles escenarios a encontrar.

El algoritmo de identificación y parametrización de curvas fue aplicado sobre dos conjuntos principales de imágenes:

1. 100 imágenes con fracturas generadas de manera sintética para comparar los resultados obtenidos por Changchun [7].
2. 45 imágenes obtenidas de 2 pozos distintos de la II región de Antofagasta proporcionados por Andes Geofísica.

5.4.1. Construcción de imágenes sintéticas

La construcción de las imágenes se realiza seleccionando valores para los tres parámetros básicos de identificación: línea base, amplitud y fase. Estos fueron generados al azar, pero considerando las restricciones de las posibles curvas a encontrar en un pozo real. Las restricciones fueron obtenidas a partir de la opinión del experto.

1. **Línea Base:** entre $y_0 = 0$ y $y_0 = 500$, considerando un trozo de 500 pixels de alto, lo que corresponde aproximadamente a 3 metros de profundidad a

partir de la resolución comúnmente utilizada.

2. **Amplitud:** entre $A = 50$ y $A = 150$ pixels, lo que corresponde a una amplitud entre 29cm. y 89cm. aproximadamente.
3. **Fase:** entre $\varphi = 0^\circ$ y $\varphi = 360^\circ$ grados, indicando que la medición se realizó con un muestreo de 1 grado de diferencia.

Para realizar los experimentos con imágenes sintéticas se construyeron alrededor de 1000 imágenes con curvas generadas aleatoriamente, pero aplicando las restricciones impuestas por los parámetros.

En la figura 5.10 se muestra una imagen artificial que contiene curvas que representan las que se podrían encontrar en un pozo en opción del experto.

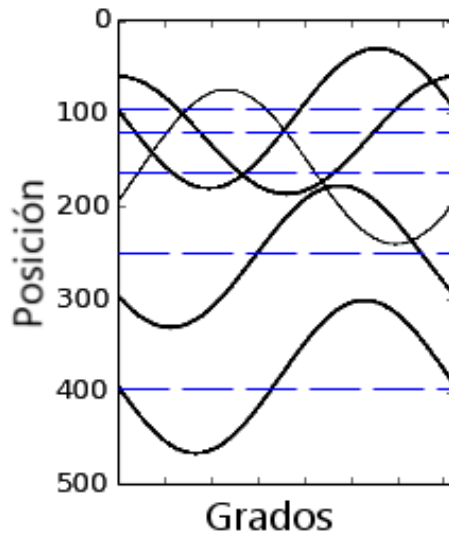


Figura 5.10: *Curvas sintéticas generadas a partir de las restricciones definidas por el experto. Se dibuja con una línea segmentada la línea base de cada curva.*

5.4.1.1. Ruido en las imágenes

Una vez generadas las imágenes artificiales de las curvas a partir de las restricciones entregadas por el experto, se procedió a agregar ruido para simular las

condiciones de prueba propuestas por Changchun[7].

El ruido es agregado para evaluar la robustez del proceso de identificación y clasificación, para ello se agregó ruido gaussiano de dos formas:

1. **Ruido Negro:** representa la información que puede ser encontrada entre las curvas, en este caso la dificultad que aplica al proceso de clasificación es la posibilidad de incorporar alguno de esos puntos a una curva no real.
2. **Ruido Blanco:** representa la posibilidad de eliminar trozos de las curvas presentes en la imagen, ésto con el fin de simular una curva con trozos faltantes.

La cantidad de ruido agregado a la imagen puede ser controlado para analizar la robustez con respecto a puntos no pertenecientes a una curva (negro) y trozos eliminados de la curva (blanco). En la figura 5.11 se puede ver un ejemplo de aplicación de ruido blanco y negro.

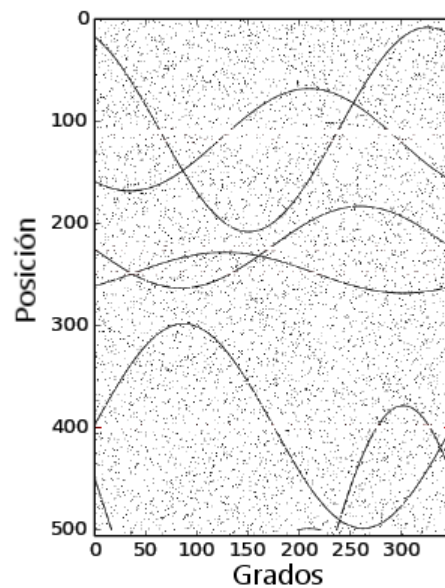


Figura 5.11: Curvas artificiales incluyendo niveles de ruido y eliminación de trozos de las mismas, simulando las condiciones naturales.

5.4.2. Imágenes clasificadas por parte del experto

Las imágenes de los pozos proporcionados por Andes Geofísica, fueron previamente procesadas por parte de los expertos, obteniendo aproximadamente 500 mt. lineales de pozo con ondas ya clasificadas (posición, amplitud y desfase). La clasificación entregada sirve como conjunto contra el cual contrastar el desempeño del algoritmo de clasificación.

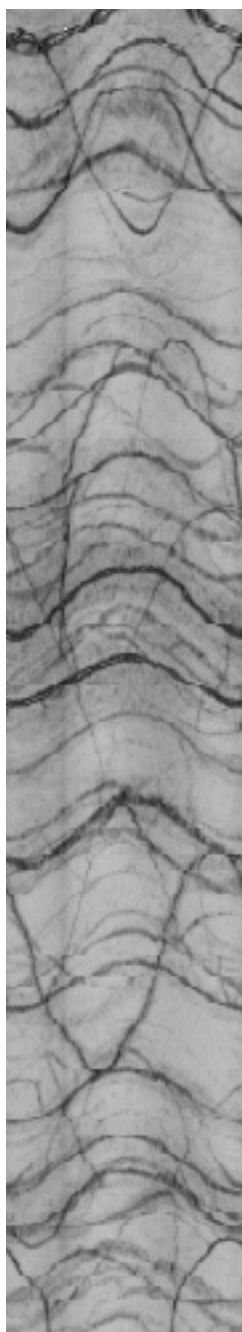
La información de los pozos entregados, proveen un espectro de análisis que puede ser separado en tres conjuntos, de acuerdo a la dificultad por parte del experto del análisis: fácil, medio y difícil.

1. **Fácil:** Las curvas son fácilmente reconocibles, ya sea por el contraste con las paredes o por el espesor de la misma.
2. **Media:** Las curvas no son fácilmente reconocibles y mediante prueba y error se ajusta la curva a los puntos que son visibles en las imágenes.
3. **Difícil:** Es complejo identificar las curvas, particularmente por el poco contraste entre éstas y las paredes del pozo, porque se encuentran parcialmente obstruidas o por un mal proceso de toma de muestra.

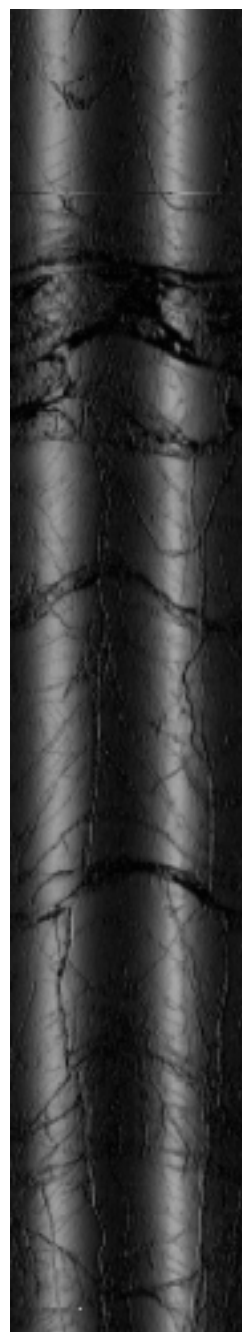
Para la evaluación del algoritmo, se seleccionaron 15 imágenes de cada clase, cada imagen de tamaño 500 x 90 pixeles. En la figura 5.12 se puede ver las distintas clasificaciones realizadas por el experto.



(a) *Fácil*



(b) *Medio*



(c) *Difícil*

Figura 5.12: *Distintos tipos imágenes de pozos clasificadas por el experto de acuerdo a la complejidad de identificación de fracturas.*

Capítulo 6

Resultados

El algoritmo de clasificación automática y las mejoras propuestas fueron aplicadas sobre dos familias de imágenes: una generadas de manera sintética a partir de los parámetros reales de las curvas que podrían ser encontradas en un pozo, pero generados al azar; y sobre las imágenes reales con fracturas pre-clasificadas por parte del experto.

6.1. Clasificación sobre Imágenes Sintéticas

Se realizó un análisis del comportamiento del algoritmo sobre un conjunto de imágenes que representan las fracturas posible de encontrar en un pozo. Las imágenes fueron generadas a partir de la información entregada por el experto y un conjunto de imágenes de pozos reales.

6.1.1. Clasificación de Fracturas Sintéticas

Para evaluar el rendimiento del algoritmo de clasificación propuesto, se generaron imágenes de fracturas con valores de los parámetros seleccionados al azar entre el rango de valores posibles.

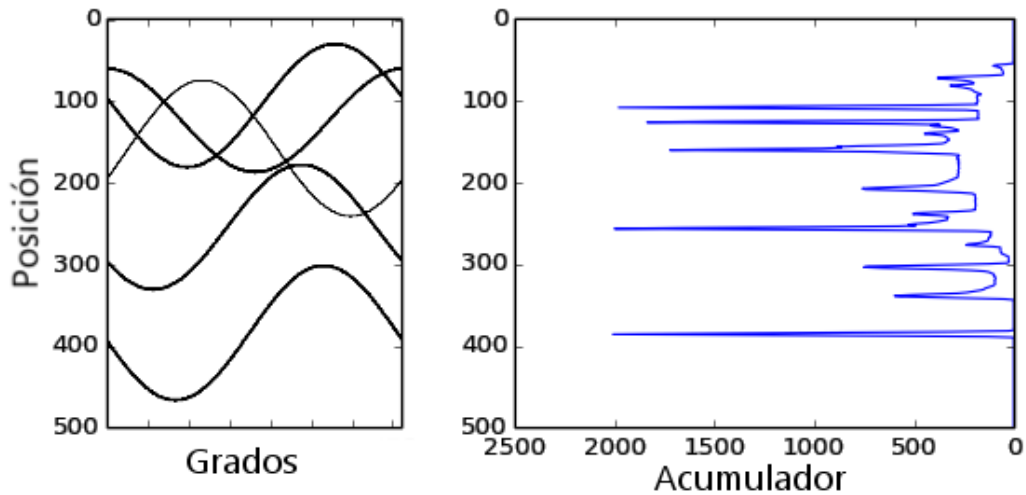


Figura 6.1: *Curvas que representan fracturas sintéticas generadas con parámetros de posición, amplitud, fase y espesor entregados por el experto (izq.). Valores obtenidos en el espacio de votación 1-D asociado a las curvas (der.), se logra identificar que cada peak en este espacio se encuentra asociado a la línea base de una fractura.*

En la figura 6.1 se puede ver el resultado de un experimento realizado, donde se generaron 5 curvas en un espacio entre 0 y 500, variando amplitud, fase y espesor de la onda que representa la fractura. Se muestra el espacio de votación 1-D generado para identificar la posible localización de líneas base. Se puede ver que en los *peaks* del espacio de votación, se encuentra asociada una línea base.

Tabla 6.1: Resultado del proceso de identificación de líneas base, amplitud y fase de curvas generadas al azar (largo del pozo $L = 500$ metros, número de curvas $N = 5$). Se resaltan los valores diferentes obtenidos en la clasificación. El error acumulado de la amplitud y ángulo es $e_\varphi = 0,28$ y $e_a = 0,2$ respectivamente; y el valor de la métrica de calidad $ExLin = 100\%$. Se resalta además la diferencia entre los valores generados y los identificados.

Valores Generados			Valores Identificados		
y_0	A	φ	y_0	A	φ
109	75	353	109	75	354 (+1)
127	63	270	126 (-1)	62 (-1)	270
161	83	154	159 (-2)	83	154
257	76	34	257	75 (-1)	34
386	82	7	386	82	7

En la tabla 6.1 se pueden ver los valores de los parámetros generados y los identificados por el algoritmo en el experimento de la figura 6.1. Al calcular el valor de la métrica de calidad del algoritmo en la identificación de líneas base, aplicado sobre las curvas de la figura 6.1, el valor obtenido es $ExLin = 100\%$ (ver ecuación 5.9). Se utilizó el valor de *resolución* = 2 (pixels por metro lineal de pozo), es decir, la posición de la línea base desviada a lo más en 2 pixels.

Se realizaron varias pruebas del algoritmo de clasificación con las mejoras propuestas (alto de ventana y *peakdetect* para el proceso de votación de líneas base), generando al azar distintas curvas, con distintos parámetros, variando la longitud del pozo y manteniendo constante el número de curvas, de esta forma las curvas se empiezan a superponer unas sobre otras. Los resultados obtenidos se puede ver en la tabla 6.2. Cada fila es el promedio de 100 experimentos sobre un pozo de largo $L = 2000$ metros.

Tabla 6.2: Resultado del proceso de identificación de líneas base, amplitud y fase para distinto número de curvas generadas la azar sobre la misma longitud de pozo ($L = 2000$). Cada fila representa el promedio de la ejecución de 100 experimentos con los mismo valores.

N° Curvas	ExLin	Precisión	e_A	e_φ
20	98 %	94.5 %	0.221	0.502
22	96 %	95.7 %	0.292	0.784
24	84 %	95.3 %	0.402	0.863
26	76 %	94.7 %	0.478	0.932
28	68 %	94.4 %	0.606	1.146
30	58 %	93.7 %	0.720	1.281
32	51 %	94.2 %	0.861	1.351
34	45 %	93.8 %	0.802	1.392
36	43 %	92.2 %	0.944	1.458
38	38 %	90.7 %	1.004	1.691
40	34 %	89.8 %	1.869	2.002

Los resultados obtenidos sugieren que a medida que se aumenta la densidad de curvas presentes, es decir, algunas de ellas sobrelapándose, la calidad del algoritmo de identificación de líneas base disminuye (ver figura 6.2), sin embargo, a pesar de ello el error acumulado de la amplitud y ángulo de fase aumenta más lentamente hasta llegar a un valor de saturación, donde crece rápidamente. Este hecho es indicativo que el proceso de votación 2-D para parametrizar la amplitud y fase es más robusto que la votación 1-D para identificar la línea base.

El valor de la precisión muestra que el algoritmo es resistente a los falsos positivos, es decir, encontrar curvas donde no existen. Esto se puede explicar debido a que si el algoritmo de identificación de línea base identifica una posición

donde no hay una curva, la fase de parametrización de amplitud y fase la descarta ya que existirán muchos máximos locales en el espacio de votación 2-D, indicativo de que en realidad no existe una curva en esa posición.

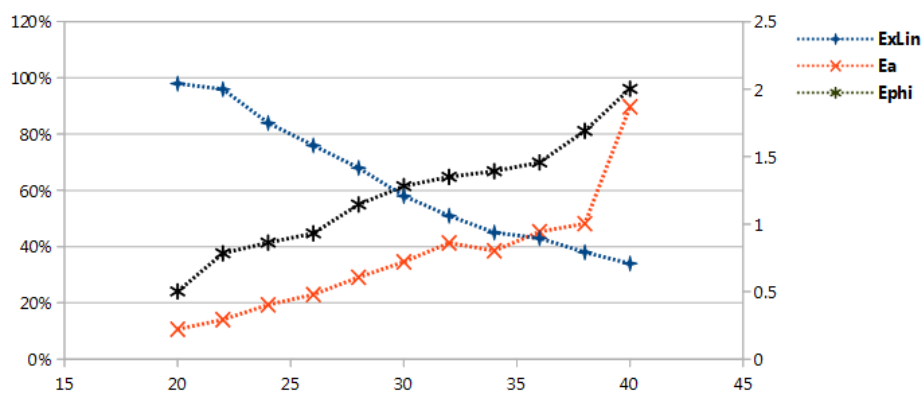


Figura 6.2: Gráfico del valor de la métrica de calidad *ExLin* a medida que aumenta la densidad de curvas. El número de curvas aumenta en el eje X y se mantiene constante la longitud del pozo.

6.1.1.1. Espacio de Votación 2-D (amplitud y fase)

En la figura 6.3 se puede ver el espacio de votación 2-D obtenido mediante la aplicación de la transformada de Hough a la parametrización de una fractura. En este caso claramente se nota el *peak* en la posición donde los parámetros de amplitud y fase se magnifican. Además se pueden ver otros puntos donde la votación es menor, indicando que en esos puntos posiblemente se logra identificar más de una senoide en la misma línea base. Este hecho sugiere que al igual que el algoritmo de votación para la línea base, es posible realizar una búsqueda de *peaks* para encontrar los valores de amplitud y fase sobre curvas que posean la misma línea base.

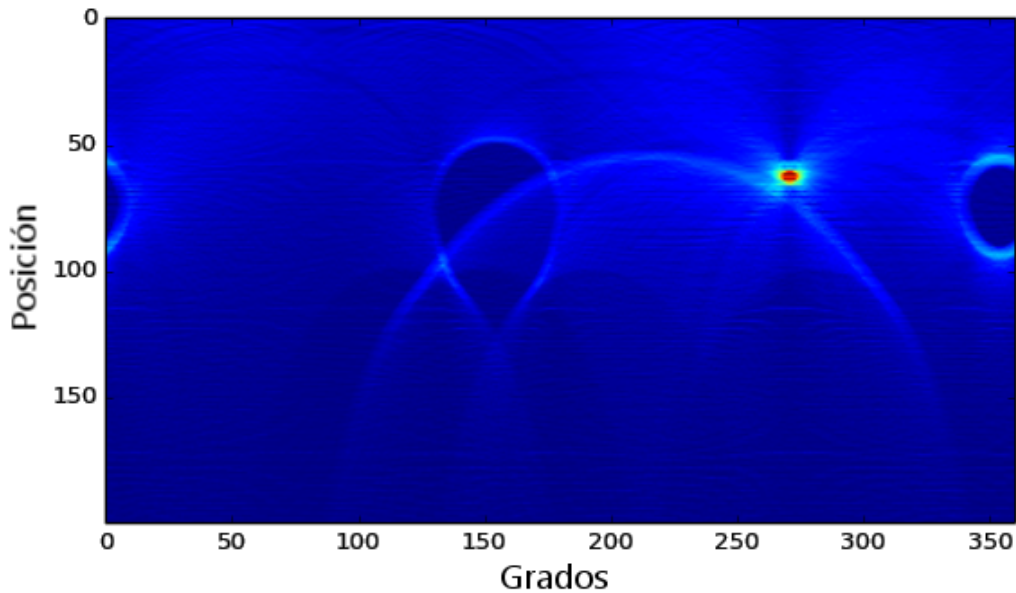


Figura 6.3: *Espacio de votación 2-D para la curva generada en la línea base $y_0 = 127$.
En este caso los parámetros identificados son $y_0 = 126$, $A = 62$ y $\varphi = 270$.*

6.1.1.2. Comparación con el algoritmo original

Al realizar una comparación entre el algoritmo original y el propuesto, se puede identificar que los valores de la métrica de calidad *ExLin* son cercanos en el caso de curvas no sobrepuestas (ver tabla 6.3), sin embargo, los tiempos de ejecución del algoritmo original es notoriamente superior. Una posible interpretación de este hecho es la inclusión del concepto de alto de ventana de trabajo, que el caso del algoritmo original no existe y por lo tanto se debe procesar toda la imagen (proceso de votación) para cada línea base identificada.

Es necesario indicar además que en el caso de algoritmo original, la posición de los máximos en el vector 1-D fueron seleccionados solamente los máximos generales, que en el caso de existir alguna curva sobre otra, elimina la posibilidad de identificarla.

Los errores de amplitud y fase también son superiores en el algoritmo original, posiblemente al incremento del espacio de votación 2-D a toda la imagen y por lo tanto, algunas curvas influyen en los parámetros de otras.

Tabla 6.3: Resultado del proceso de identificación en términos de la función de calidad *ExLin* de identificación de línea base y tiempos de ejecución. Fase 1 representa el tiempo en encontrar las posibles líneas base. Fase 2 es el tiempo necesario para identificar los parámetros de amplitud y fase. El número de curvas buscadas $N = 20$ y largo del pozo $L = 2000$ metros. Los valores obtenidos es el promedio de 100 ejecuciones de cada experimento.

Algoritmo	ExLin	e_A	e_φ	Fase 1	Fase 2
Mejorado	97 %	0.226	0.496	594ms	2242ms
Original	94 %	0.952	2.108	18093ms.	4777ms.

6.1.2. Clasificación con Ruido

Para probar la robustez del proceso de identificación y clasificación, se procedió a agregar ruido normal blanco y negro de manera separada. Se realizó de esta forma para eliminar trozos de la senoide y además generar elementos de ruido en la misma, tratando de simular las condiciones en que se puede encontrar en la imagen original.

6.1.2.1. Clasificación Sin Limpieza

Se realizó el experimento de aplicar el algoritmo sobre la imagen original con ruido con distribución normal $\sigma = 0,1$ binarizada sin incluir ningún proceso de limpieza. El valor de la métrica de calidad obtenida fue $ExLin = 57,7\%$. Este valor indica que el ruido introducido afecta adversamente la identificación de líneas base y por lo tanto, se requiere el proceso de limpieza de la imagen con

el fin de disminuir el ruido introducido.

6.1.2.2. Clasificación con Limpieza

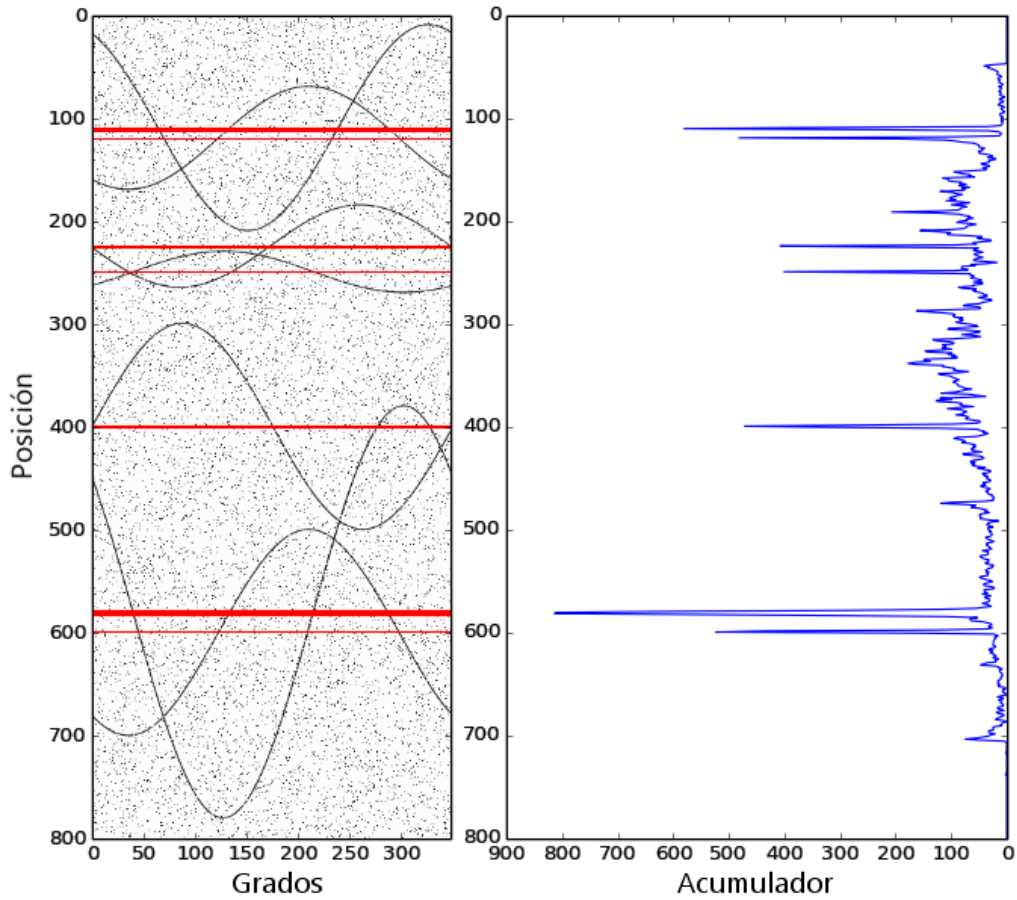


Figura 6.4: *Espacio de votación 1-D para siete curvas generadas al azar y a la cual se ha agregado ruido ($\sigma = 0,1$). Se puede identificar que los peaks identifican las líneas base de la imagen original.*

La limpieza de la imagen comenzó con la binarización vía *threshold* sobre el histograma de los valores encontrados [18], luego se aplicó un filtro de medianas

con el fin de eliminar el ruido presente. En la figura 6.4 se puede ver un experimento donde se agregaron 7 curvas ($N = 7$) a un pozo de largo $L = 800$ sobre una imagen con ruido ($\sigma = 0,1$). Se puede observar que a pesar de la presencia de ruido, los *peaks* en el vector de votación 1-D claramente reflejan la posición de las líneas base. Este hecho es indicativo de la robustez del algoritmo de identificación de líneas base.

Una vez identificadas las líneas base, se procedió a identificar los parámetros de amplitud y fase, por medio del espacio de votación 2-D construido aplicando la transformada de Hough. Los valores obtenidos se pueden encontrar en la tabla 6.4, siendo éstos exactamente los mismos utilizados para generar las curvas.

Tabla 6.4: Resultado del proceso de identificación de líneas base, amplitud y fase sobre un conjunto de curvas generadas al azar, superpuestas y con ruido $\sigma = 0,1$ (ver figura 6.4).

Valores Generados			Valores Identificados		
y_0	A	φ	y_0	A	φ
110	100	295	110	100	295
119	50	54	119	50	54
224	40	2	224	40	2
249	20	129	249	20	129
399	100	180	399	100	180
580	200	319	580	200	319
599	100	54	599	100	54

Para identificar la calidad del algoritmo de identificación, se generaron varias curvas al azar en una imagen de largo representativo al largo de un sondaje ($L = 2000$), sobre las cuales se realizó la identificación. En la figura 6.5 se puede ver un subconjunto de las pruebas realizadas. Las líneas horizontales son indicativos

donde el algoritmo identificó una posible línea base.

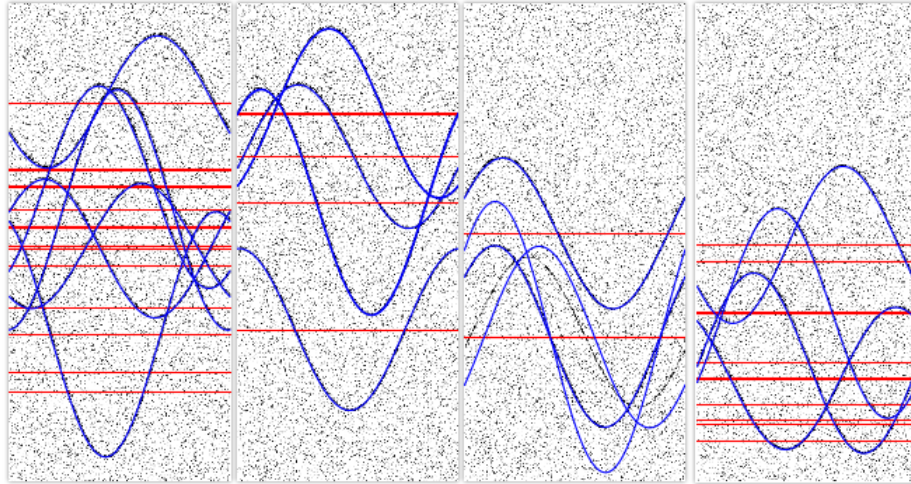


Figura 6.5: Resultados para distintos valores de curvas generadas al azar a las cuales se les agregó ruido. Se puede apreciar la correcta identificación de las líneas base en todos los casos.

6.1.2.3. Sensibilidad al Ruido

Para medir la sensibilidad del proceso de limpieza, identificación y parametrización de fracturas, se aplicó el algoritmo a imágenes con distintos niveles de ruido. Los resultados obtenidos se puede ver en la figura 6.6.

El experimento consistió en ir incrementando la distribución de ruido con valores entre 0 y 1, en pasos de 0.1. Se puede ver que para valores de ruido inferiores o igual a $\sigma = 0,4$ el algoritmo no presenta problemas en detectar las líneas base. A partir de $\sigma = 0,4$ la calidad de la identificación de línea base (*ExLin*) decae y los valores de error de amplitud y fase suben rápidamente.

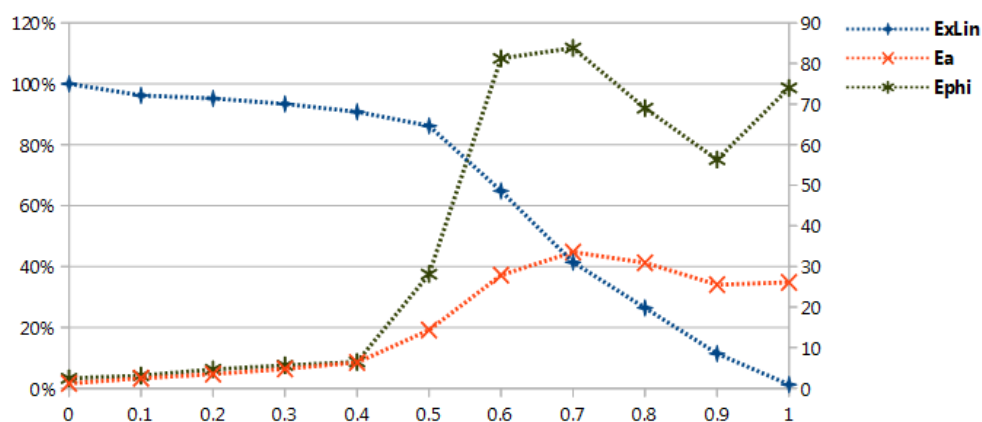


Figura 6.6: Gráfico del valor de la métrica de calidad *ExLin*, el error de amplitud y fase a medida que aumenta los niveles de ruido presente en la imagen. Los valores presentados con el promedio de la ejecución de 100 experimentos.

6.1.2.4. Resultados

Se puede identificar a partir de los experimentos realizados (ver figura 6.6 y 6.7, tabla 6.5), que el algoritmo no posee problemas en identificar curvas superpuestas o con la misma línea base, además de ser resistente al ruido presente en la imagen hasta cierto punto, sin embargo, la métrica de calidad baja rápidamente

a medida que aumenta el número de curvas sobrepuestas, la cantidad de ruido presente, además el error en la amplitud y fase sube rápidamente.

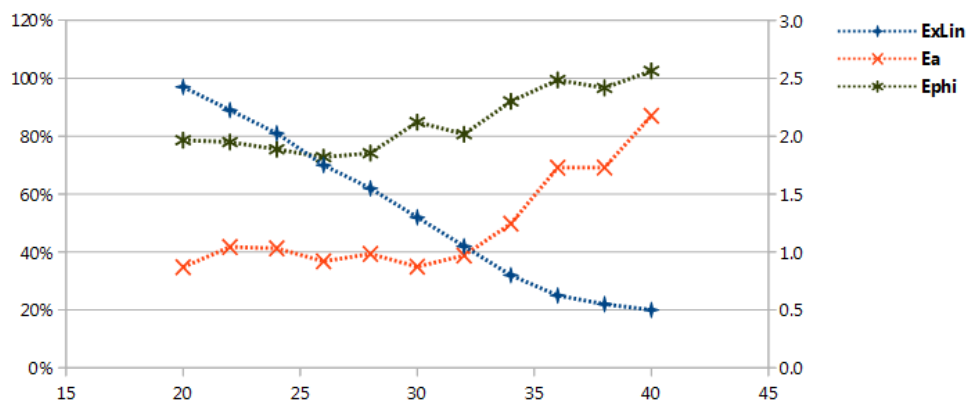


Figura 6.7: Gráfico del valor de la métrica de calidad *ExLin* a medida que aumenta el número de curvas sobrepuestas. El eje *X* indica el número de curvas presentes manteniendo constante el largo del pozo ($L = 2000$ metros). Los valores obtenidos son el promedio de realizar cada experimento 100 veces.

Tabla 6.5: Resultado del proceso de identificación de líneas base, amplitud y fase para distinto número de curvas generadas la azar sobre la misma longitud de pozo ($L = 2000$ metros). Cada fila representa el promedio de la ejecución de 100 experimentos con los mismo valores.

N° Curvas	ExLin	e_A	e_φ
20	97 %	0.869	1.965
22	89 %	1.043	1.949
24	81 %	1.032	1.887
26	70 %	0.920	1.819
28	62 %	0.984	1.854
30	52 %	0.874	2.122
32	42 %	0.970	2.017
34	32 %	1.245	2.300
36	25 %	1.728	2.485
38	22 %	1.730	2.416
40	20 %	2.176	2.564

6.2. Identificación de Fracturas en Imágenes Reales

Previo al desarrollo de experimentos de identificación de líneas base y parametrización sobre las imágenes seleccionadas por el experto, se procedió a aplicar el algoritmo sobre dos trozos de imágenes representativas de la información a encontrar en pozos.

6.2.1. Clasificación Sobre Imágenes Específicas

Se aplicó el algoritmo de clasificación a dos imágenes extraídas de pozos diferentes, que juicio del experto, son representativos del tipo de información a encontrar:

1. **Imagen Pozo GT-37:** Entre los 291 y 295 mt. de profundidad. De alta calidad por contener un alto contraste entre el fondo y las curvas buscadas.
2. **Imagen Pozo GT-39:** Entre los 320 y 324 mt. de profundidad. De baja calidad por contener una gran cantidad de onda superpuestas y con alto ruido de fondo.

En ambos casos fue necesario un proceso de preparación de la imagen, antes de pasar por el algoritmo de clasificación. Se siguió el proceso propuesto en [18] debido a los resultados obtenidos:

1. **Contraste:** Se utilizó ecualización de histograma, pero de manera local, es decir, en cada ventana deslizante utilizada. El propósito de este paso es ayudar al proceso de binarización.
2. **Binarización/Detección de Bordes:** Se investigaron varias formas de realizar la binarización de la imagen: porcentual propuesto en [18], la identificación de un valor de corte por medio de OTSU (global y local) [24], el método adaptativo de Chan [6] y *adaptive thresholding* incluido en scikit-image [29]. Se utilizó además el operador de gradiente Sobel en unión con

el filtro Canny para detección de bordes, sin embargo, luego de ajustar los parámetros de los mismos, recién se lograron resultados satisfactorios.

3. **Eliminación de Ruido de Fondo:** Para la eliminación del ruido de fondo se probaron varias técnicas: filtro de medianas, filtros morfológicos (Erosión y Dilatación) y filtros de ruido gaussiano.

En la figura 6.8 y 6.9 se puede ver el resultado de aplicar el proceso de preparación de la imagen. De acuerdo a esta información, se podría afirmar que el proceso de clasificación aplicado sobre el trozo GT-37 debería ser más preciso que sobre el trozo GT-39, en particular por la gran cantidad de ruido que posee este segundo.

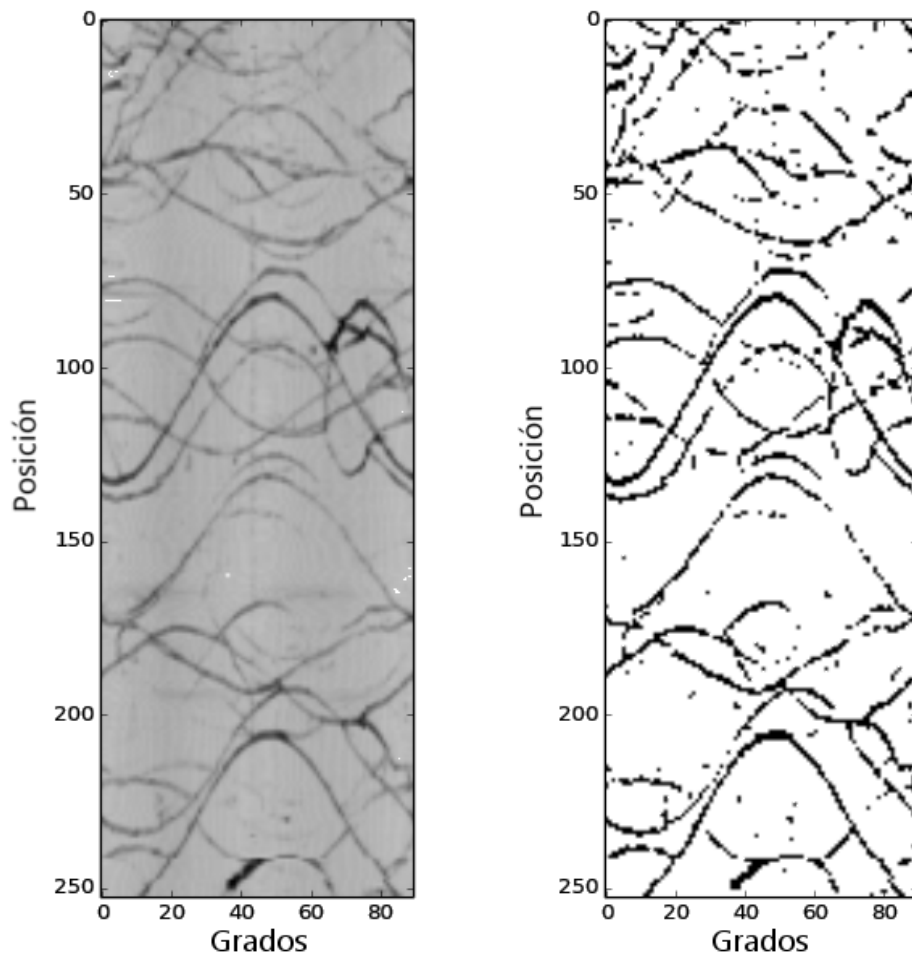


Figura 6.8: Resultado de aplicar el proceso de limpieza de imagen al trozo GT-37.

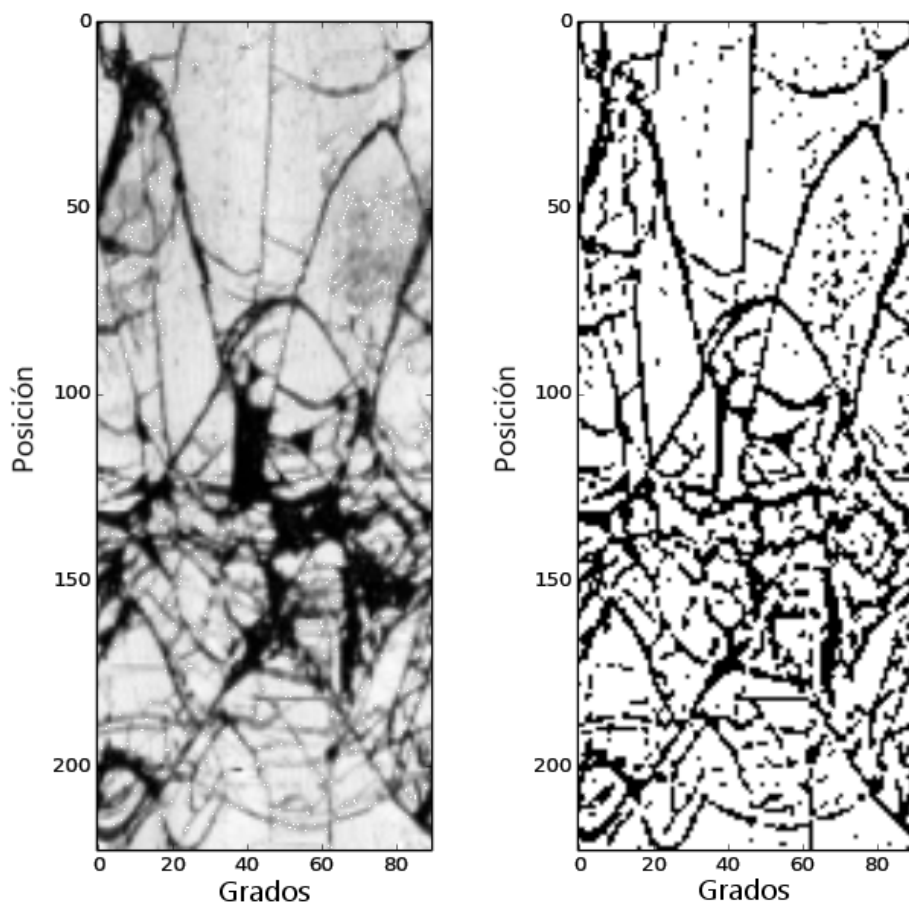


Figura 6.9: Resultado de aplicar el proceso de limpieza de imagen al trozo GT-39.

6.2.1.1. Resultados

Los resultados de la clasificación sobre los dos trozos se pueden visualizar en la figura 6.10 y 6.11. Las líneas horizontales indican la línea base encontrada y la línea continua la curva obtenida con los parámetros de amplitud y fase. Los valores de la métrica de calidad y error se pueden en en la tabla 6.6.

En el caso del trozo GT-37 se puede ver que el algoritmo de identificación de líneas base logró encontrar 7 de las 10 posibles líneas base, sin embargo,

se equivocó en no encontrar otras líneas bases presentes y en el valor de los parámetros de amplitud y fase.

En el trozo GT-39 a pesar de tener una mayor cantidad de ruido y mayor cantidad de fracturas sobrelapadas, logró un valor de la métrica de calidad $ExLin = 100\%$, sin embargo, el error de fase es notoriamente mayor. No es de extrañar este resultado, debido a que al aumentar la densidad de fracturas, la matriz de votación 2-D empieza a tener varios máximos locales y por lo tanto distintos valores de fase.

El valor de $ExLin$ obtenido por el trozo GT-39 sugiere que el algoritmo se comportó excelente en identificar las líneas base, sin embargo, es importante mencionar que además el algoritmo identificó más líneas base incorrectas, es decir, donde no existían fracturas.

Tabla 6.6: Resultado del proceso de identificación en términos de la métrica de calidad, error de amplitud y fase a dos trozos de imágenes acústicas de pozos reales.

Trozo	ExLin	e_A	e_φ
GT-37	70 %	32.484	13.154
GT-39	100 %	13.548	33.840

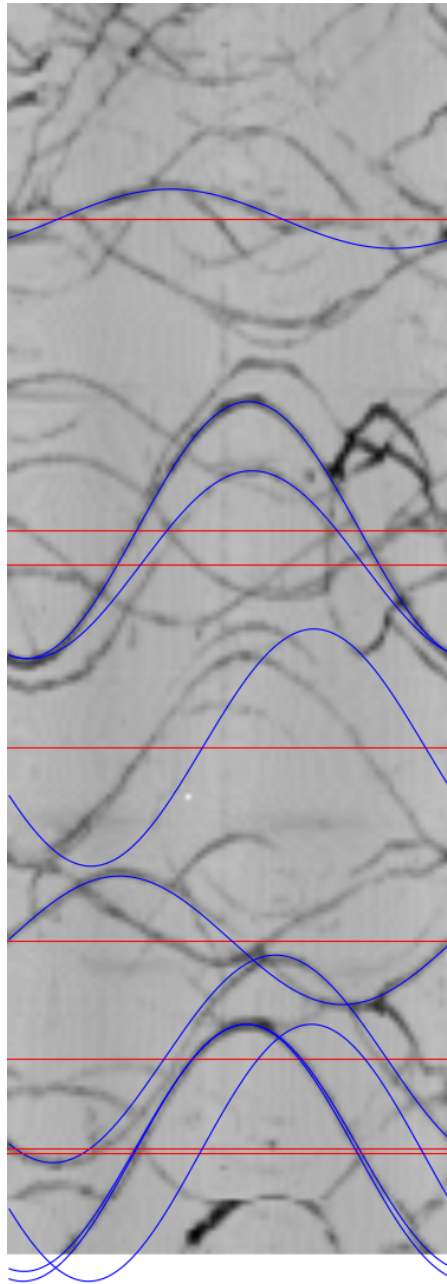


Figura 6.10: Resultados de la clasificación en el trozo GT-37. Se puede ver algunas de las curvas que el algoritmo no identificó, además de un pequeño error de fase.

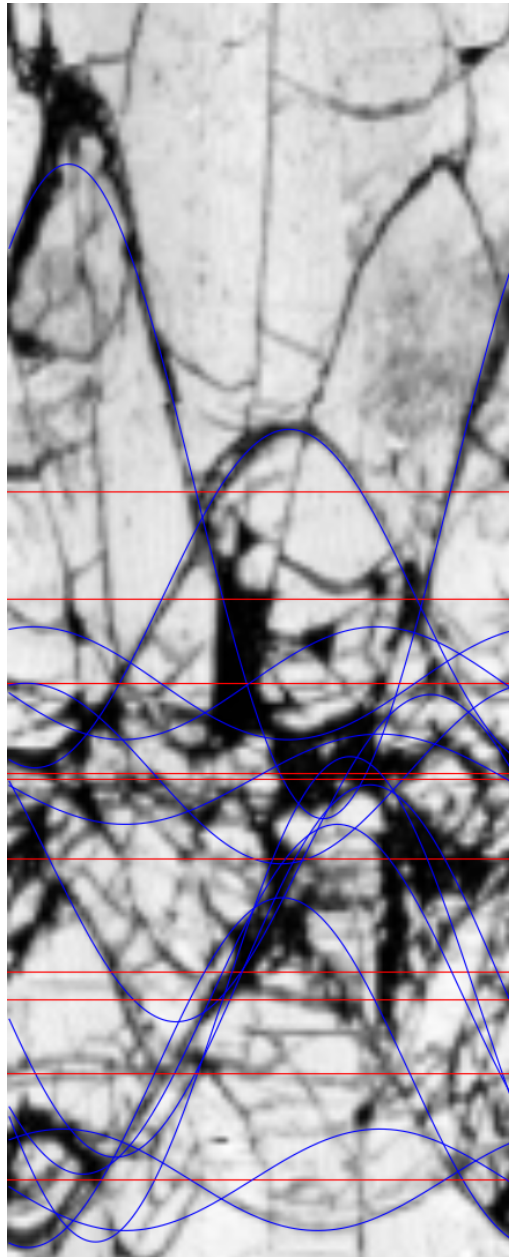


Figura 6.11: Resultados de la clasificación en el trozo GT-39. Se puede ver que el algoritmo realizó un buen trabajo encontrando las líneas bases y los parámetros, esto a pesar de la cantidad de ruido presente en la misma.

6.2.2. Identificación Sobre Imágenes Pre-Clasificadas

Se aplicó el algoritmo de clasificación a los tres tipos de imágenes clasificadas por el experto: fácil, medio y difícil.

- **Fácil:** Las curvas son fácilmente reconocibles, ya sea por el contraste con las paredes y/o por el espesor de la misma. El nivel de ruido presente en las imágenes permite un fácil reconocimiento de las fracturas. La cantidad de curvas sobrepuestas es nula o muy pequeña.
- **Medio:** Las curvas no son fácilmente reconocibles, ya sea por el nivel de ruido presente en la imagen o por encontrarse incompletas. Se debe realizar un proceso de prueba y error para ajustar la curva a los puntos visibles. Existen curvas sobrepuestas.
- **Difícil:** Es complejo identificar las curvas, existe bajo nivel de contraste entre la fractura y la pared de pozo, las fracturas se encuentra obstruidas por otras o la imagen posee errores asociado a un mal proceso de tomar la muestra. El número de curvas sobrepuestas es alto y con mucha variabilidad de amplitud.

A cada grupo se le aplicó el mismo procedimiento de limpieza visto en la sección anterior. En el caso del grupo difícil, fue necesario un paso más de limpieza, incluyendo un filtro pasa-bajos vertical para eliminar la des-centricidad de la probeta al momento de obtener a la muestra.

El experimento de clasificación se realizó sobre cada grupo de 15 imágenes, promediando el valor de la métrica de calidad y los errores de amplitud y fase obtenidos. Los resultados se pueden ver en la tabla [6.7](#).

Tabla 6.7: Resultado del proceso de identificación en términos de la métrica *ExLin*, error de amplitud y fase a los distintos grupos de imágenes clasificadas por el experto.

Complejidad	ExLin	e_A	e_φ
Fácil	66 %	15.202	11.818
Media	96 %	24.820	15.386
Difícil	23 %	40.178	19.948

En el caso de las imágenes de fácil clasificación, se puede ver que el rendimiento del algoritmo es similar al ser aplicado sobre el trozo GT-37. En el caso de las imágenes de complejidad media, el comportamiento también es similar al del grupo GT-39. En el caso del grupo difícil, la calidad es bastante menor, indicando que para este tipo de imágenes se hace necesario alguna modificación al algoritmo o al proceso de limpieza para mejorar la métrica de calidad.

6.3. Problemas Encontrados

En esta sección se discuten los principales problemas encontrados en el desarrollo de la aplicación del algoritmo de identificación y parametrización.

6.3.1. Semejanza por Vecindad

En el proceso de la modificación del algoritmo de identificación de líneas base, para incluir el concepto de vecindad y así descartar elementos que influyen en el comportamiento del algoritmo, como son el brillo o contraste, se encontraron resultados dispares: para el mismo conjunto de imágenes sintéticas, donde la precisión del algoritmo lograda es $ExLin = 97 \%$, con la modificación propuesta el resultado obtenido es $ExLin = 20 \%$.

A raíz del valor de *ExLin* obtenido, se realizó un análisis del algoritmo con el fin de identificar las razones de este comportamiento. Se identificaron dos problemas posibles: La implementación de la ventana móvil y la implementación de LBP utilizada.

6.3.1.1. Implementación de la ventana móvil

La implementación de la ventana móvil (*sliding window*) para comparar los *patches* o texturas de las imágenes, se basa en los ejemplos de `itertools` de Python. La implementación utilizada se puede ver en el código 6.1.

Código Fuente 6.1: *Sliding Window*

```
1 from itertools import islice
2
3 def window(seq, n=2):
4     "Returns a sliding window (of width n) over data from the iterable"
5     " s -> (s0,s1,...s[n-1]), (s1,s2,...,sn), ..."
6     it = iter(seq)
7     result = tuple(islice(it, n))
8     if len(result) == n:
9         yield result
10    for elem in it:
11        result = result[1:] + (elem,)
12        yield result
```

Al analizar la forma en que se mueve la ventana, se sospecha que la versión implementada no lo hace de acuerdo al algoritmo planteado.

6.3.1.2. Implementación de LBP

Se utilizó la librería de Python Mahotas [8], la cual posee ya la implementación de LBP¹ basada en [23]. La implementación declara que el vector retornado,

¹<http://mahotas.readthedocs.org/en/latest/lbp.html>

es el histograma del número de características encontradas. Al intentar aplicar la semejanza vía el ángulo que forman estos dos vectores, los valores son dispares, indicando que se debe realizar una implementación propia del mismo algoritmo. La implementación local de LBP tampoco mejoró la calidad de la clasificación.

6.3.2. Sensibilidad de *peakdetect*

Un elemento clave del algoritmo de identificación de líneas base, es identificar los *peaks* más importantes en el vector de votación 1-D. Esta identificación es realizada vía la librería *peakdetect*, la cual requiere dos parámetros muy sensibles en la identificación: **lookahead** y **delta**. Si sus valores son elegidos incorrectamente, la métrica de calidad del algoritmo es baja, particularmente por la forma en que identifica los valores máximos presentes.

En la implementación planteada, estos valores fueron calculados de manera empírica en $lookahead = 50$ y $delta = max/4$, donde max es el valor del máximo *peak* encontrado en el vector.

Capítulo 7

Conclusiones y Trabajo Futuro

A continuación se describen los resultados obtenidos, conclusiones y trabajo futuro con respecto al algoritmo y las distintas técnicas utilizadas.

7.1. Resultados Obtenidos

A pesar que la calidad de la identificación y parametrización de fracturas obtenida, no es comparable a la realizada manualmente por un humano, la aplicación del método propuesto muestra ser una opción viable para un pre-proceso de identificación de fracturas sobre imágenes de pozos obtenidas vía televisión acústica.

7.1.1. Limitaciones

Se debe tener especial cuidado en los parámetros que utiliza el algoritmo, tales como: altura de ventana de proceso, altura de ventana de análisis y valores de binarización. Ésto debido a la sensibilidad que posee la calidad de la clasificación asociada a los valores de estos parámetros.

7.2. Algoritmo

A partir de los resultados obtenidos en la identificación y parametrización de las curvas sinusoidales generadas de manera sintética, se puede deducir que el

algoritmo realiza una correcta identificación de la línea base y la parametrización de la onda.

7.2.1. Mejora del Algoritmo

Las mejoras introducidas al algoritmo propuesto por [7] realizado en el presente trabajo, se puede dividir en tres elementos claves:

1. **Identificación del tamaño de la ventana:** Para descartar falsos positivos, se propone utilizar como tamaño de ventana el doble de la amplitud máxima de una senoide a identificar. Esto con el fin de no procesar votos asociados a ondas que no pertenezcan a una misma línea base.
2. **Identificación de los peaks en el espacio de votación 1-D:** La mejora del algoritmo en este paso, implicó la utilización del método **peakdet**, sin embargo, a partir del estudio realizado por Yang [35] y Du [10] en un problema similar: búsqueda de peaks en señales de espectrógrafos de masa, se propone utilizar la **Transformada Wavelet Continua**, de la cual ya existe una implementación en `scipy.signal.cwt`.
3. **Distribución de la votación en el espacio de votación 2-D:** En el caso del espacio de votación 2-D la mejora pasó por descartar los *peaks* locales que se alejan mucho de un *peak* global. Además, vía un análisis de la distribución de los puntos pertenecientes a una curva, se puede descartar más votos que no se acerquen a ese valor. En este caso, una posible mejora sería identificar los *peaks* vía interpolación cuadrática y aplicar algunas de las técnicas presentadas en [35].

7.2.2. Limitaciones

El algoritmo y las mejoras propuestas poseen resistencia al ruido a partir de los resultados obtenidos, sin embargo, para niveles de ruido altos y/o superposición de gran número de curvas, el algoritmo obtiene valores bajos de la métricas de calidad.

7.3. Proceso de Limpieza y Binarizado de la Imagen

Debido a que el algoritmo utilizado realiza el proceso de votación identificando de forma binaria si un pixel pertenece o no a una onda, la calidad de la binarización de la imagen se encuentra directamente relacionada con el rendimiento del algoritmo.

Este procedimiento es crítico para el algoritmo, debido a ello y en relación a que las imágenes representan las paredes de un pozo, donde las ondas encontradas podrían no estar relacionadas entre sí; se propone un método de binarización local, el cual puede ser mejorado incluyendo en la fase previa mejoras en el brillo y contraste.

7.3.1. Limitaciones

El proceso de limpieza y binarizado necesariamente descarta información, que para algunos tipos de pozos o fracturas es necesario mantener, por ejemplo, con fracturas de ancho variable.

7.4. Utilización de Información Local

En la fase de binarización de la imagen, necesario para el algoritmo, se puede identificar un problema que podría ser clave para la mejora de la calidad de

identificación y parametrización del algoritmo:

Al binarizar la imagen, se elimina parte de la información de vecindad de un punto, es decir, su textura. Debido a que los puntos de la senoide representan la intersección entre un plano y las paredes del pozo, sería esperable que cada uno de los puntos que pertenezcan a la onda, tengan más o menos la misma textura. La razón de ello es que la textura representa en este caso el material constituyente del tipo de roca que intersecta al pozo.

Al analizar las imágenes reales, se logra identificar la semejanza, por lo tanto, es posible modificar el algoritmo para contemplar esta información:

1. No realizar el proceso de binarización, pero sí uno de realce de brillo y contraste, esto con el fin de mejorar la diferencia entre distintas zonas de la imagen.
2. Modificar la búsqueda de puntos que pertenecen a una sinusoidal, por la búsqueda de zonas con cierto nivel de semejanza que pertenezcan a la onda. Para ello se propone cambiar la comparación punto a punto, por una realizada vecindad a vecindad.

7.4.1. Semejanza por Vecindad

Debido a los dispares resultados obtenidos vía la implementación de semejanza de vecindad por medio del valor del ángulo formado por los vectores LBP, se propone un estudio más detallado de posibles formas de implementar semejanza de texturas, incluyendo elementos importantes a considerar como: cambios de iluminación, distorsión, cambios de perspectiva o rotación. Justamente elementos que abundan en las zonas que pertenecen a un fractura.

7.4.2. Limitaciones

La utilización de información local necesariamente implica definir un tamaño de vecindad a analizar. En el caso de las diferencias entre pozos, este valor puede ser clave para mejorar la calidad de clasificación del algoritmo. A raíz de ello debería ser un valor que el analista de datos debería poder modificar a voluntad.

7.5. Métricas

Se ha presentado una métrica de medición de calidad de la identificación de líneas base, la cual pone énfasis en los falsos positivos (líneas base no identificadas) sobre los falsos negativos, particularmente porque el error de no identificar una línea base es más importante que encontrar una donde realmente no existe.

Las métricas para medir el error de amplitud y fase representan la calidad de la parametrización, sin embargo, al ser valores asociados al alto de la onda o el número de grados de desfase, es difícil utilizarlos como medida de comparación directa contra otros algoritmos bajo distintos tipos de fracturas.

7.6. Materiales Utilizados

7.6.1. Conjunto de Imágenes

Las imágenes utilizadas representan pozos de la II región, que poseen cierta composición en común. A raíz de ello se hace difícil asegurar que el comportamiento del algoritmo y sus mejoras sean aplicables a otros tipos de pozos.

7.6.2. Lenguaje de Programación

La utilización de *Python* como lenguaje para implementar el algoritmo y sus mejoras, lleva asociado características positivas desde el punto de vista de la facilidad de implementación, sin embargo, el mismo algoritmo re-escrito en C++ debería suponer mejoras de rendimiento en velocidad y utilización de CPU.

Como IDE de desarrollo se utilizó *PyCharm 4.0*¹, herramienta de desarrollo de software que facilitó el desarrollo en el lenguaje Python.

7.6.3. WellCAD

WellCAD como herramienta de visualización, análisis y anotación es de gran utilidad para los analistas de fracturas. Gracias a contar con la posibilidad de automatizar las tareas de proceso de datos por medio de script de VisualBasic, es posible permitir que el algoritmo de identificación y clasificación se comunique directamente con WellCAD, ésto con el fin de leer datos de los pozos y publicar las fracturas encontradas directamente a la interfaz gráfica utilizada por el analista.

¹<https://www.jetbrains.com/pycharm/>

Bibliografía

- [1] Said Assous, Peter Elkington, Stuart Clark, and James Whetton. Automated detection of planar geologic features in borehole images. *Geophysics*, 79(1):D11–D19, 2014.
- [2] Colleen A Barton and Mark D Zoback. Stress perturbations associated with active faults penetrated by boreholes: Possible evidence for near-complete stress drop and a new technique for stress magnitude measurement. *Journal of Geophysical Research: Solid Earth (1978–2012)*, 99(B5):9373–9390, 1994.
- [3] Eli Billauer. Peak detection using matlab. <http://billauer.co.il/peakdet.html>, 2013.
- [4] John Canny. A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):679–698, 1986.
- [5] John Francis Canny. Finding edges and lines in images. *Massachusetts Inst. of Tech. Report*, 1, 1983.
- [6] Francis HY Chan, Francis K Lam, and Hui Zhu. Adaptive thresholding by variational method. *Image Processing, IEEE Transactions on*, 7(3):468–473, 1998.
- [7] Zou Changchun and Shi Ge. A hough transform-based method for fast detection of fixed period sinusoidal curves in images. In *Signal Processing, 2002 6th International Conference on*, volume 1, pages 909–912. IEEE, 2002.

- [8] Luis Pedro Coelho. Mahotas: Open source software for scriptable computer vision. *arXiv preprint arXiv:1211.4907*, 2012.
- [9] SR Deans. Radon transform and its applications. *Interscience, New York*, 1983.
- [10] Pan Du, Warren A Kibbe, and Simon M Lin. Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching. *Bioinformatics*, 22(17):2059–2065, 2006.
- [11] Richard O Duda and Peter E Hart. Use of the hough transformation to detect lines and curves in pictures. *Communications of the ACM*, 15(1):11–15, 1972.
- [12] Electromind. Acoustic borehole televiewer abi42. http://www.electromind.eu/products.html#anchor-a_televiewer, 2013.
- [13] Ann M. Fugl-Meyer. Optical and acoustic televiewer. Master’s thesis, Technical University of Denmark, 2007.
- [14] Robertson Geologging. Borehole, slimhole and wireline logging equipment. <http://www.geologging.com>, 2013.
- [15] Bryn Hubbard, Sam Roberson, Denis Samyn, and Derek Merton-Lyn. Digital optical televiewing of ice boreholes. *Journal of Glaciology*, 54(188):823–830, 2008-12-10T00:00:00.
- [16] Bryn Hubbard, Jean-Louis Tison, Frank Pattyn, Marie Dierckx, Thierry Boereboom, and Denis Samyn. Optical-televiewer-based identification and characterization of material facies associated with an antarctic ice-shelf rift. *Annals of Glaciology*, 53(60):137–146, 2012-11-01T00:00:00.

- [17] Marek Jarosinski. Contemporary stress field distortion in the polish part of the western outer carpathians and their basement. *Tectonophysics*, 297(14):91 – 119, 1998.
- [18] Fabiana Rodrigues Leta, Esteban Clua, Mauro Biondi, Toni Pacheco, and Maria do Socorro de Souza. An automatic process to identify features on boreholes data by image processing techniques. In *Experimental and Numerical Investigation of Advanced Materials and Structures*, pages 249–262. Springer, 2013.
- [19] Tony Lindeberg. Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):117–156, 1998.
- [20] G Louloudis, B Gatos, I Pratikakis, K Halatsis, et al. A block-based hough transform mapping for text line detection in handwritten documents. In *Tenth International Workshop on Frontiers in Handwriting Recognition*, 2006.
- [21] Stepan M Luthi and Philippe Souhaite. Fracture apertures from electrical borehole scans. *Geophysics*, 55(7):821–833, 1990.
- [22] Terry Malone, Bryn Hubbard, Derek Merton-Lyn, Paul Worthington, and Reyer Zwiggelaar. Borehole and ice feature annotation tool (bifat): A program for the automatic and manual annotation of glacier borehole images. *Computers & Geosciences*, 2012.
- [23] Timo Ojala, Matti Pietikäinen, and Topi Mäenpää. Gray scale and rotation invariant texture classification with local binary patterns. In *Computer Vision-ECCV 2000*, pages 404–420. Springer, 2000.
- [24] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *Automatica*, 11(285-296):23–27, 1975.

- [25] Karthik Raghupathy. *Curve tracing and curve detection in images*. PhD thesis, Cornell University, 2004.
- [26] M.H. Rider and M. Kennedy. *The geological interpretation of well logs*. Rider-French, 2011.
- [27] Sam Roberson and Bryn Hubbard. Application of borehole optical televiewing to investigating the 3-d structure of glaciers: implications for the formation of longitudinal debris ridges, midre lovénbreen, svalbard. *Journal of Glaciology*, 56(195):143–156, 2010-04-01T00:00:00.
- [28] Andes Geofísica S.A. Página web corporativa. <http://www.andesgeofisica.cl/>, March 2013.
- [29] scikit-image development team. Adaptive thresholding. <http://scikit-image.org/>, 2013.
- [30] ABIM Solutions. Demo hole, acoustic televiewer interpretation, 2010.
- [31] Advanced Logic Technology. Wellcad. <http://www.alt.lu/wellcad.htm>, April 2013.
- [32] Advanced Logic Technology. Wellcad module image. http://www.alt.lu/pdf/Image_Module.pdf, April 2013.
- [33] Peter Aundal Toft and John Aasted Sørensen. *The radon transform-theory and implementation*. PhD thesis, Technical University of Denmark Danmarks Tekniske Universitet, Department of Information Technology Institut for Informationsteknologi, 1996.
- [34] John H. Williams and Carole D. Johnson. Acoustic and optical borehole-wall imaging for fractured-rock aquifer studies. *Journal of Applied Geophysics*, 55(1), 2004. Non-Petroleum Applications of Borehole Geophysics.

- [35] Chao Yang, Zengyou He, and Weichuan Yu. Comparison of public peak detection algorithms for maldi mass spectrometry data analysis. *BMC bioinformatics*, 10(1):4, 2009.

Apéndice A

Formato LGX

[General]

Version=101

The LGX format is only implemented in files with Version=101 or greater.

Previous versions have log data in previous RG log type formats (PCL2 compatible).

DataFiles=n

To cope with sondes which acquire vector (waveform) data at a different sample interval from normal scalar data (e.g. Full-waveform sonic), it is necessary to record a separate file for the different data types. This value determines the number of File blocks which will appear.

[Filen]

Each file block will describe the file name, data recording units, logging direction, start and end depth etc. A couple of headings require amplification here, when the string value is an enumeration

ReferenceUnit=nn

Metres=0, Feet=1, Seconds=2

Mode=nn

Log modes can be TIME=10, UP=11, DOWN=12, ASYNC=13, TOWARDS=14, AWAY=15.

TOWARDS and AWAY modes are used by sondes that are propelled in sub-horizontal boreholes by the semi-rigid cable system. Logs in these modes can only be made by the OPTV and HiRAT at this time.

Interval=nn

The sample interval is always expressed as an integer value whose base unit is 1/1000 of the ReferenceUnit (see above) i.e. millimetres, millifeet, milliseconds. Expect the OPTV log sample interval to be 1 or 2mm for metric and 3 or 6 mft for imperial (may change to 4/8mft in future).

RecLineSize=nnnn

This field describes the length of the records in the data file. The position of the individual channels, and the method of decoding them is given in the Channel sections - see below.

[WellInfo]

Well information contains the normal descriptive information entered by the logging engineer. Some fields from the [File] blocks may be repeated here for compatibility reasons. All fields are adequately described by their profile string names, and no field is required to understand the data encoding in the log file.

[Channeln]

Each data channel will have its own block. It is the key to reading the data and its presentation. The profile strings essential to reading the data will be described. Others speak for themselves, and may relate only to the WinLogger presentation. There is no requirement to specify presentation details, since the format is prescriptive only for data exchange purposes, but the WinLogger application will use the profile strings to control its log presentation in replay.

It is expected that the first channel will be depth (or time) but this is not mandatory. However, when it does appear IT WILL BE DESCRIBED AS A DATA CHANNEL.

File=nn

For datasets which contain multiple files, this entry shows from which file the data must be drawn. It will refer to the [Filen] group as described above.

DataOffset=nn

Position (bytes) within the data record at which the data will be found.

SampleType=nn

Encoding details for the data object.

char = 0, unsigned char=1, short integer = 2, unsigned short integer =3, long integer = 4, unsigned long integer = 5, int64 = 6, unsigned int64 = 7, float = 8, double = 9, RGB (3 bytes) = 10, Amp/TT doublet (BHTV) = 11, Amp/TT/Gain triplet (HRAT) = 12.

SampleCount=nn

The number of data items which make a sample, for scalars SampleCount=1, otherwise it will contain the vector length.

[Trackn]

These entry blocks will refer only to the WinLogger presentation and may be ignored.

Imagery from the High Resolution Acoustic Televiewer is sampled in the time domain, but with depth as an attribute - asynchronous mode - as is the case with the earlier BHTV, whose data is not yet recorded in LGX format. However, the data has programmable radial resolution of 1,2 or 4 degree steps under surface control. Logs will not change their radial resolution.

The data from the HRAT includes raw orientation vectors. The data channels will be:

HX GX HY GY HZ GZ.

where H refers to magnetic data and G refers to accelerometer (gravity) data.

See the Applied Physics Systems Model 544 Miniature Orientation Sensor technical reference for full details of axis conventions and terminology.

The acoustic data is recorded under the channel name Scan. The profile string ImageAligned indicates the data alignment. With ImageAligned=1 the first sample will be orientated to North. When ImageAligned=0 the data is not aligned (like data from the MW BHTV) and the vectors must be used to obtain the sonde orientation. Image alignment would normally only be disabled for inspection of magnetic casing or steel-reinforced piles. When ImageAligned=1 there is an additional parameter that indicates whether correction for magnetic declination has also been applied: MagneticDeclination=nn. When the value is zero, no correction has been applied.

The number of samples is variable (see above) between logs, and the SampleCount value will give the number of samples per rotation. The data is encoded as sample type 12 which comprises:

TransitTime SampleType = 3 bits 0-15 (numbered from left) Units 100nSec

Amplitude SampleType = 1 bits 16-23

Gain SampleType = 1 bits 24-31

Depth and Time are available in every record. Note that the asynchronous mode will result in multiple records having the same depth value. There is no constraint upon cable direction, and the file may contain 'yo-yo' features where actual cable direction has reversed twice before the original depth is regained. True log direction can therefore be inferred only from the start and end depths in the [Filen] block (repeated in Wellinfo for compatibility).