PONTIFICIA UNIVERSIDAD CATÓLICA DE VALPARAÍSO

FACULTAD DE INGENIERÍA

ESCUELA DE INGENIERÍA INFORMÁTICA

**SISTEMA DE RECONOCIMIENTO GESTUAL DEL ALFABETO LENGUA DE SEÑAS CHILENA MEDIANTE CAMARA DIGITAL**

**CARLOS GUILLERMO GONZÁLEZ RIVEROS**

**FRANCISCO JAVIER YIMES INOSTROZA**

INFORME DE AVANCE PROYECTO 1

PARA OPTAR AL TITULO PROFESIONAL DE

INGENIERO DE EJECUCIÓN EN INFORMÁTICA

OCTUBRE 2015

Pontificia Universidad Católica de Valparaíso

Facultad de Ingeniería

Escuela de Ingeniería Informática

**SISTEMA DE RECONOCIMIENTO GESTUAL DEL ALFABETO LENGUA DE SEÑAS CHILENA MEDIANTE CAMARA DIGITAL**

**CARLOS GUILLERMO GONZÁLEZ RIVEROS**

**FRANCISCO JAVIER YIMES INOSTROZA**

Profesor Guía: **Claudio Cubillos Figueroa**

Profesor Co-referente: **Iván Mercado Bermúdez**

Carrera: **Ingeniería de Ejecución Informática**

Octubre 2015

**Resumen**

**Abstract**

**Índice**

[Glosario de Términos v](#_Toc433133707)

[Lista de Abreviaturas o Siglas vi](#_Toc433133708)

[Lista de Figuras vii](#_Toc433133709)

[Lista de Tablas viii](#_Toc433133710)

[Introducción 1](#_Toc433133711)

[1 Planteamiento del Tema 2](#_Toc433133712)

[1.1 Motivación 2](#_Toc433133713)

[1.2 Definición de Objetivos 2](#_Toc433133714)

[1.2.1 Objetivo general 2](#_Toc433133715)

[1.2.2 Objetivos específicos 2](#_Toc433133716)

[1.3 Planificación del Proyecto 3](#_Toc433133717)

[1.4 Estructura del Documento 5](#_Toc433133718)

[1.5 Metodología de Desarrollo 5](#_Toc433133719)

[1.5.1 Metodología Rapid Application Development 5](#_Toc433133720)

[2 Estado del Arte 7](#_Toc433133721)

[2.1 Reconocimiento Gestual 7](#_Toc433133722)

[2.1.1 Clasificación 7](#_Toc433133723)

[2.1.2 Dificultades 8](#_Toc433133724)

[2.1.3 Aplicaciones 8](#_Toc433133725)

[2.1.4 Herramientas y/o Técnicas 8](#_Toc433133726)

[2.1.4.1 Modelos Ocultos de Markov (HMM) 8](#_Toc433133727)

[2.1.4.2 Aproximación por máquinas de Estados Finitos (FMS) 8](#_Toc433133728)

[2.1.4.3 Máquinas de Vectores de Soporte (SVM) 8](#_Toc433133729)

[2.1.4.4 Otras Herramientas 8](#_Toc433133730)

[2.2 Modelado de Gestos 8](#_Toc433133731)

[2.2.1 Taxonomía (Clasificación) Gestual 8](#_Toc433133732)

[2.2.2 Modelado Temporal de los Gestos 8](#_Toc433133733)

[2.2.3 Modelado Espacial de los Gestos 8](#_Toc433133734)

[2.3 Análisis de Gestos 8](#_Toc433133735)

[2.3.1 Detección de Características 8](#_Toc433133736)

[2.3.2 Estimación de Parámetros 8](#_Toc433133737)

[2.3.3 Detector Basado en Características Haar-Like 8](#_Toc433133738)

[2.3.4 Descriptores de Histogramas de Gradientes Orientados 8](#_Toc433133739)

[2.4 Sistemas de Reconocimiento Gestual 8](#_Toc433133740)

[2.4.1 LeapMotion 8](#_Toc433133741)

[2.4.2 Wisee 8](#_Toc433133742)

[2.4.3 Kinect 8](#_Toc433133743)

[2.4.4 Air Gesture 8](#_Toc433133744)

[2.4.5 Openarch 8](#_Toc433133745)

[2.5 Alternativas de Desarrollo Gestual 8](#_Toc433133746)

[2.5.1 AForge.NET 8](#_Toc433133747)

[2.5.2 OpenCV 8](#_Toc433133748)

[2.5.3 JMF 8](#_Toc433133749)

[2.5.4 Motion.py 8](#_Toc433133750)

[2.5.5 SoftCollection VideoMotionDetection 8](#_Toc433133751)

[2.5.6 VisionLab VCL 8](#_Toc433133752)

[3 Marco Teórico 9](#_Toc433133753)

[3.1 Lengua de Señas 9](#_Toc433133754)

[3.2 Imágenes Digitales 10](#_Toc433133755)

[3.2.1 Tipo de Imágenes Digitales 10](#_Toc433133756)

[3.2.1.1 Imágenes de Mapa Bits 10](#_Toc433133757)

[3.2.1.2 Imágenes Vectoriales 10](#_Toc433133758)

[3.2.2 Colores 11](#_Toc433133759)

[3.2.2.1 Espacio RGB 11](#_Toc433133760)

[3.2.2.2 Espacio HSI 12](#_Toc433133761)

[3.3 Procesamiento de Imágenes 13](#_Toc433133762)

[3.3.1 Manipulación de Contraste 13](#_Toc433133763)

[3.3.2 Eliminación de Ruido 13](#_Toc433133764)

[3.3.3 Realce de Bordes 14](#_Toc433133765)

[3.3.3.1 Método de Desplazamiento y Sustracción 15](#_Toc433133766)

[3.4 Extracción de Características 15](#_Toc433133767)

[3.4.1 Detección de Bordes 15](#_Toc433133768)

[3.4.1.1 Técnicas Basadas en el Gradiente 15](#_Toc433133769)

[3.4.1.2 Operadores de Segundo Orden 15](#_Toc433133770)

[3.4.1.3 Métodos basados en las Derivadas de Gaussianas 15](#_Toc433133771)

[3.4.1.4 Operadores de Ajuste al Modelo 15](#_Toc433133772)

[3.4.1.5 Detección de Líneas 15](#_Toc433133773)

[3.4.1.6 Detección de Esquinas 15](#_Toc433133774)

[3.4.2 Texturas 15](#_Toc433133775)

[3.4.2.1 Análisis Estadístico de Texturas 15](#_Toc433133776)

[3.4.2.2 Análisis Frecuencial de Texturas 15](#_Toc433133777)

[3.4.3 Detección de Movimiento 15](#_Toc433133778)

[3.5 Segmentación de Imágenes 15](#_Toc433133779)

[3.5.1 Segmentación basada en la detección de bordes 15](#_Toc433133780)

[3.5.1.1 Método basado en el gradiente 15](#_Toc433133781)

[3.5.1.2 La transformada de Hough 15](#_Toc433133782)

[3.5.2 Segmentación de Regiones 15](#_Toc433133783)

[3.5.2.1 Umbralización 15](#_Toc433133784)

[3.5.2.2 Etiquetado 15](#_Toc433133785)

[3.5.2.3 Crecimiento de Regiones 15](#_Toc433133786)

[3.5.2.4 División y Unión de Regiones 15](#_Toc433133787)

[4 Conclusiones 16](#_Toc433133788)

[5 Referencias Bibliográficas 17](#_Toc433133789)

# Glosario de Términos

# Lista de Abreviaturas o Siglas

# Lista de Figuras

[Figura 1.1 Metodología RAD [Elab. Propia] 5](file:///C:\Users\CarlosGuillermo\Desktop\Proyecto-1-Avance-Carlos-G.-Francisco-Y.%20Versión%202.0.0.docx#_Toc433133697)

[Figura 1.2 Metodología Clásica de Desarrollo de Software. [Elab. Propia] 6](file:///C:\Users\CarlosGuillermo\Desktop\Proyecto-1-Avance-Carlos-G.-Francisco-Y.%20Versión%202.0.0.docx#_Toc433133698)

[Figura 2.1 Ejemplo de Reconocimiento Gestual. 7](file:///C:\Users\CarlosGuillermo\Desktop\Proyecto-1-Avance-Carlos-G.-Francisco-Y.%20Versión%202.0.0.docx#_Toc433133699)

[Figura 3.1 Alfabeto de la Lengua de Señas Chilena. [10] 9](file:///C:\Users\CarlosGuillermo\Desktop\Proyecto-1-Avance-Carlos-G.-Francisco-Y.%20Versión%202.0.0.docx#_Toc433133700)

[Figura 3.2 Dos rejillas de 5x5 píxeles. [5] 10](file:///C:\Users\CarlosGuillermo\Desktop\Proyecto-1-Avance-Carlos-G.-Francisco-Y.%20Versión%202.0.0.docx#_Toc433133701)

[Figura 3.3 Representación tridimensional del espacio RGB. 12](file:///C:\Users\CarlosGuillermo\Desktop\Proyecto-1-Avance-Carlos-G.-Francisco-Y.%20Versión%202.0.0.docx#_Toc433133702)

[Figura 3.4 HSI: (a) Tono, (b) Saturación, (c) Intensidad. [6] 12](file:///C:\Users\CarlosGuillermo\Desktop\Proyecto-1-Avance-Carlos-G.-Francisco-Y.%20Versión%202.0.0.docx#_Toc433133703)

[Figura 3.5 (a) Imagen Original, (b) Aumento de Contraste. [2] 13](file:///C:\Users\CarlosGuillermo\Desktop\Proyecto-1-Avance-Carlos-G.-Francisco-Y.%20Versión%202.0.0.docx#_Toc433133704)

[Figura 3.6 Diferencia entre imagen sin ruido y con ruido. [9] 14](file:///C:\Users\CarlosGuillermo\Desktop\Proyecto-1-Avance-Carlos-G.-Francisco-Y.%20Versión%202.0.0.docx#_Toc433133705)

[Figura 3.7 Realce de bordes, (a) imagen original, (b) imagen con filtro gaussiano para eliminar ruido, (c) realce de bordes sin filtro gaussiano, (d) realce de bordes con filtro gaussiano. 14](file:///C:\Users\CarlosGuillermo\Desktop\Proyecto-1-Avance-Carlos-G.-Francisco-Y.%20Versión%202.0.0.docx#_Toc433133706)

# Lista de Tablas

[Tabla 2.1 Carta Gantt Primer Avance Proyecto 1. 4](file:///C:\Users\CarlosGuillermo\Desktop\Proyecto-1-Avance-Carlos-G.-Francisco-Y.%20Versión%202.0.0.docx#_Toc433132599)

# Introducción

# Planteamiento del Tema

## Motivación

El reconocimiento gestual es un problema que se encuentra en una etapa de investigación y desarrollo un tanto inmadura, pero algunos de los avances que se han logrado hoy en día han dado mucho de qué hablar, que decir con la llegada de la “Kinect”, la cual ha dado ese empuje al reconocimiento gestual teniendo mayor relevancia a la hora de la investigación sobre ésta temática, y se han construido importantes librerías de software tanto libres como privadas. Es así como nosotros esperamos constituir una base sólida de conocimientos para el posterior desarrollo de tecnologías que permitan utilizar reconocimiento gestual en diferentes ámbitos de progreso.

La motivación que nos lleva a la realización de este proyecto es el hecho de poder ayudar este sector de la población que tienen problemas para comunicarse y se ven un tanto excluidos por la sociedad.

La forma más común de poder interactuar con este grupo minoritario de incapacidad auditiva, pero no menor, es a través de cursos presenciales y/o cursos E-Learning. Dentro del grupo de personas con hipoacusia hay distintos grados de pérdida auditiva y en algunos casos las posibilidades de comunicación con su entorno se vuelve más complejo, peor aun cuando se nace con pérdida total, la comunicación se vuelve un problema mayor.

Mediante este proyecto se pretende aportar una contribución a los avances en este campo creando y desarrollando una nueva forma de aplicar el reconocimiento gestual a través una cámara digital e implementar un pequeño curso introductorio de la lengua de señas lo más didáctico y entretenido posible, todo esto por medio de dispositivos móviles Smartphone o Teléfonos inteligentes, que permitan que una persona con capacidad limitada y/o una persona de su entorno se comuniquen entre sí, sin la necesidad de saber la lengua de señas.

## Definición de Objetivos

### Objetivo general

Desarrollar un sistema de reconocimiento de señas, implementando algoritmos de procesamiento y segmentación de imágenes, pertenecientes a la lengua de señas chilena, a través de una cámara digital.

### Objetivos específicos

Los objetivos específicos del actual proyecto son:

* Investigar sobre sistemas y algoritmos de reconocimiento gestual ya existentes.
* Indagar sobre metodologías de enseñanza de lengua de señas.
* Determinar la factibilidad de desarrollo del proyecto de reconocimiento.
* Clasificar y determinar los algoritmos necesarios para el procesamiento de imágenes.
* Especificar y establecer la (s) librería (s) de desarrollo para el sistema de reconocimiento gestual.
* Proponer un algoritmo que permita detectar el movimiento de la(s) mano(s) dentro de la imagen.

## Planificación del Proyecto

El proceso de gestión de un proyecto de software comienza con un conjunto de actividades que, globalmente, se denominan planificación del proyecto. "No podemos pedir exactitud a la fase de planificación, es solo una idea de cómo van a transcurrir las cosas. Hay que planificar el trabajo, los recursos humanos y la tecnología. Planificar es estimar."

Es por eso que hemos desarrollado una carta Gantt para la correcta elaboración de este primer avance de proyecto, en donde se definieron a grandes rasgos las tareas que se realizaron, no se entró en detalles de la investigación, se tomaron los puntos más relevantes e importantes a desarrollarse a lo largo de este primer avance durante mes y medio de trabajo.

El presente proyecto se basa en dos partes:

1. Proyecto 1: En proyecto 1 existen dos partes de evaluación una presentación de avance y final donde en el avance participan el Profesor guía y el correferente, en la presentación final existe además un examinador invitado.
2. Proyecto 2: En el proyecto 2 existen tres etapas de evaluación avance, presentación del software y presentación final. Todas forman parte de una sola nota.

A continuación se presenta en la tabla 2.1 el plan de trabajo del presente primer avance de proyecto en donde se ve como primera etapa el análisis del problema en cuestión y la estructuración del documento entregable y posteriormente el desarrollo de este.

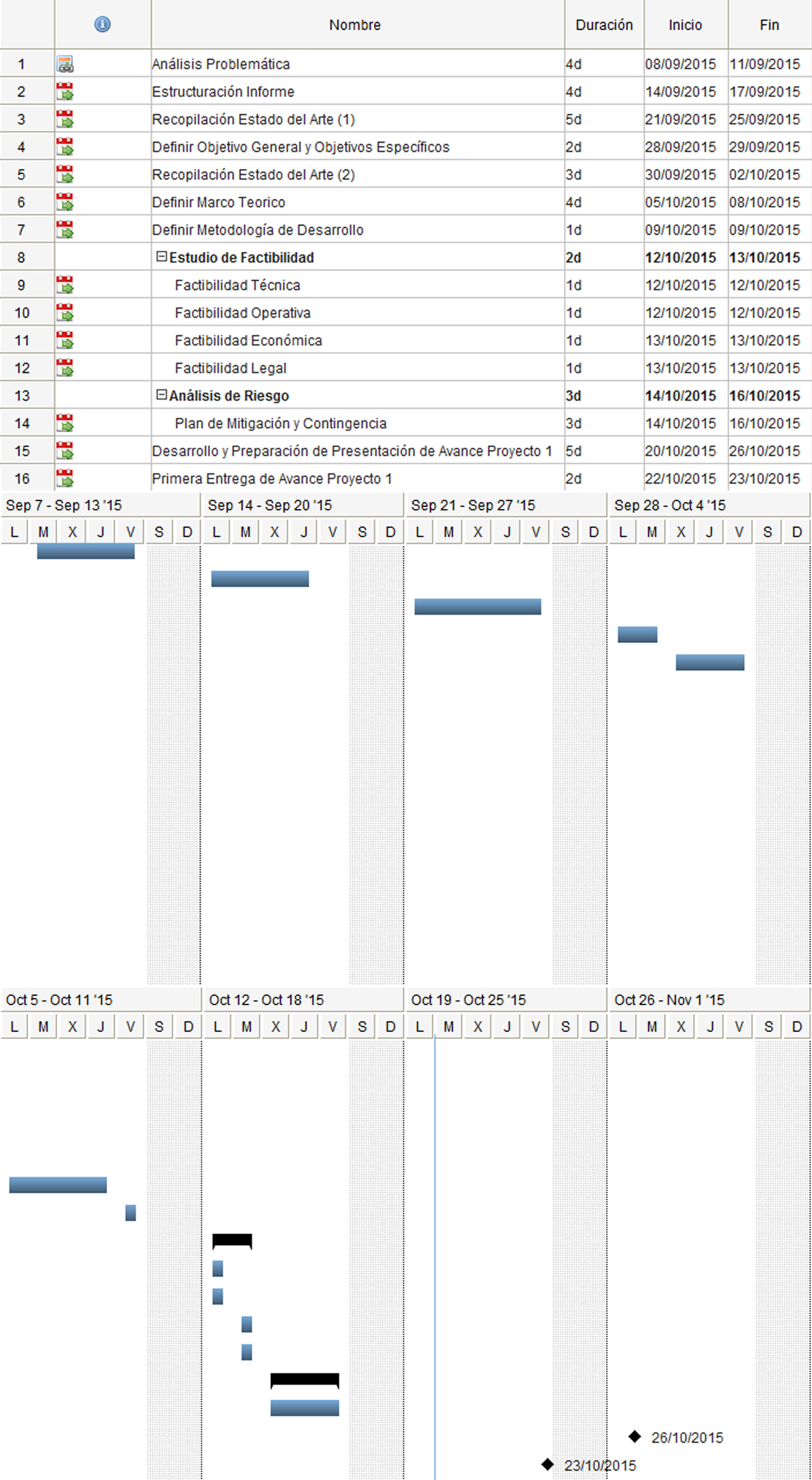


Tabla 2.1 Carta Gantt Primer Avance Proyecto 1.

## Estructura del Documento

El presente documento se divide en 4 capítulos, los cuales serán descritos de manera breve y precisa a continuación:

Capítulo 1, Planteamiento del tema: En este capítulo, se describirá la motivación, se establecerán y definirán objetivos, junto con la metodología de desarrollo, la planificación del proyecto y finalmente estudios de factibilidad.

Capítulo 2, Estado del arte: En este capítulo, se verán las tecnologías que ya se han desarrollado sobre reconocimiento gestual, se definirán el concepto de reconocimiento gestual y las alternativas que tendremos para el desarrollo de nuestro proyecto para su posterior implementación.

Capítulo 3, Marco teórico: En este capítulo, se podrán ve en primera instancia la definición de la lengua de señas chilenas, y se verán los principales aspectos del reconocimiento de imágenes digitales.

Capítulo 4, Conclusiones: En este capítulo, se presentaran las conclusiones a las que se ha llegado tras el marco de investigación de reconocimiento gestual, en donde se concluye también cual herramienta nos facilita mayor documentación.

## Metodología de Desarrollo

El desarrollo de software no es una tarea fácil. Prueba de ello es que existen numerosas propuestas metodológicas que inciden en distintas dimensiones del proceso de desarrollo. Por una parte tenemos aquellas propuestas más tradicionales que se centran especialmente en el control del proceso, estableciendo rigurosamente las actividades involucradas, los artefactos que se deben producir, y las herramientas y notaciones que se usarán. Estas propuestas han demostrado ser efectivas y necesarias en un gran número de proyectos, pero también han presentado problemas en muchos otros.

Una posible mejora es incluir en los procesos de desarrollo más actividades, más artefactos y más restricciones, basándose en los puntos débiles detectados. Sin embargo, el resultado final sería un proceso de desarrollo más complejo que puede incluso limitar la propia habilidad del equipo para llevar a cabo el proyecto.

### Metodología Rapid Application Development

El desarrollo rápido de aplicaciones (RAD) es una metodología de desarrollo de software, que implica el desarrollo iterativo y la construcción de prototipos. El desarrollo rápido de aplicaciones es un término originalmente utilizado para describir un proceso de desarrollo de software introducido por James Martin en 1991.

Figura 1.1 Metodología RAD [Elab. Propia]

Ciclo de Prototipos

Figura 1.2 Metodología Clásica de Desarrollo de Software. [Elab. Propia]

# Estado del Arte

## Reconocimiento Gestual

El gesto en sí es un movimiento corporal con significado, que puede ser realizado con los dedos, las manos, los brazos, la cabeza, con todo el tronco; o incluso con una variación de la expresión de la cara o mediante el movimiento de ojos. Por sí mismo constituye una parte importante de la comunicación natural cotidiana, en muchas ocasiones como refuerzo al lenguaje hablado y, cuando éste no es posible, como la primera alternativa de comunicación. [8]

El reconocimiento gestual es un tema de la visión por computador o bien de la visión artificial, área de la inteligencia artificial, además de la tecnología del lenguaje que tiene como objetivo reconocer e interpretar gestos y/o patrones humanos a través de algoritmos matemáticos aplicados a la imagen que se reconoce mediante ciertas técnicas de visión, por medio de dispositivos de entrada como cámaras, webcams, o dispositivos móviles.

Figura 2.1 Ejemplo de Reconocimiento Gestual.

Como se puede ver en la Figura 2.1, no solo se trata de reconocer manos o cara, sino toda actividad corporal que produzca un cambio significativo a la imagen, pero como nuestro proyecto está centrado en el reconocimiento del alfabeto de lengua de señas chilena, éste está enfocado en el rescate de la forma y movimiento de manos.

### Clasificación

Los gestos se pueden clasificar como gestos estáticos y dinámicos:

* Estáticos: El usuario asume cierta postura, o configuración de su cuerpo o de la cara, fija.
* Dinámicos: En los que la información esta codificada en el movimiento, a través de la evolución temporal de las posturas, o bien una mezcla de ambos, como, por ejemplo la lengua de signos. [8]

### Dificultades

### Aplicaciones

### Herramientas y/o Técnicas

#### Modelos Ocultos de Markov (HMM)

#### Aproximación por máquinas de Estados Finitos (FMS)

#### Máquinas de Vectores de Soporte (SVM)

#### Otras Herramientas

## Modelado de Gestos

### Taxonomía (Clasificación) Gestual

### Modelado Temporal de los Gestos

### Modelado Espacial de los Gestos

## Análisis de Gestos

### Detección de Características

### Estimación de Parámetros

### Detector Basado en Características Haar-Like

### Descriptores de Histogramas de Gradientes Orientados

## Sistemas de Reconocimiento Gestual

### LeapMotion

### Wisee

### Kinect

### Air Gesture

### Openarch

## Alternativas de Desarrollo Gestual

### AForge.NET

### OpenCV

### JMF

### Motion.py

### SoftCollection VideoMotionDetection

### VisionLab VCL

# Marco Teórico

## Lengua de Señas

La lengua de señas es la lengua natural de las personas sordas, se basa en movimientos y expresiones a través de ojos, rostro, boca, cuerpo y principalmente manos, siendo esta última nuestra principal base a estudiar.

Las personas sordas nacen con la capacidad biológica del lenguaje intacta. Sin embargo, su limitación auditiva impide que puedan apropiarse del idioma o lengua oral que se habla a su alrededor.

Al igual que los idiomas orales, existen diferentes lenguas de señas. Esto debido a que como ha pasado con los demás idiomas, las lenguas de señas han surgido de acuerdo a los procesos de educación o de influencia de unos países sobre otros al igual que a las características propias de cada región, así por ejemplo existe una lengua de señas colombiana, una lengua de señas argentina, una lengua de señas española, una lengua de señas americana, que no son iguales. Lo mismo que sucede con los idiomas orales, (español, francés, inglés de estados unidos, inglés de Inglaterra, portugués, etc.) las lenguas de señas también presenta diferencias. [4]

Figura 3.1 Alfabeto de la Lengua de Señas Chilena. [10]

## Imágenes Digitales

La imagen digital está formada por un conjunto definido de puntos llamados píxeles. La imagen tradicional, sobre material posible, también está formada por puntos o granos; la gran diferencia es que en la fotografía tradicional, los granos son irregulares y están situados de forma aleatoria, mientras que en la digital forman una matriz con filas y columnas.

### Tipo de Imágenes Digitales

Por la forma de manejar los datos en un archivo de imagen, se puede hablar de dos modos principales para manipular la información que integra una imagen digital. Estos modos son las imágenes de mapa de bits y las imágenes vectoriales. Dado que cada uno se adapta mejor a un tipo de imagen, antes de conocer los diferentes espacios, debe conocerse el funcionamiento tanto de imágenes vectoriales como de imágenes de mapa de bits.

#### Imágenes de Mapa Bits

Las imágenes de mapa de bits están formadas por una rejilla de celdas. A cada una de estas celdas, que se denominan píxeles, se le asigna un valor de color y luminancia propios. Por esto, cuando vemos todo el conjunto de celdas, tenemos la ilusión de una imagen de tono continuo.

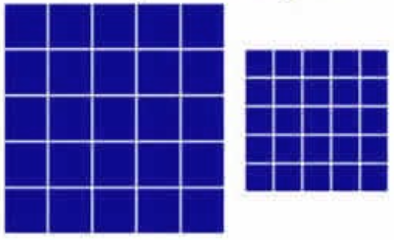
El pixel es una unidad de información, no una unidad de medida, ya que no se corresponde con un tamaño concreto. Un píxel puede ser muy pequeño (0.1 mm) o muy grande (1 cm). Cuando creamos una imagen de mapa de bits se genera una rejilla específica de píxeles; por esto, al modificar su tamaño, transformamos, a su vez, la distribución y coloración de los píxeles, por lo que los objetos, dentro de la imagen, suelen deformarse. Así los objetos pierden o ganan algunos de los píxeles que los definen, como se muestra en la figura 3.2.

Figura 3.2 Dos rejillas de 5x5 píxeles. [5]

Gracias a esta característica, que siempre hay que tener en cuenta, las imágenes de mapa de bits se crean con un tamaño determinado y pierden calidad si se modifican sus dimensiones.

#### Imágenes Vectoriales

Los llamados gráficos orientados a objetos son las imágenes vectoriales. Su tamaño es mucho más reducido, en comparación con los mapas de bits, porque el modo como organizan la información de una imagen es más simple que en aquellos. Dicha simplicidad radica en generar los objetos que conforman una imagen a través de trazos geométricos determinados por cálculos y fórmulas matemáticas. De manera tal que los gráficos vectoriales se visualizan a partir de las coordenadas de una línea guardadas como referencia, mismas que forman los objetos a partir de la definición matemática de los puntos y líneas rectas o curvas.

Mientras que en un mapa de bits los objetos se definen píxel por píxel, por lo que no pueden manipularse individualmente, caso contrario sucede con las imágenes vectoriales, con ellas cada objeto, dentro de una imagen, puede modificarse sin que se alteren los demás. En un gráfico de vectores se delimitan por la posición de los puntos inicial y final y por la trayectoria de la línea que los une. Como la información se guarda matemáticamente, esto hace posible que las imágenes vectoriales sean independientes de la resolución del monitor, ya que, al contrario de los mapas de bits, no dependen de los píxeles. Por lo tanto, tienen la mínima o máxima resolución que permita el formato en que se almacenen. Asimismo, al aumentar o reducir la resolución de un gráfico vectorial, tampoco se pierde definición en la imagen, porque la computadora solo tiene que redefinir las coordenadas y vectores que la imagen contenga pera redimensionar los objetos. [5]

### Colores

Aunque el color es una de las características más importantes que definen a los objetos, hasta hace poco tiempo se le prestaba poca atención debido al coste computacional y a la memoria necesaria para procesar imágenes de color.

Para obtener una imagen hay que transformar primero los parámetros cromáticos en eléctricos y representar los colores, lo cual puede hacerse de distintas maneras, dando lugar a los distintos espacios de color.

Un espacio de color es un método por el que se puede especificar, crear o visualizar cualquier color. Dependiendo del tipo de sensor y aplicación se han desarrollado diversos espacios de colores.

#### Espacio RGB

El espacio RGB se basa en la combinación de tres señales de luminancia cromática distinta: rojo, verde y azul (Red, Green, Blue). La manera más intuitiva de conseguir un color concreto es determinar la cantidad de color rojo, verde y azul que se necesita combinar, para ello se realiza la suma aritmética de las componentes: X = R + G + B. [3]

Si tomamos en cuenta que cada una de las componentes tiene variaciones de tonalidad que van desde el 0 al 255 (28), entonces el cubo posee (28)3 = 16,777,216 colores representados vectorialmente; de esta manera podemos distinguir los extremos como el rojo (255,0,0), verde (0,255,0), azul (0,0,255), negro (0,0,0) y blanco (255,255,255).

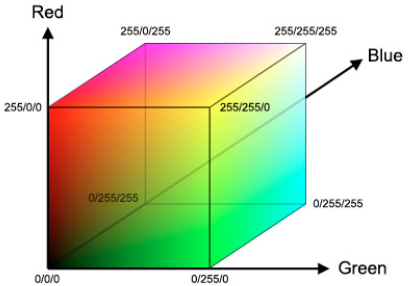


Figura 3.3 Representación tridimensional del espacio RGB.

#### Espacio HSI

El espacio de color HSI se basa en el modo de percibir los color que tenemos los humanos. Dicho sistema caracteriza el color en términos de tinte (Hue), saturación (Saturation) y brillo (Intensity).

El tono está asociado con la longitud de onda dominante en una mezcla de ondas luminosas. Así el tono representa el color dominante tal y como lo percibimos; cuando decimos que un objeto es rojo, verde o café, estamos indicando su tono en la Figura 3.1 (a) Tono.

La intensidad representa la iluminación percibida. Da la sensación de que algún objeto refleja más o menos luz. Este atributo lo podemos ver claramente en un televisor en blanco y negro en la Figura 3.1 (b) Saturación.

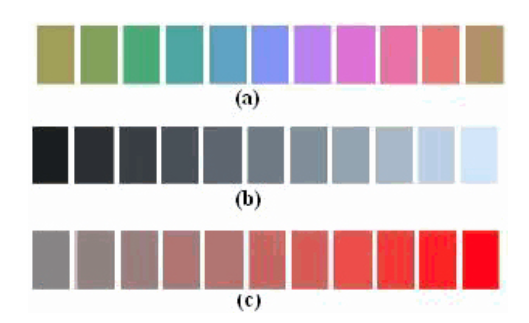
La saturación se refiere a la cantidad de luz blanca mezclada con el color dominante. La saturación es un atributo que nos diferencia un color intenso de uno pálido. Cada uno de los colores primarios tiene su mayor valor de saturación antes de ser mezclados con otros. Así el azul cielos es muy claro (menos saturado), mientras que el azul marino es más opaco (más saturado). Otro ejemplo es el color rosa (rojo y blanco) que está menos saturado; mientras que el color rojo está totalmente saturado en la Figura 3.1 (c) Intensidad. [6]

Figura 3.4 HSI: (a) Tono, (b) Saturación, (c) Intensidad. [6]

## Procesamiento de Imágenes

El preprocesamiento de imágenes comprende aquellos algoritmos cuya finalidad es conseguir una mejora en la apariencia de la imagen original. Esta mejora en la apariencia consiste en resaltar determinadas características de la imagen o en eliminar aquello que las oculta. No se trata de mejorar la imagen de manera que quede lo más ideal posible sino que pueda ser analizada de la forma más simple.

### Manipulación de Contraste

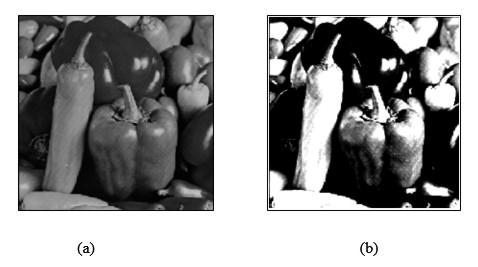
El contraste muestra las variaciones locales del brillo, incrementa el cambio de luminosidad entre las zonas más oscuras o más claras de una fotografía, simulando a su vez, un mejor enfoque y claridad de imagen [1]. Su manipulación busca mejorar unas zonas en prejuicio de otras, y así distanciar más los píxeles con valores bajos o a la inversa. En una imagen en escala de grises, consiste en aumentar el rango dinámico de los niveles de gris de la imagen [2].

Figura 3.5 (a) Imagen Original, (b) Aumento de Contraste. [2]

### Eliminación de Ruido

Todas las imágenes tienen una cierta cantidad de ruido, valores distorsionados, bien debidos al sensor de la cámara o al medio de transmisión de la señal. El ruido se manifiesta generalmente en píxeles aislados que toman un valor gris deferente al de sus vecinos. El ruido puede clasificarse en cuatro tipos:

* + Gaussiano: Produce pequeñas variaciones en la imagen. Es debido, por ejemplo, a las diferentes ganancias en el sensor, ruido en los digitalizadores, perturbadores en la transmisión, etc.
  + Impulsional: El valor del píxel no tiene relación con el valor ideal sino con el valor del ruido que toma valores muy altos o bajos. Se caracteriza entonces porque el píxel toma un valor máximo, causado por una saturación del sensor, o mínimo si se ha perdido la señal.
  + Frecuencial: La imagen obtenida es la suma entre imagen ideal y otra señal, la interferencia, caracterizada por ser una senoide de frecuencia determinada.
  + Multiplicativo: La imagen obtenida es la multiplicación de dos señales.

Figura 3.6 Diferencia entre imagen sin ruido y con ruido. [9]

### Realce de Bordes

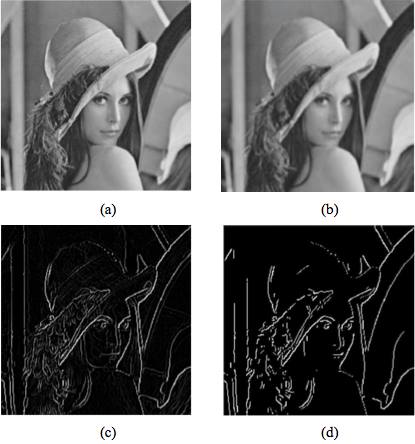
El realce de bordes transforma una imagen de manera que exhibe sólo el detalle de bordes o fronteras. Los bordes aparecen como las líneas de contorno de los objetos dentro de la imagen (Figura 3.7 (c) y Figura 3.7 (d)). Estos contornos pueden utilizarse en posteriores operaciones de análisis de imágenes para el reconocimiento de objetos o rasgos. [7]

Figura 3.7 Realce de bordes, (a) imagen original, (b) imagen con filtro gaussiano para eliminar ruido, (c) realce de bordes sin filtro gaussiano, (d) realce de bordes con filtro gaussiano.

El realce de bordes tiene un efecto opuesto a la eliminación de ruido, ya que de lo que se trata es de resaltar aquellos píxeles que representan un valor de gris distinto al de sus vecinos. Por ello, si la imagen es ruidosa, el efecto de ruido se multiplicará; por lo que antes de resaltar los bordes, habrá que eliminar el ruido como se muestra en la Figura 3.7.

Los realces de bordes son implementados a través de filtros espaciales. Uno de los más usados es el método de desplazamiento y sustracción.

#### Método de Desplazamiento y Sustracción

Permite realzar información de bordes verticales u horizontales, desplazando una imagen a la izquierda en un píxel y restando esta nueva imagen de la original, se resaltarán los bordes verticales. Esto se debe a que el valor de brillo de cada pixel de entrada es restado de su vecino horizontal, dando un valor de la diferencia de brillo entre ambos, conocida como pendiente en el contexto de una imagen.

Si dos pixeles adyacentes tienen brillos muy diferentes (existencia de borde), la diferencia resultará grande. Por el contrario, si dos pixeles adyacentes tienen valores similares de brillo (no existencia de borde), la diferencia de brillos resultante será pequeña. El efecto es una imagen donde aparecen los contornos direccionales. Análogamente, el realce de borde horizontal se implementa desplazando la imagen un pixel hacia arriba y restando [7].

## Extracción de Características

### Detección de Bordes

#### Técnicas Basadas en el Gradiente

#### Operadores de Segundo Orden

#### Métodos basados en las Derivadas de Gaussianas

#### Operadores de Ajuste al Modelo

#### Detección de Líneas

#### Detección de Esquinas

### Texturas

#### Análisis Estadístico de Texturas

#### Análisis Frecuencial de Texturas

### Detección de Movimiento

## Segmentación de Imágenes

### Segmentación basada en la detección de bordes

#### Método basado en el gradiente

#### La transformada de Hough

### Segmentación de Regiones

#### Umbralización

#### Etiquetado

#### Crecimiento de Regiones

#### División y Unión de Regiones

# Conclusiones

# Referencias Bibliográficas