



DESARROLLO DE SOFTWARE PARA LA DETERMINACIÓN DE LA CALIDAD DEL AZÚCAR UTILIZANDO VISIÓN ARTIFICIAL

TRABAJO DE GRADUACIÓN ELABORADO PARA LA
FACULTAD DE INGENIERÍA

PARA OPTAR AL GRADO DE
INGENIERO EN MECATRÓNICA

PRESENTADO POR
IVAN ERNESTO LÓPEZ RAMOS
GERARDO ANTONIO LÓPEZ RAMOS
RICARDO ALEXANDER ARGUETA MORÁN

ASESOR
Msc. MANUEL CARDONA


SOYAPANGO, EL SALVADOR
AGOSTO DE 2015

UNIVERSIDAD DON BOSCO
FACULTAD DE INGENIERÍA

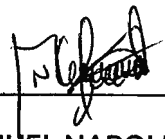


EVALUACIÓN DEL TRABAJO DE GRADUACIÓN

**DESARROLLO DE SOFTWARE PARA LA DETERMINACIÓN
DE LA CALIDAD DEL AZÚCAR UTILIZANDO VISIÓN
ARTIFICIAL**



ING. DAVID ARMANDO FLAMENCO
LECTOR



MSC. MANUEL NAPOLEÓN CARDONA
ASESOR



ING. MARIA CELIA PARADA
ADMINISTRADORA DEL PROCESO

UNIVERSIDAD DON BOSCO

RECTOR

OSCAR RODRÍGUEZ BLANCO

SECRETARIA GENERAL

MG. YESENIA XIOMARA MARTÍNEZ

DECANO DE FACULTAD DE INGENIERÍA

ING. OSCAR GIOVANNI DURÁN VIZCARRA

ASESOR DEL TRABAJO

MSC. MANUEL NAPOLEON CARDONA

LECTOR

ING. DAVID ARMANDO FLAMENCO

ADMINISTRADORA DEL PROCESO

ING. MARIA CELIA PARADA

Resumen

El objetivo de este trabajo es dar a conocer la base teórica y las herramientas que se necesitaron para desarrollar un software dedicado a la extracción y análisis de propiedades geométricas de cristales de azúcar, como proceso de control de calidad en la industria azucarera. Utilizando técnicas de visión artificial y herramientas estadísticas para el control de calidad para la obtención de datos útiles y concluyentes, concernientes a las variables de producción a las cuales fueron sometidas las muestras estudiadas.

El software ha sido desarrollado en MatLab haciendo uso de *toolboxes* específicas en los campos de procesamiento digital de imágenes y cálculos estadísticos; con lo cual se logró un producto con funciones e interfaces específicas para esta necesidad, reduciendo en gran medida el tiempo del proceso de control de calidad y obteniendo datos comparables con los obtenidos mediante sistemas, tanto homólogos como diferentes, que son estándares en la industria.

Abstract

The goal of this work is to present the theoretical concepts and tools that were necessary to develop a software dedicated to the extraction and analysis of geometric properties of sugar crystals, as quality control process in the sugar industry. Using computer vision techniques and statistical tools for quality control, to obtain useful and conclusive data, concerning to the production variable to which the studied samples were submitted.

The software have been developed in MatLab using specific toolboxes in the fields of digital image processing and statistical calculation; with this, we achieve a product with specific functions and interfaces for this need, significantly reducing the quality control process time and obtaining data comparable with those obtained by systems, both homologous and different, that are standard in the industry.

Agradecimientos

Ante este logro quiero agradecer:

A Dios por la vida, cariño, bendiciones, personas y oportunidades que ha puesto a lo largo de todo mi camino.

A mis padres Mario Antonio López y María Patricia Ramos de López por todo su esfuerzo y apoyo incondicional durante toda mi vida. Por ser un pilar fundamental durante mi desarrollo personal y profesional; y por ser un ejemplo a seguir.

A mi hermano Gerardo Antonio López por ser un amigo, colega y apoyo durante toda mi vida.

A mi nana Ana Hernández por dedicarnos su vida y cariño; y por ser un ejemplo de amor y dedicación.

A mis compañeros de proyecto Ricardo Argueta y José Gálvez por su amistad y por su apoyo y participación a lo largo del desarrollo de este trabajo.

A la institución Salesiana, a todos sus miembros y formadores, que a lo largo de 20 años me han guiado con su educación y sabiduría.

A los ingenieros Manuel Cardona, David Flamenco y Ricardo García por compartir con nosotros sus conocimientos y estar siempre dispuestos a ayudarnos.

Y a todos mis amigos y personas que forman parte de mí vida día a día, por motivarme constantemente y creer en mí. Por apoyarme en las buenas y en las malas e inspirarme a seguir creciendo y mejorando.

Ivan Ernesto López Ramos

Agradecimientos

A Dios, por haberme dado la sabiduría y fortaleza para culminar mi carrera, a mi padre por apoyarme siempre y cuyos consejos han sido pilares fundamentales para convertirme en un profesional. A mi madre, por acompañarme y guiarme en el camino correcto hasta el final de mi carrera. A mi hermano, por ser un compañero y amigo a lo largo de esta experiencia. A mi tía Ana Hernández, por brindarme su apoyo y cariño, y por llevarme siempre en su oración. A mis amigos y amigas que estuvieron siempre a mi lado apoyándome en distintas ocasiones y quienes confiaron que esta obra sería posible. Y finalmente a mis profesores, quienes invirtieron su tiempo y transmitieron su sabiduría a lo largo de mi formación profesional. En fin, muchas gracias a todas estas personas con la que comparto este logro tan importante que ha sido fruto de la dedicación y sacrificio de todos.

A todos ellos mis agradecimientos

Gerardo Antonio López Ramos

Agradecimientos

Gracias primeramente a Dios por todas las bendiciones que me ha dado, por su apoyo y todas las oportunidad que me ha presentado en el camino.

Gracias de corazón a mis padres José Ricardo Argueta y Ana Isabel de Argueta por brindarme la oportunidad de estar aquí ya que sin ellos nada de esto sería posible, a mis hermanos porque sin su apoyo y compañía hubiera decaído en el camino.

Gracias a la comunidad estudiantil de la Universidad Don Bosco y todas esas personas que durante mi formación académica aportaron a mi crecimiento profesional e individual, ya que hay cosas que ningún libro puede enseñar.

Gracias a nuestro compañero y amigo José Alfredo Gálvez por el apoyo brindado durante todo el desarrollo de la investigación como sus ideas, aportes y recomendaciones. Ha sido un privilegio contar con su apoyo y su amistad.

Gracias a nuestro tutor Manuel Napoleón Cardona por orientarnos en el camino, apoyarnos en el desarrollo, despejar dudas, su paciencia y la motivación brindada durante el desarrollo de toda la investigación.

Gracias a todas aquellas personas que me apoyaron durante el desarrollo de esta tesis de grado especialmente a mis compañeros de tesis Iván Ernesto López y Gerardo Antonio López por su amistad, apoyo, paciencia y mucho más.

Ricardo Alexander Argueta Morán

Índice

CAPÍTULO I: ESTADO DEL ARTE **1**

1.1 INTRODUCCIÓN	1
1.2 OBJETIVOS	3
1.2.1 OBJETIVO GENERAL	3
1.2.2 OBJETIVOS ESPECÍFICOS	3
1.3 JUSTIFICACIÓN	4
1.4 ANTECEDENTES	6
1.4.1 MÉTODO CILAS <i>PARTICLE LASER SIZE</i>	6
1.4.2 TAMIZADO MEDIANTE EL MÉTODO DE BUTLER	7

CAPÍTULO II: VISIÓN POR COMPUTADORA **10**

2.1 ¿QUÉ ES LA VISIÓN POR COMPUTADORA?	10
2.1.1 ÁREAS DE APLICACIÓN DE LA VISIÓN POR COMPUTADOR	10
2.1.1.1 Procesos Industriales	11
2.1.1.2 Percepción remota.	11
2.1.1.3 Robótica	12
2.1.1.4 Medicina	12
2.2 HISTORIA	12
2.3 ANALOGÍA ENTRE VISIÓN HUMANA Y VISIÓN ARTIFICIAL	14
2.4 CONCEPTOS BÁSICOS	16
2.4.1 IMAGEN DIGITAL	16
2.4.2 IMAGEN BINARIA	18
2.4.3 IMAGEN RGB	18
2.4.4 PIXEL	19
2.4.5. VECINDAD ENTRE PÍXELES	19
2.4.6 CONECTIVIDAD	20

2.4.7 HISTOGRAMA	21
2.4.8 ILUMINACIÓN	21
2.4.9 CONTRASTE	22
2.4.10 DINÁMICA	23
2.4.11 ETIQUETADO DE OBJETOS EN IMÁGENES BINARIAS	24
2.4.12 CARACTERÍSTICAS DE OBJETOS BINARIOS	25
2.4.13 PERÍMETRO	25
2.4.14 ÁREA	26
2.4.15 BOUNDING BOX	26
2.4.16 MOMENTO DE INERCIA	26
2.5 SISTEMA DE PROCESAMIENTO DIGITAL DE IMÁGENES	26
2.5.1 ADQUISICIÓN DE LA IMAGEN	27
2.5.2 ALMACENAMIENTO	27
2.5.3 PROCESAMIENTO	28
2.5.4 VISUALIZACIÓN	28

CAPÍTULO III: CONTROL DE CALIDAD DEL AZÚCAR **29**

3.1 CALIDAD DENTRO DE LA INDUSTRIA ALIMENTICIA.	29
3.1.1 HISTORIA	29
3.1.2 TIPOS DE CALIDAD PRESENTES EN LA INDUSTRIA ALIMENTICIA	30
3.1.3 NORMAS QUE RIGEN LA CALIDAD DE LOS ALIMENTOS.	31
3.2 OBJETIVO Y FINALIDAD DEL USO DE MÉTODOS ESTADÍSTICOS PARA EL CONTROL DE CALIDAD	33
3.2.1 ¿QUÉ ES LA ESTADÍSTICA?	33
3.2.2 PRESENCIA DE LA ESTADÍSTICA EN EL CONTROL DE CALIDAD	34
3.2.2.1. Detección de defectos	34
3.2.2.2. Inclusión de técnicas de muestreo	35
3.2.2.3. Análisis de múltiples variables	35
3.2.3. IMPORTANCIA DEL CONTROL ESTADÍSTICO DE PROCESOS.	37

3.3 PRINCIPIOS ESTADÍSTICOS EMPLEADOS PARA EL PROCESO DE CONTROL DE CALIDAD	37
3.3.1 HERRAMIENTAS ESTADÍSTICAS DE MUESTREO	37
3.3.2 HERRAMIENTAS ESTADÍSTICAS DE EXPOSICIÓN DE RESULTADOS	39
3.3.3 HERRAMIENTAS ESTADÍSTICAS UTILIZADAS PARA CREAR LA TABLA DE FRECUENCIAS	41
3.3.3.1. Medidas de dispersión.	41
3.3.3.2. Medidas de tendencia central	42
3.3.3.3. Elementos de la tabla de frecuencia	43
3.4 ANÁLISIS CUALITATIVO Y DESCRIPTIVO DEL PROCESO.	44
3.4.1 DESARROLLO DE SELECCIÓN DE LA MUESTRA A ANALIZAR	44
3.4.1.1 Selección del método de muestreo a utilizar	44
3.4.1.2 Asignación de nivel de inspección	46
3.4.1.2 Tamaño de la muestra.	47
3.4.2 ANÁLISIS DE RESULTADOS DE LA MUESTRA	49
3.4.2.1 Construcción de tabla de frecuencias y cálculo de medidas de dispersión	49
3.4.2.2 Coeficiente de variación ponderado por volumen	54
3.4.3 VISUALIZACIÓN DE RESULTADOS	56
3.4.4 ANÁLISIS DE RESULTADOS.	58
 CAPÍTULO IV: DESARROLLO DE SOFTWARE	 59
4.1 ¿QUÉ ES MATLAB?	59
4.1.1 DESCRIPCIÓN GENERAL	59
4.1.2 HISTORIA	59
4.2 HERRAMIENTAS DE VISIÓN ARTIFICIAL DE MATLAB	61
4.2.1 GENERALIDADES SOBRE LAS <i>TOOLBOX</i>	61
4.2.2 <i>TOOLBOX</i> : PROCESAMIENTO DE IMÁGENES Y VISIÓN ARTIFICIAL	62
4.3 DEFINICIÓN DEL SOFTWARE	67
4.3.1 ¿QUÉ ES SUGARVISION?	67
4.3.2 CONSIDERACIONES DE DESARROLLO DEL SOFTWARE	67
4.3.2.1 Requerimientos del software	67

4.3.2.2 Pasos	68
4.4 INTERFAZ Y FUNCIONAMIENTO DE SUGARVISION	70
4.4.1 INTERFAZ	70
4.4.2 FUNCIONES Y PASOS	72
4.4.2.1 Importar archivos de imagen	72
4.4.2.2 Conversión de imagen a matriz	73
4.4.2.3 Cambio de Umbral	74
4.4.2.4 Transiciones en Navegador de Funciones	76
4.4.2.5 Definición de línea de referencia para calibración	77
4.4.2.6 Limpieza de imagen	78
4.4.2.7 Análisis de regiones	79
4.4.2.8 Visualización de Estadística	82
4.4.3 BARRA DE HERRAMIENTAS	84
4.5 CONSIDERACIONES DE HARDWARE	89
 <u>CAPÍTULO V: VALIDACIÓN DE RESULTADOS</u>	 <u>93</u>
 5.1 DESCRIPCIÓN DE MÉTODOS UTILIZADOS	 93
5.1.1 MEDICIÓN MANUAL UTILIZANDO SOFTWARE NX	93
5.1.2 MEDICIÓN MANUAL UTILIZANDO ESTEREOSCOPIO	94
5.2 COMPARACIÓN DE RESULTADOS	96
5.2.1 COMPARACIÓN NX vs. SUGARVISION	96
5.2.2 COMPARACIÓN ESTEREOSCOPIO vs. SUGARVISION	97
5.2.3 RESULTADO DE LAS COMPARACIONES	98
 <u>CONCLUSIONES</u>	 <u>99</u>
 <u>BIBLIOGRAFÍA</u>	 <u>100</u>

Capítulo I

Estado del Arte

1.1 Introducción

En la actualidad, la mayoría de herramientas que la industria nacional utiliza para suplir sus necesidades son de carácter privativo; es decir, son software o aplicaciones desarrolladas y parametrizadas por una empresa tercera, sometiendo al usuario a características generales que en la mayoría de casos, solo cumplen de manera parcial los requisitos del usuario. Debido a esto surge la propuesta de que la industria empiece a dejar de ser solamente un usuario y se convierta en el desarrollador de sus propias herramientas, para poder abordar de mejor manera sus necesidades, y de esta manera incentivar la investigación y desarrollo nacional.

Esto despierta nuestro interés en desarrollar un proyecto para la industria nacional, y al encontrar una oportunidad en la industria azucarera, nos motiva a contribuir con una alternativa para los procesos de control de calidad morfológica. Las características del proyecto nos empujan a realizar investigaciones en las áreas de la visión artificial, estadística y procesos de control de calidad; así como también a hacer uso de conocimientos previos de lógica, programación y desarrollo de interfaces.

Este proyecto se desarrolla a lo largo de tres etapas que son fácilmente identificables en el presente documento:

1. Investigación: En los Capítulos II y III encontrará la investigación realizada a los campos de visión artificial, estadística y control de calidad en la industria alimenticia. En estos capítulos se brinda la

información necesaria para la comprensión de aspectos y términos técnicos utilizados a lo largo del documento.

2. Desarrollo: Estos aspectos, consideraciones y conceptos son utilizados en el desarrollo del software que se aborda en el Capítulo IV, donde además se dan a conocer las razones de la elección de MatLab como herramienta de desarrollo y compilación. También se guía al lector a través de un recorrido por las pantallas, funciones y herramientas de SugarVision, explicando los procesos que se realizan en cada función y dando a conocer los comandos más importantes utilizados para el desarrollo de SugarVision.
3. Validación: En el capítulo V se explica el proceso de validación de los datos obtenidos con el SugarVision en comparación a otros métodos, dando a conocer en qué medida son comparables uno del otro.

1.2 Objetivos

1.2.1 Objetivo General

- Desarrollar un software que sea capaz de extraer propiedades geométricas de cristales de azúcar y someterlas a modelos estadísticos para la determinación de su calidad.

1.2.2 Objetivos Específicos

- Realizar investigaciones sobre procesos de control de calidad, visión artificial y modelos estadísticos usados en la industria alimenticia.
- Determinar los métodos y herramientas necesarios para un software que cumpla con funciones y resultados estándares en el control de calidad.
- Brindar al usuario una interfaz sencilla e intuitiva, enfatizando en la visualización de la imagen y su procesamiento.
- Reducir tiempos en los procesos de control de calidad.
- Validar el software realizando comparaciones de resultados con otros sistemas utilizados en la industria.

1.3 Justificación

Los procesos de control de calidad se han vuelto indispensables en la industria para la verificación de los productos finales que se entregarán a los consumidores y debido a que los tiempos de producción y evaluación son de suma importancia para la solvencia de las empresas, se han encontrado en la necesidad de hacerse de herramientas cada vez más eficientes, rápidas y accesibles.

Por otro lado se ha identificado que las herramientas existentes para el control de calidad poseen deficiencias dependiendo de su naturaleza: Una herramienta de control de calidad por medios físicos, como el peso, nos brinda datos calculados mediante funciones donde se realizan suposiciones, dando como resultado datos menos reales e imprecisos; una herramienta por medio de imágenes sin visión por computador, nos añade certeza a los datos, pero es un proceso que lleva demasiado tiempo y se vuelve engorroso para el operario; y las herramientas con visión artificial disponibles en el mercado nacional poseen grandes deficiencias, debido a que los parámetros variables son muy limitados, sin mencionar que adquirir el software e instrumentos necesarios es muy costoso.

Debido a esto se desarrolla SugarVision como una alternativa que solvete las carencias antes mencionadas, logrando un software en el cual el usuario puede intervenir a lo largo del procesamiento de la imagen que contiene los elementos a ser analizados, además de visualizar cómo influye cada etapa en la imagen. Esto nos brinda flexibilidad para la realización de los estudios, además, al realizar el procesamiento de la información de manera digital, se logran tiempos de análisis sumamente reducidos.

Las características de este software benefician en gran medida al área de control de calidad de compañías azucareras, al dotarlas de herramientas más rápidas que reducen el tiempo de estudio y aumentan la capacidad del

procesamiento de los indicadores de calidad. Además de reducir los costos al no necesitar equipos de alta tecnología especiales para la adquisición de imágenes.

Cabe resaltar que SugarVision posee un flujo de trabajo definido, donde el usuario navega por las funciones de manera secuencial, de manera que el paso anterior es necesario para el siguiente y así sucesivamente. Además en cada uno de estos pasos se muestran los conceptos de visión artificial utilizados en la pantalla, así como la información y descripción sobre el proceso que se realiza en este, para que el usuario sea consciente de su función y utilice de mejor manera el software.

1.4 Antecedentes

En la industria alimenticia existen una serie de métodos que se han utilizado tradicionalmente para realizar el control de calidad, estos métodos consisten en el uso de una serie de equipos y modelajes matemáticos que permiten obtener resultados confiables, el desarrollo de estos métodos se explicaran a continuación.

1.4.1 Método CILAS *Particle Laser Size*

El método *Particle Laser Size* consiste en la inspección visual de partículas mediante la refracción de la luz, fue desarrollado por la empresa CILAS alrededor del año 1967 y utiliza un equipo que lleva el mismo nombre CILAS *Laser particle size*. Para llevar a cabo el método se toma una pequeña muestra de cristales de azúcar y se ponen en una maquina especial llamada *Particle Size Analyzer*.

Una vez puesto los cristales en la cámara superior de la máquina, se configura el tipo de muestra si es una muestra seca o liquida, se configura como muestra seca y comienza el funcionamiento de la maquina; los cristales de azúcar son insertados uno por uno dentro de una cámara especial equipada con rayos laser capaz de detectar cristales desde 40 nm hasta 25mm, los cristales entran a la cámara y mediante la refracción de luz generada por los láseres, se obtiene el diámetro equivalente de cada uno de los cristales que se van insertando.

Posteriormente a que la información es tomada y transmitida a un ordenador con el software que lleva el nombre de CILAS, se procede al análisis de la información, donde cada cristal primeramente es modelado como una esfera perfecta tomando como radio el diámetro equivalente y posteriormente se procede a calcular el CVV, como se realizó en el capítulo 3.

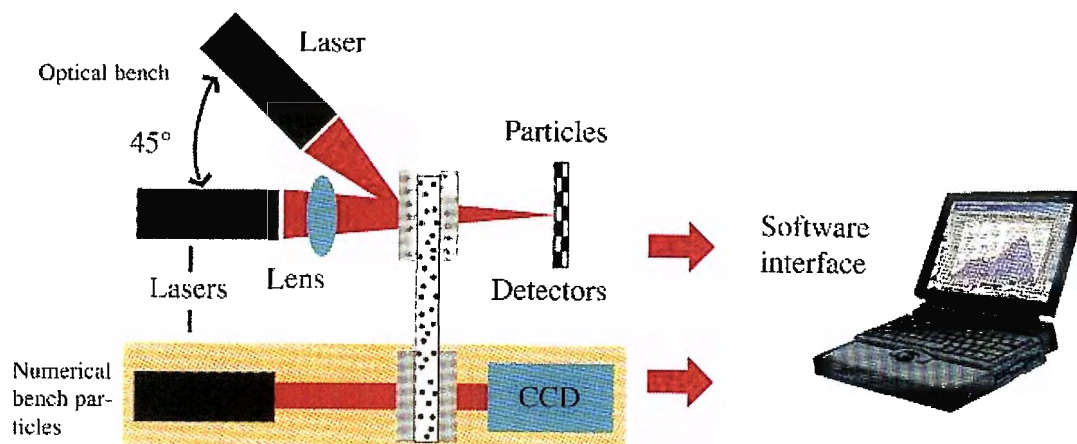


Imagen 1.1 - Proceso de obtención de información. (Tomada de <https://commons.wikimedia.org/>)

Finalmente el software genera un reporte que nos entrega la información obtenida de la muestra de cristales en forma de un histograma gausseano, y nos detalla algunos datos que son relevantes a la hora de considerar los resultados como son: el diámetro equivalente general de la muestra, el coeficiente de variación ponderado por volumen, etc.

1.4.2 Tamizado mediante el método de Butler

Este método es uno de los distintos métodos existentes para realizar tamizado normado por el ICUMSA existente desde 1987 pero en este caso se utilizara la edición 2011; este método nos permite calcular el coeficiente de variación (C.V.) y la apertura media (M.A.), y se basa en que el tamaño medio de cada cristal es igual al promedio de la apertura de la malla del tamiz, solo cabe resaltar que puede haber más de una malla para realizar este método y los resultados obtenidos son ponderados por peso.

Por lo que el procedimiento consiste en la toma de una muestra de azúcar la cual se vierte en una serie de tamices donde el numero puede variar dependiendo de los análisis a realizar, y luego se toma la muestra que solamente haya pasado nuestro tamiz de interés para pesarla y obtener el porcentaje que

representa de la muestra total; con esos valores se procede a obtener el tamaño medio de la partícula y posteriormente a realizar los cálculos de M.A. y C.V. basándose en siguiente ecuaciones.

Se procede a calcular el valor de abertura media (M.A.)

$$M.A. = \frac{\sum(f * d'')}{\sum f} \quad (1.1)$$

Donde f representa el porcentaje de cristales que traspaso cada uno de los tamices y d'' el tamaño de cada una de las aperturas de los tamices (cabe resaltar que cada abertura del tamiz es cuadrada).

El cálculo del coeficiente de variación se realiza en dos pasos por lo que primero se calcula la desviación típica (S.D.) tal como se muestra a continuación.

$$S.D. = \sqrt{\frac{\sum\{f * (M.A. - d'')^2\}}{\sum f}} \quad (1.2)$$

Donde $\sum\{f * (M.A. - d'')^2\}$ es la sumatoria de la multiplicación del porcentaje de cristales que pasa por el cuadrado de la resta entre la abertura media menos el tamaño de la apertura.

Se procede calcular el coeficiente de variación ponderado por peso (C.V.W.) con la siguiente ecuación.

$$C.V.W. = \frac{S.D. * 100\%}{M.A.} \quad (1.3)$$

Posteriormente con cada los cálculos obtenidos se realiza una tabla donde se pueda apreciar la información y tomarla como análisis de resultados, a continuación se muestra un ejemplo.

Cálculo de *MA* y *CV* con el método de Butler

Tamaño de la apertura (mm)	Tamaño medio de partícula, d'' (mm)	% retenido f	$f \times d''$	$MA - d''$	$f \times (MA - d'')^2$
0,71	0,780	11,3	8,814	-0,287	0,9308
0,60	0,655	19,3	12,642	-0,162	0,5065
0,50	0,550	14,6	8,030	-0,057	0,0474
0,43	0,465	14,6	6,789	0,028	0,0114
0,36	0,395	17,6	6,952	0,098	0,1690
0,30	0,330	6,3	2,079	0,163	0,1674
0,25	0,275	5,9	1,623	0,218	0,2804
bandeja de fondo	0,230	10,4	2,392	0,263	0,7194
Total		100,0	49,321		2,8323

Imagen 1.2 - Ejemplo de tabla expositora de resultados (Tomada de Métodos Imcusa para el análisis de azúcar, Método de Butler, Editorial Esleiver, Edición 2011.)

Descrito los procedimientos de los métodos anteriores se puede partir de que el desarrollo de herramientas que ayuden al control de calidad han existido desde ya hace algunos años y que las representaciones de los resultados pueden ser ponderados de distintas maneras en estos casos por partículas y por peso, pero también podría llegar a ser ponderados por volumen.

El modelaje de estos métodos permitirá observar cual es la información necesaria para realizar el control de calidad y los distintos tratamientos que se le deben realizar para obtener resultados concretos; marcando así un punto de partida en la investigación referente a la parte de control de calidad estadístico.

Capítulo II

Visión por computadora

2.1 ¿Qué es la visión por computadora?

La visión por computadora es un término muy utilizado en los últimos años que tiende a ser confundido con otro tipo de conceptos similares como: procesamiento digital de imágenes, reconocimiento de formas, inteligencia artificial, álgebra difusa, etc. Que si bien la visión por computadora utiliza algunas técnicas o herramientas de ellas, no pueden ser catalogados como iguales.

El término visión por computadora se refiere a la ciencia que hace uso de bases teóricas y algorítmicas, para obtener información sobre el mundo real y desarrollar sistemas capaces de interpretar automáticamente su contenido, a partir de una o varias imágenes digitales.

La visión por computadora tiene como objetivo principal la interpretación de la estructura y propiedades de las imágenes suministradas por cámaras, que captan la realidad de un escenario o mundo tridimensional; utilizando para ello, la gran capacidad de procesamiento y las funciones de un computador digital.

2.1.1 Áreas de aplicación de la visión por computador

A medida que pasa el tiempo y la tecnología va avanzando, el área de aplicación de la visión por computador se va expandiendo cada vez más. Por esa razón en el presente apartado se muestran las principales áreas de aplicación de la visión por computador; las cuales son: procesos industriales, percepción remota, robótica y apoyo al diagnóstico médico.

2.1.1.1 Procesos Industriales

En el ámbito de las aplicaciones en procesos industriales, la visión por computador es aplicada en dos grupos: el control de calidad de productos manufacturados y el control e inspección de los propios procesos de fabricación. Pero la que tiene mayor auge o se utiliza más es el control de calidad de productos por medio de visión por computador. Los principales objetivos de este tipo de aplicación son:

- Verificar la presencia de las características esperadas de los objetos producidos.
- Verificar las dimensiones de dichas características.
- Verificar las interrelaciones entre características (por ejemplo distancias entre centros de gravedad y ángulos entre planos)

Generalmente el proceso de control de calidad se realiza mediante la inspección virtual de un objeto; esto se refiere a que se verifica por medio de un software si un objeto cumple con determinados criterios, utilizando un proceso de comparación con un patrón o un objeto modelo que posee las características relevantes del objeto analizado.

2.1.1.2 Percepción remota.

Entre las aplicaciones de percepción remota se pueden mencionar en geología, la detección de movimientos de terrenos captando dos imágenes en diferentes momentos de tiempo y la reconstrucción de áreas 3-D mediante visión estereoscópica. En meteorología, la detección y predicción del movimiento y la evolución de masas nubosas, u otros fenómenos meteorológicos, a través de imágenes recibidas vía satélite. Entre otras aplicaciones como el control de asentamientos urbanísticos, monitorización de recursos naturales, control de incendios e inundaciones, etc.

2.1.1.3 Robótica

Esta es una de las áreas donde se explota más el recurso de la visión por computador; el cual se basa en un sistema multisensorial que integra la información de diferentes tipos de sensores y la información procedente de los sistemas de visión estereoscópica que posee el robot, para la reconstrucción de una escena 3-D o la identificación de la presencia de determinados objetos. A partir de esta información el robot puede tomar decisiones dependiendo de qué tipo de robot sea; si es un robot de transporte o de movimiento, puede evitar o dirigirse hacia estos objetos; o si es un robot industrial puede manipularlos según su función.

2.1.1.4 Medicina

Entre las aplicaciones de apoyo al diagnóstico médico se pueden destacar el análisis e interpretación de radiografías y tomografías, la adquisición de muestras menos dañinas para el paciente, intervención quirúrgica asistida, la obtención del entramado de vasos capilares, o nervios en un determinado tejido mediante la extracción de bordes, la detección de lesiones vasculares a partir de angiogramas renales, la diferenciación de tejidos sanos de tejidos cancerígenos o infectados por el color, la detección de cánceres de piel mediante técnicas de color y extracción de bordes etc.

2.2 Historia

- 1958: Para resolver los problemas que presentaba el desarrollo de la visión por computador, Frank Rosenblatt presento un nuevo algoritmo, el *perceptron*, el cual era una clase de red neuronal que clasificaba automáticamente las imágenes que contenían en ellas tanques camuflados en bosques, en comparación con las imágenes que mostraban los bosques solamente.

- 1960: Se introduce el procesamiento de imágenes y reconocimiento de patrones dentro de la inteligencia artificial.
- 1963: Para su tesis de doctorado en el MIT, Larry Roberts presenta la "Máquina de Percepción de sólidos tridimensionales". En su tesis, Larry Robert propuso la idea de extraer información geométrica tridimensional, a partir de vistas en dos dimensiones. Este hecho contribuyo significativamente; ya que estableció los cimientos del campo de la visión por computador.
- 1972: David Marr propuso su enfoque ascendente de comprensión de escena, el cual se basaba en la detección de bordes y la segmentación de imágenes. Este es considerado la contribución más influyente en la visión por computador.
- 1980-1990: Se utilizan conceptos de matemática, estadística y teoría de control enfocados en la visión artificial. También se desarrollan estudios para comprender la visión humana y como emular esa percepción artificialmente.
- 1990-2000: En este periodo de tiempo se da el nacimiento de la visión por computadora moderna. Se desarrollan una gran cantidad de métodos para abordar problemas de visión por computador, como clasificación de objetos, detección de objetos, segmentación, reconocimiento de rostros, etc. Además se desarrolla el uso de aprendizaje estadístico.
- 1999: Se desarrolla el algoritmo de extracción y representación de características SIFT, el cual se caracterizaba por ser una herramienta eficiente y precisa que permitía realizar una comparación precisa entre el mismo objeto en diferentes imágenes. Lo cual abrió las puertas al reconocimiento de objetos.

En la actualidad la visión por computador ha evolucionado hasta el desarrollo de redes neuronales, también conocidas como redes de aprendizaje profundas; las cuales son capaces de clasificar el contenido de una imagen de manera muy precisa. De hecho, se podría afirmar que las computadoras están llegando a un nivel humano de precisión en ciertas tareas.

2.3 Analogía entre visión humana y visión artificial

En cuanto al reconocimiento o clasificación de objetos se sabe que el cerebro humano todavía procesa las imágenes mejor que un sistema computacional robusto. El ser humano procesa la información visual en el espacio semántico, es decir, que extrae las características semánticamente significativas, tales como segmentos lineales, bordes, formas, etc. Pero aun con las técnicas actuales de procesamiento de información de las computadoras, este tipo de características no pueden ser detectados con facilidad por los ordenadores digitales. Las computadoras procesan la información visual en el espacio de datos formado por características de menor importancia que detecta, tales como color, textura etc. A pesar de esto, todo proceso de visión por computadora siempre se ha basado en la visión del ser humano y siempre tratara de emularla, por lo que ambas poseen ciertas similitudes. Pero este proceso no es sencillo, ya que intervienen muchos factores y partes del ser humano que deben de trabajar el conjunto.

La visión humana se realiza en cuatro fases:

- a. Percepción: Esta etapa es puramente óptica, donde la luz entra en el ojo humano atravesando órganos transparentes como la córnea, humor acuoso, cristalino y humor vítreo, donde se busca, sigue y enfoca la imagen. Esta etapa es comparable con la función de una cámara digital donde la luz atraviesa el lente de la cámara y un

sensor de captación de imagen descompone esos haces en colores rojo, verde y azul.

- b. Transformación: En esta etapa la energía luminosa llega a la retina, donde los conos y bastones, que son células foto receptoras sensibles a la luz, responsables de la visión a colores y la visión en baja luminosidad, se activan y transforman la luz en energía nerviosa. Estas células foto receptoras son similares a los sensores de captación de imagen de las cámaras digitales; los cuales transforman las señales de colores rojo, verde y azul de la luz que capta, en cargas eléctricas analógicas. Posteriormente la electrónica de la cámara transforma estos datos en formato digital, específicamente en un arreglo de matrices, de tal forma que el ordenador pueda reconocerlos e interpretarlos.
- c. Transmisión: Los impulsos nerviosos son transportados por medio del nervio óptico hacia la corteza cerebral. Esta etapa es comparada con cualquier método de transmisión de datos; como por ejemplo cable USB, Ethernet, memoria, etc. Que transmita los datos de la cámara digital hacia el ordenador o PC.
- d. Interpretación: En el caso de la visión humana, las señales o impulsos neuronales son interpretados por el lóbulo occipital del cerebro, lo que desencadena una serie de acciones o decisiones que se toman en base a lo interpretado. He aquí donde se presenta la mayor diferencia entre la visión humana y la visión por computador, ya que aunque la interpretación de la información de las imágenes es realizada por parte de los procesadores de las computadoras, no se ha llegado a desarrollar algún tipo de algoritmo o función que emule la capacidad de procesamiento del cerebro humano.

Por lo que se puede decir que todavía existe una gran brecha entre la capacidad de reconocimiento y procesamiento de imágenes del ser humano y la visión por computador. Ya que el ser humano tiene la capacidad de recibir e interpretar, en un tiempo muy corto, una gran cantidad de información proveniente de una imagen ya sea estática o dinámica, vista desde diferentes puntos de vista, en condiciones de luz cambiantes, etc. A diferencia de las capacidades limitadas de la visión por computadora, que puede emular procesos visuales elementales, tales como reconocimiento de bordes o cambios de estado, pero con tiempos de procesamiento elevados. Aunque, la visión por computador posee funciones que el ser humano no puede realizar o se le dificulta; por ejemplo, la medición de magnitudes físicas de forma precisa, tales como longitud, área, elongación, etc. Además la visión por computador presenta ventajas en tareas rutinarias o que requieran de gran concentración, como los procesos de control de calidad. Por lo que se puede concluir que los sistemas de visión por computador son un complemento a las capacidades visuales de los seres humanos.

2.4 Conceptos Básicos

2.4.1 Imagen digital

La imagen digital es el elemento primario del procesamiento digital de imágenes, de la cual se puede extraer información muy útil del entorno físico. La imagen es captada por medio de un proceso que involucra captura, muestreo, cuantificación y codificación. La información obtenida se convierte en información alfanumérica que cuantifica la intensidad de luz de la imagen, mediante la utilización de algoritmos matemático, y es representada visualmente por pixeles de un determinado tamaño o resolución. Las imágenes normalmente son representadas por medio de una matriz $I(x, y)$ de la siguiente forma:

$$I(x, y) = \begin{bmatrix} I(1,1) & I(2,1) & \dots & I(N,1) \\ I(1,2) & I(2,2) & \dots & I(N,2) \\ \vdots & \vdots & \ddots & \vdots \\ I(1,M) & I(2,M) & \dots & I(M,N) \end{bmatrix} \quad (2.1)$$

Donde M y N representan las dimensiones de la imagen.

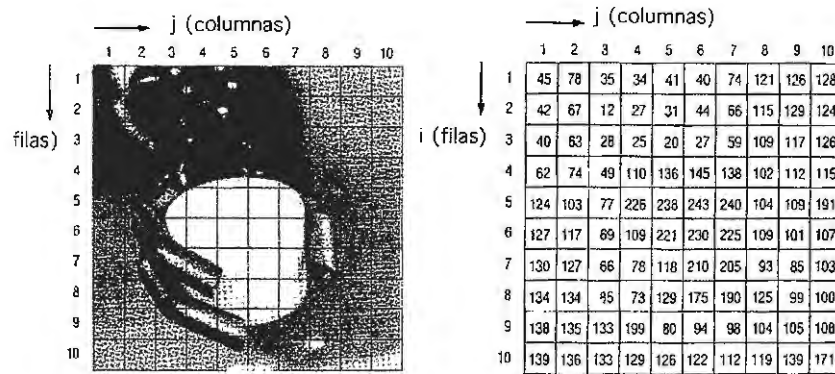


Imagen 2.1 - Imagen original y su correspondiente digitalización (Tomado de R. Ezkenazi, Video Signal Input, Robotics Age, Vol 3, No. 2, March/April 1981)

Las imágenes digitales se pueden presentar en los siguientes formatos y extensiones:

Tabla 2.1

Tipos de imágenes y sus extensiones

Tipo de Imagen	Extensión
TIFF	.tiff o .tif
JPEG	.jpg o .jpeg
GIF	.gif
BMP	.bmp
PNG	.png
XWD	.xwd

Nota: Recuperado de Procesamiento digital de Imágenes usando MATLAB & Simulink. Copyright 2010 por Alfaomega Grupo Editores, SA de C.V.

2.4.2 Imagen binaria

Una imagen binaria es aquella cuyos pixeles solo pueden tener el valor de 0 o 1 dependiendo del valor de umbral que se utilice en la segmentación de la imagen. Si el nivel de gris de un pixel es menor que el valor de umbral, se le asigna el valor de 0, que visualmente es representado por el color negro, y si es mayor, se le asigna el valor de 1, que representa el color blanco; teniendo como resultado una imagen formada por estructuras en blanco y negro.

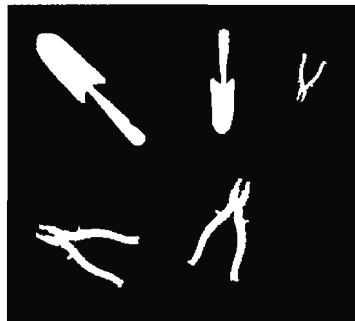


Imagen 2.2 - Imagen Binaria

2.4.3 Imagen RGB

Las imágenes RGB, como sus siglas en Inglés lo indican (Red, Green, Blue), se basan en la combinación de los colores primarios rojo, verde y azul para formar imágenes a colores, cuya iluminación y tono dependen de la intensidad de cada uno de estos componentes. El modelo RGB se basa en la adición de forma individual de cada componente de color en forma de capas que se traslapan, de tal forma que su combinación forme una sola imagen a color.

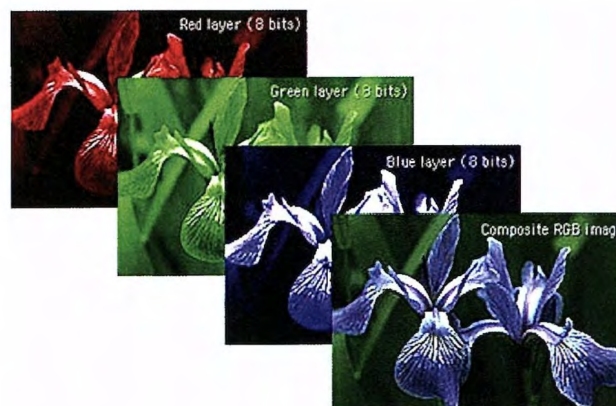


Imagen 2.3- Imagen RGB

2.4.4 Pixel

El termino Pixel se deriva de los vocablos en ingles *Picture Elements*, o elementos de imagen. El pixel es la unidad más pequeña que compone a una imagen digital, la cual contiene el nivel de luminosidad de un punto correspondiente de una escena digital codificado en formato binario. Esta codificación se le llama resolución, la cual contiene el número de bits que emplea un pixel para representar un nivel de intensidad. Así, 8 bits permiten manejar 256 niveles de gris, que es la resolución en escala de grises. Y la resolución de color, que se vale de tres bytes para establecer un color, cuya combinación permite tener 16, 777, 216 opciones de color.

2.4.5. Vecindad entre pixeles

La vecindad entre pixeles se refiere a la relación en cuanto a posición que tiene un pixel con los pixeles más cercanos a él. Existen dos tipos de vecindad, la vecindad 4-vecinos y la 8-vecinos.

La vecindad 4-vecinos es aquella en la cual un pixel central, denominado P, tiene como vecinos un pixel arriba, abajo, a la izquierda y derecha de él. La vecindad 8-vecinos es aquella que aparte de los pixeles correspondientes a la vecindad de 4-vecinos, también posee como vecinos 4 pixeles más en posición diagonal al pixel central.



Imagen 2.4 - Definición de la vecindad de un pixel P. (a) Vecindad 4, (b) Vecindad 8 (Tomado de E. Cuevas, D. Zaldívar y M. Pérez-Cisneros, "Procesamiento digital de imágenes usando MatLab & Simulink")

2.4.6 Conectividad

La conectividad entre píxeles es utilizada para establecer los límites de objetos en ciertas regiones dentro de una imagen, y se da cuando los píxeles pertenecientes a una región determinada cumplen con las condiciones de adyacencia que establecen; ya sea la vecindad a 4-vecinos o la vecindad a 8-vecinos y además sus respectivos niveles de gris cumplen con un criterio de similitud especificado o entran en un rango definido por un conjunto de valores. Por ejemplo, en la imagen binaria ilustrada en la Imagen 2.5, se puede determinar si existen uno o dos objetos dependiendo si se cumplen las condiciones de conectividad anteriormente descritas. En la Imagen 2.5 (a) se observan 2 objetos; ya que, aunque la imagen cumple con una vecindad 8-vecinos, los niveles de gris de ambos objetos no son los mismos. A diferencia de la Imagen 2.5 (b) donde se observa un solo objeto; ya que se cumple una vecindad a 8 y se tiene el mismo nivel de gris en todo el objeto.

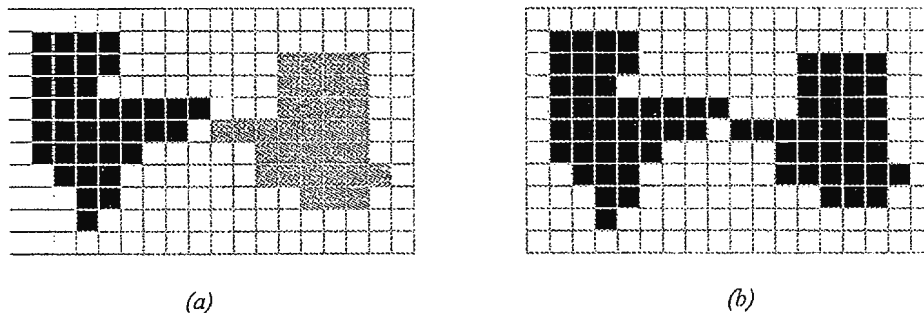


Imagen 2.5 - Definición de conectividad de un pixel. (a) Imagen que no cumple criterio de conectividad, (b) Imagen con conectividad-8 (Tomado de E. Cuevas, D. Zaldívar y M. Pérez-Cisneros, "Procesamiento digital de imágenes usando MatLab & Simulink")

2.4.7 Histograma

Es un tipo de distribución estadística que describe el número de veces que un pixel con determinado valor de intensidad, generalmente en escala de grises, se repite en una imagen determinada. Esta medida estadística muestra características importantes de una imagen, como el contraste, la iluminación y dinámica, que ayudan a determinar errores o problemas que se producen durante la toma de imagen y que deben de corregirse antes de que estas sean analizadas en las etapas posteriores de procesamiento para evitar consecuencias negativas.

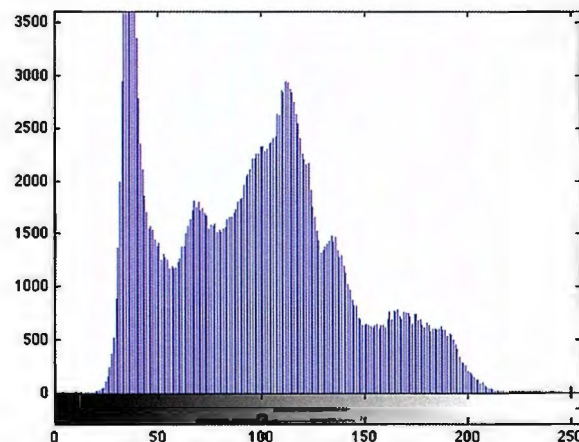


Imagen 2.6 - Imagen a escala de grises con su respectivo histograma (Tomado de M. Cardona, Material de clase de la Cátedra "Sistemas de Visión en Manufactura", Universidad Don Bosco)

2.4.8 Iluminación

La iluminación en una imagen digital es representada por la frecuencia de los valores de intensidad de los extremos izquierdo y derecho de un histograma, que representan los píxeles más oscuros y claros respectivamente. Por lo que los errores de iluminación en una imagen son fácilmente detectables en un histograma, ya que la región final o inicial de la escala de intensidad no es utilizada.

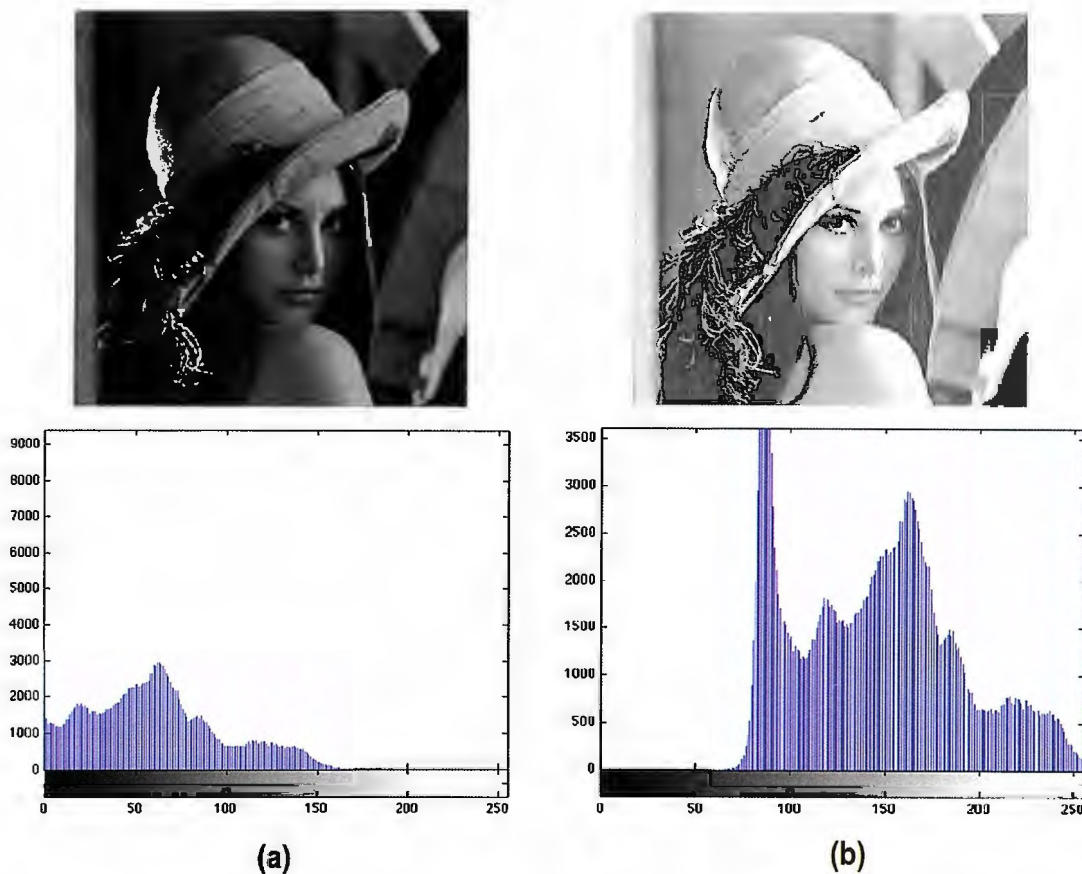


Imagen 2.7 - Errores de Iluminación, (a) Imagen con baja iluminación (b) con alta iluminación, (Tomado de M. Cardona, Material de clase de la Cátedra "Sistemas de Visión en Manufactura", Universidad Don Bosco)

2.4.9 Contraste

El contraste se refiere al campo de valores de intensidad que son utilizados en una imagen, y que se obtienen mediante la diferencia entre el valor máximo y mínimo de intensidad de los píxeles presentes en la imagen. Los problemas más comunes de contraste se deben a un rango de intensidad reducido en comparación al rango disponible de niveles de gris; el cual puede mejorarse mediante el re-escalamiento de la intensidad de cada píxel al aplicar una función específica que relacione la intensidad de los píxeles de la imagen de entrada, y la convierta en una imagen de salida con un mayor rango de intensidades y un mejor contraste.

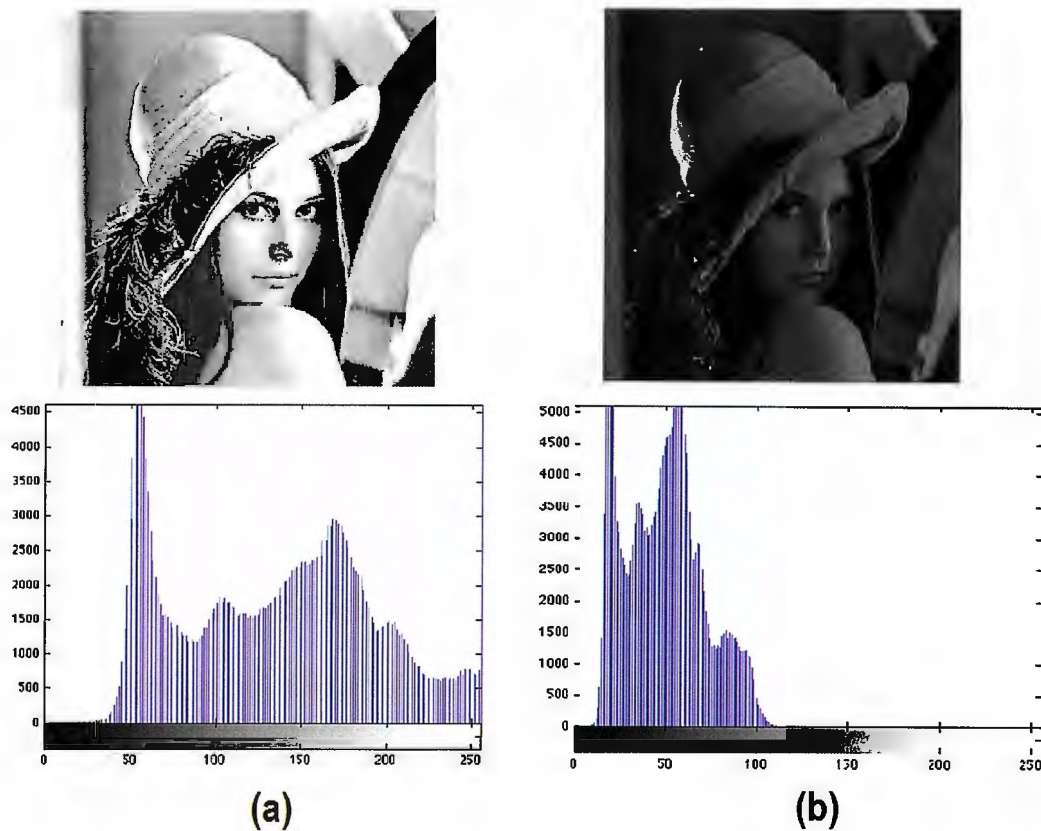


Imagen 2.8 - Imágenes con diferentes contrastes y su respectivo histograma (a) aumento de contraste, (b) reducción de contraste (Tomado de M. Cardona, Material de clase de la Cátedra "Sistemas de Visión en Manufactura", Universidad Don Bosco)

2.4.10 Dinámica

La dinámica se refiere a la cantidad de píxeles de diferente intensidad que contiene una imagen en particular. El mejor de los casos se presenta cuando se utiliza el rango completo de valores de intensidad disponibles en una imagen. LA dinámica está directamente relacionada con la calidad de la imagen; ya que, una dinámica alta reduce el riesgo de que la imagen pierda calidad a través de las siguientes etapas de procesamiento.

2.4.11 Etiquetado de Objetos en Imágenes binarias

El etiquetado de objetos, o mejor conocido como “*region labeling*”, es una técnica de procesamiento digital de imágenes que busca aislar cada uno de los objetos contenidos en una imagen con el objetivo de determinar el número de objetos presentes y describir e interpretar sus características estructurales. Esta técnica se basa en un algoritmo que consta de dos partes: 1) el etiquetado temporal de los objetos y 2) la solución de las colisiones entre etiquetas que pertenecen al mismo objeto.

En el primer paso, se recorre la imagen de izquierda a derecha y de arriba hacia abajo, asignándole en cada recorrido a cada pixel una etiqueta temporal que depende del tipo de vecindad presente; ya sea 4-vecinos u 8-vecinos.

En el segundo paso, se busca resolver las colisiones entre etiquetas temporales que se producen en el paso 1. Las colisiones de etiquetas se producen cuando dos o más pixeles que se agrupan para formar un mismo objeto poseen etiquetas diferentes. Este segundo paso posee un mayor grado de complejidad que el primero; ya que, dos etiquetas que pertenecen tentativamente a dos objetos diferentes, pueden estar conectadas entre sí por un tercer objeto. Al final de este proceso se tiene un objeto representado por pixeles que poseen el mismo valor de etiqueta.

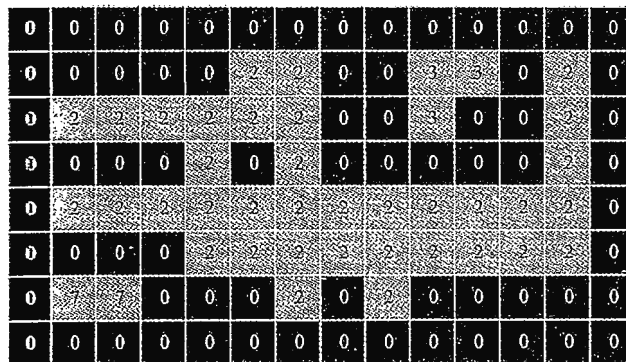


Imagen 2.9 - Resultado del proceso de etiquetado de un objeto. (Tomado de E. Cuevas, D. Zaldívar y M. Pérez-Cisneros, “Procesamiento digital de imágenes usando MatLab & Simulink”)

2.4.12 Características de objetos binarios

Las características de los objetos binarios son medidas cuantitativas que describen de forma precisa los objetos binarios que se encuentran en una imagen. Estas pueden ser: perímetro, área, *bounding box*, centroide, momento de inercia, etc. Dichas características dependen directamente de los píxeles que conforman a cada objeto; ya que, muchas de ellas pueden ser obtenidas mediante el cálculo directo de dichos píxeles. Además, para describir concretamente a un objeto, sus características se agrupan en un vector que contiene a todas y cada una de ellas.

2.4.13 Perímetro

El perímetro de un objeto binario se refiere a la longitud de su contorno exterior. En el caso de una vecindad 8- vecinos los movimientos horizontales y verticales del contorno tienen una distancia de 1, mientras que los diagonales de $\sqrt{2}$.

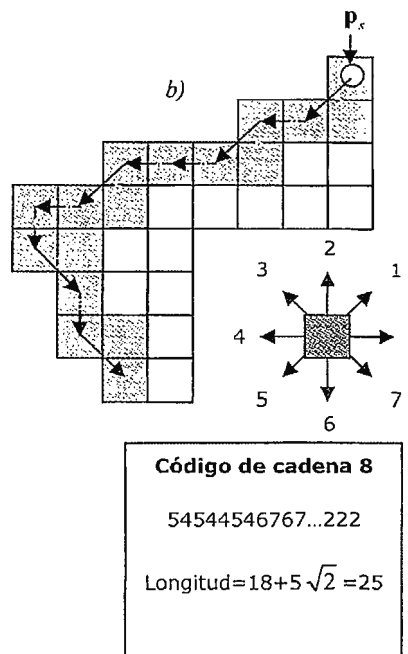


Imagen 2.10 - Perímetro basado en un código cadena 8-vecinos (Tomado de E. Cuevas, D. Zaldívar y M. Pérez-Cisneros, "Procesamiento digital de imágenes usando MatLab & Simulink")

2.4.14 Área

El área de un objeto binario es simplemente el número de píxeles que lo conforman.

2.4.15 Bounding Box

Es el cuadrado mínimo que es capaz de contener al objeto binario.

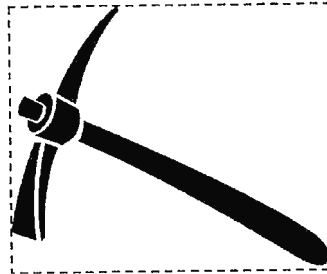


Imagen 2.11 - Bounding Box (Tomado de E. Cuevas, D. Zaldívar y M. Pérez-Cisneros, "Procesamiento digital de imágenes usando MatLab & Simulink")

2.4.16 Momento de Inercia

El momento de inercia es una característica que permite determinar el ángulo de orientación de un objeto binario en la imagen.

2.5 Sistema de Procesamiento digital de imágenes

El procesamiento digital de imágenes es una tecnología asociada a las ciencias de la computación y la inteligencia artificial, que se define como la operación de imágenes bidimensionales haciendo uso de un computador. Por tanto, al conjunto de elementos de software y hardware que se utilizan para procesar una señal visual, como una imagen digital, se le conoce como sistema de procesamiento digital de imágenes; el cual consta de las siguientes fases:

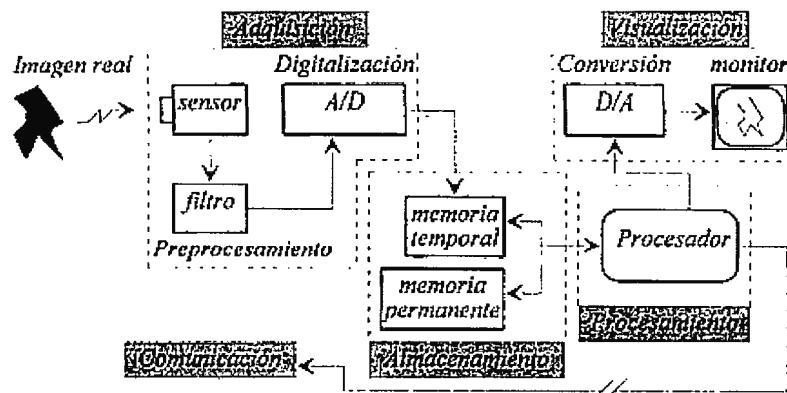


Imagen 2.12 - Configuración básica de los componentes de un sistema de procesamiento digital de imágenes (Tomado de Gianfranco Passariello, Fernando Mora (1995). Imágenes Médicas: Adquisición, Análisis, procesamiento e Interpretación)

2.5.1 Adquisición de la Imagen

En la fase de adquisición de la imagen se deben de utilizar 3 componentes de hardware para captar una imagen digital de la mejor manera. El primero es algún tipo sensor que registre cualquier tipo de radiación electromagnética, como: luz visible, rayos X, rayos ultravioletas, etc. Y la transforme en una señal eléctrica proporcional al nivel de energía percibido. El segundo es un dispositivo de filtraje que proteja la señal eléctrica de cualquier tipo de alteración en su información. Y el tercero se trata de un convertidor análogo-digital, el cual transforma la señal eléctrica en datos binarios con los que trabajara el software de procesamiento y almacenaje.

2.5.2 Almacenamiento

Esta fase utiliza dos tipos de memoria: la memoria de alta velocidad o temporal y la memoria lenta o permanente. La memoria temporal únicamente sirve como soporte para la fase de procesamiento y como mecanismo de transición entre la digitalización de la imagen y la memoria de almacenamiento permanente. Este último tipo de memoria debe de tener una gran capacidad de almacenamiento; ya que, es en esta donde se guardan tanto las imágenes originales las resultantes de la fase de procesamiento.

2.5.3 Procesamiento

Es la fase donde a través de algoritmos bien estructurados se desarrolla un software capaz de extraer información útil de las imágenes captadas. Este proceso es realizado por un procesador, ya sea un computador o un procesador específico para el tipo de operación, el cual accede al bloque de almacenamiento y realiza los cálculos requeridos sobre la memoria temporal para posteriormente guardar la información resultante en la memoria permanente del sistema.

Cabe resaltar que el procesamiento no es una acción general del sistema; si no que se caracteriza por ser específico para cada aplicación, ya que el software utilizado en cierta aplicación no será funcional en otra.

2.5.4 Visualización

En esta fase final se presentan las imágenes al usuario a través de una interfaz, que generalmente son monitores de video. Para esta fase se utiliza un convertidor digital-análogo para convertir la información digital en una señal analógica compatible con los dispositivos de visualización. Terminando así con todas las fases del procesamiento digital de imágenes.

Capítulo III

Control de calidad del azúcar

3.1 Calidad dentro de la industria alimenticia.

3.1.1 Historia

Al igual que muchas cosas la gestión de calidad presente dentro de la industria alimenticia nace por la necesidad de saber si los productos son buenos, confiables y del gusto de las personas, mediante técnicas de control, simple inspección o chequeo.

El proceso evolutivo comienza con la producción artesanal de los cristales de azúcar donde las mismas personas se encargaban de revisar la calidad del producto; conforme el progreso seguía avanzando durante la revolución industrial la demanda del producto fue creciendo dando lugares a nuevos trabajos como supervisores y expertos que estudiaran a detalle el producto.

La industria primitiva de los productos alimenticios fue creciendo cada vez más, de tal manera que fueron requeridas la implementación de técnicas estadísticas; que dieron como resultado el nacimiento del control estadístico de calidad. Las ventajas del nuevo modelo era el uso de técnicas de muestreo para analizar una población, en lugar de analizar el 100% como era tradicionalmente.

Posterior a la segunda guerra mundial los procesos de calidad se enfocaron en reducir el mayor número de pérdidas posibles y generar productos que satisficiera la necesidad del consumidor; por lo que se empezó dentro de la industria a analizar de manera individual cada uno de los procesos que conllevan a la extracción de materia prima hasta el procesamiento de los alimentos dentro de la planta.

Una década antes del cambio de siglo, la calidad se vio en la necesidad de integrar un nuevo término: servicio de calidad total; el cual no sólo pone a prueba los procesos de fabricación y recolección del producto, sino que de igual manera la presentación de los productos finalizados con el fin de brindar un producto que no solo fuese bueno, sino atractivo para todas las personas.

Finalmente con el progreso de los avances tecnológicos, poco a poco se observa que se van simplificando trabajos de supervisión complicados gracias a máquinas que permiten visualizar el estado de seguridad de los sistemas y determinan si existen presencias de imperfecciones no deseadas dentro de los lotes de alimento, al igual que procesos que permiten mejorar la calidad de vida y cualidades de los productos alimenticios.

3.1.2 Tipos de calidad presentes en la industria alimenticia

La calidad no solo se limita a un área, existen distintos tipos de estándares de calidad para determinar si un producto es totalmente bueno; aunque el tipo de calidad en el que se profundizará a lo largo del desarrollo del trabajo es la calidad estructural o morfológica.

A continuación se mostraran los distintos tipos de calidad analizados comúnmente en la industria alimenticia.

- Calidad higiénica y sanitaria: Consiste en una serie de normas que permiten garantizar la seguridad y salubridad de los productos alimenticios en todos los procesos que lleva su producción.
- Calidad bromatológica: Consiste en una serie de parámetros que nos permiten determinar si el producto alimenticio tiene valores nutricionales, toxicidad, contaminantes y cumple las características físicas deseadas.
- Calidad organoléptica: Consiste en una serie de normas que permiten garantizar que el producto alimenticio no generará problemas, rechazo o

disgustos a los sentidos de los consumidores (tacto, olfato, vista, gusto y el oído).

- Calidad ética o emocional: Consiste en una serie de propiedades de importancia creciente que por lo general son de interés para el consumidor para decidirse a adquirir el producto como: prácticas ecológicas, conservación de los recursos naturales, etc.
- Calidad tecnológica: Consiste en el uso de métodos que nos permiten prolongar la vida de los productos, aumentar la confianza de los productos y obtener productos con mejores cualidades mediante el uso de herramientas de innovación y tecnología.
- Calidad económica: Consiste en la capacidad de mantener la estabilidad de los precios del producto alimenticio cuando existen fluctuaciones o escasez debido a factores ambientales como sequías, tormentas y temporales fríos.
- Calidad morfológica: Consiste en una serie de parámetros que permiten determinar si las características geométricas y morfológicas de nuestro producto alimenticio cumplen con los estándares esperados.
- Calidad de servicio: Consiste en una serie de normas que ayudan a establecer los estándares del producto para que sea uniforme y encargado de generar el factor psicológico de seguridad sobre los consumidores.

3.1.3 Normas que rigen la calidad de los alimentos.

Existen en la actualidad una serie de modelos de gestión de calidad que nos permiten decir efectivamente si el producto alimenticio final cumple o no con los estándares de calidad, estos van desde las normas ISO hasta normas realizadas por asociaciones del rubro alimenticios, dichos modelos son los siguientes.

- *Norma ISO 22000.2005 Gestión de la Inocuidad de los alimentos:*
Este modelo define y especifica los requerimientos para desarrollar e implementar un sistema de gestión de inocuidad de los alimentos, con el fin de lograr una armonización internacional que permita una mejora de la seguridad alimentaria durante el transcurso de toda la cadena de suministro [1].

- *Análisis de Peligros y Puntos de Control Críticos (APPCC):*
Este modelo posee un enfoque sistemático en la identificación, evaluación y control de aquellas etapas en el manufacturado de los alimentos que son críticas para la seguridad del producto [2].

El APPCC identifica los riesgos de los procesos de producción que conducen hacia productos no aptos para el consumo, además de diseñar medidas que reduzcan esos riesgos hasta niveles aceptables. [3]

- *British Retail Consortium (BRC):*
Este modelo fue creado bajo una serie de normas regidas por el Consorcio Británico de Minoristas, elaboradas para todas aquellas compañías proveedoras a empresas minoristas de productos agroalimentarios. Ha sido desarrollado para ayudar a los minoristas en cuanto al cumplimiento de sus obligaciones legales y de protección al consumidor, mediante la creación de una base común para la certificación de empresas [4].

- *International Featured Standards (IFS):*
Este modelo fue creado por asociaciones alemanas, francesas e italianas con el fin de desarrollar un estándar de gestión de la calidad y seguridad de los alimentos para marcas propias, que define los requisitos para gestionar la seguridad de los alimentos y el nivel de calidad de los productores y elaboradores.

Puede utilizarse en todas las etapas de fabricación que guarden relación con la producción, la elaboración y preparación de los alimentos y

distribución. Además, afecta a los elaboradores y procesadores de productos alimentarios bajo marcas de minoristas o mayoristas [5].

3.2 Objetivo y finalidad del uso de métodos estadísticos para el control de calidad

3.2.1 ¿Qué es la estadística?

La estadística posee su origen de la palabra francesa *statistique* que significa préstamo y procede del latín *statisticus* que significa relativo al estado, ya que inicialmente la estadística era solamente aplicada para fenómenos relacionados con el estado como son los inventarios y censos. Posteriormente en el siglo XIX nace el significado moderno de estadística bajo la palabra *statistics*, que ya hace referencia a la manipulación de datos numéricos para obtener información.

La estadística moderna se puede describir con las siguientes definiciones:

- Ciencia que utiliza conjuntos de datos numéricos para obtener, a partir de ellos, inferencias basadas en el cálculo de probabilidades.
- Estudio que reúne, clasifica y recuenta todos los hechos que tienen una determinada característica en común, para poder llegar a conclusiones a partir de los datos numéricos extraídos.
- Conjunto de los datos o los hechos recogidos y clasificados.
- Censo o recuento del número de habitantes de un lugar, de los recursos naturales e industriales, del tráfico o de cualquier otra manifestación relacionada con un estado, una provincia, una localidad, etcétera.
- Ciencia formal y una herramienta que estudia el uso y los análisis provenientes de una muestra representativa de datos, busca explicar las correlaciones y dependencias de un fenómeno físico o natural, de ocurrencia en forma aleatoria o condicional.

3.2.2 Presencia de la estadística en el control de calidad

La calidad a medida que fue evolucionando fue requiriendo formas de saber si los resultados obtenidos tenían correlación y cumplían con las características deseadas, por lo que recurre a la estadística y a sus múltiples formas de analizar, visualizar y presentar información de manera detallada y ordenada. Algunas ventajas de análisis estadístico son las siguientes:

3.2.2.1. Detección de defectos

Una de las principales razones de la integración del control de calidad estadístico en la industria alimenticia fue la detección de productos defectuosos, ya que mediante la combinación de distintas herramientas puede detectar los productos que incumplen con una característica específica de calidad respecto a un límite especificado.

De igual manera mediante los resultados obtenidos del control de calidad estadístico se puede determinar el posible origen de la formación de productos defectuosos, y actuar según sea necesario para reducir la cantidad de productos de baja calidad, ya que representan pérdidas para la industria.

Los defectos en los productos alimenticios pueden tener distintos atributos por lo que se pueden clasificar de la siguiente manera:

- Defectos críticos: Son aquellos que violan leyes, agreden al consumidor o hacen inservible al producto.
- Defectos mayores: producen una disminución en el correcto funcionamiento o utilización del producto y es notado por el consumidor.

- Defectos menores: producen una disminución leve en el correcto funcionamiento o utilización del producto, probablemente no lo note el consumidor. Pero si lo nota, el personal calificado de producción y de control de calidad.

3.2.2.2. Inclusión de técnicas de muestreo

La integración del control de calidad estadístico trajo una herramienta que facilita y agiliza los procesos de control de calidad, la cual es el muestreo. El muestreo consiste en tomar una fracción de la población total para realizarle una serie de análisis cuyos resultados brindarán un estimado bastante aceptable de los resultados que se obtendrían si se analiza la población completa.

La toma de una muestra nos permite analizar de manera más rápida una población a diferencia de analizar el 100% de la población, permitiendo analizar un mayor número de muestras en el menor tiempo posible, de igual manera al analizar solo una fracción de los resultados los procesos de análisis son más rápidos, obteniendo reportes de resultados con mayor rapidez y permitiendo a la industria actuar ágilmente acorde a los resultados de las características obtenidas en los reportes.

3.2.2.3. Análisis de múltiples variables

El control de calidad estadístico permite el análisis simultáneo de distintas variables o de características que conforman el producto alimenticio a analizar, dotándolos así de más información sobre las características finales de nuestro producto a diferencia de sólo analizar de manera general el producto final.

Una mayor cantidad de variables permiten obtener un mejor análisis de resultados, y al obtenerse mejores resultados se puede atacar de mejor manera las razones por la cual existen variaciones en la formación de los productos alimenticios, haciendo más efectiva la gestión y control de calidad.

Se resalta que cada característica es analizada de manera individual puesto que las imperfecciones en nuestro producto final pueden manifestarse en distintas formas y tener distintos efectos, algunas características notables en la industria alimenticia de cristales y granos son las siguientes:

- Forma: hace referencia a la geometría del producto alimenticio.
 - o Bordes: Normales, Cuadrados, Faba de colmillo o Truncados.
 - o Longitudinales: Normales, Anchos o Alargados.
 - o Transversales: Secciones llenas o secciones semi-llenas
 - o Concavidades: Normales, Rectilíneas o Arriñonadas.
 - o Tamaño: Pequeños, Medianos, Grandes.

- Color: hace referencia a la tonalidad y pigmentación del producto alimenticio.
 - o Coronación: Presencia de corona o ausencia de corona.
 - o Tegumentario: Presencia de tegumento o Ausencia de tegumento.
 - o Veteado: Presencia de vetas o Ausencia de vetas.
 - o Brillo: Cuerpo brillante, cuerpo sin brillo o cuerpo mate.
 - o Coloración: Sin color o con color (característico de cada producto).

- Defectos: hace referencia a características que sobresalen de las características convencionales pero que no son aceptables.
 - o Superficies manchadas.
 - o Superficies rugosas no deseadas.
 - o Pigmentación de colores no deseados sobre la superficie.

Existen una gran variedad de características notables mediante las técnicas de control de calidad estadístico, las cuales cambian dependiendo del producto alimenticio a analizar.

3.2.3. Importancia del control estadístico de procesos.

Habiendo conocido la importancia de la estadística dentro de los procesos de control de calidad estadístico, se debe definir ahora el control de cada uno de los procesos que conforman nuestra producción, ya que para obtener un buen producto se debe realizar un análisis estadístico del control de procesos.

El control estadístico de procesos permite usar criterios objetivos para distinguir variaciones en las distintas fases de la producción de los alimentos, su principal tarea reside monitorizar el proceso y determinar cada una de las variaciones mediante la recopilación de datos de mediciones en diferentes sitios presentes en el proceso.

Gracias al control estadístico de procesos se pueden detectar y corregir variaciones en el proceso que puedan afectar a la calidad del producto alimenticio final, reduciendo así la cantidad de pérdidas, de igual manera el control de procesos permite la detección precoz de defectos y prevención de problemas antes que sea demasiado tarde.

3.3 Principios estadísticos empleados para el proceso de control de calidad

3.3.1 Herramientas estadísticas de muestreo

Durante el proceso de desarrollo se abordaron dos métodos de muestreo para la realización de la investigación, el muestreo estratificado y el muestreo opinático, los cuales se explicarán a continuación.

➤ *Muestreo estratificado:*

El muestreo estratificado es una técnica de muestreo probabilístico en donde primero se toma una muestra aleatoriamente de una población y luego esta muestra se divide en submuestras o estratos los cuales poseen

una característica común que los une o asemeja, algunas de estas características pueden ser el color, forma o tamaño.

Las características del muestreo estratificado son las siguientes:

- o Permite resaltar un subgrupo específico dentro de la población.
- o Observar relaciones entre dos o más subgrupos.
- o Se puede probar de forma representativa hasta a los subgrupos más pequeños y más inaccesibles de la población.
- o Precisión estadística más elevada en comparación con el muestreo aleatorio simple.

➤ *Muestreo opinático aleatorio:*

El muestreo opinático aleatorio es la combinación de dos técnicas de muestreo una técnica de muestreo probabilístico y una técnica de muestreo no probabilístico donde en primera instancia se extrae de la población una muestra aleatoria, y posteriormente de la muestra se extraen los elementos que cumplen con una característica específica ya sea geometría, tamaño o color.

Las características del muestreo opinático aleatorio son las siguientes:

- o Permite resaltar un subgrupo específico dentro de la población.
- o Toma en cuenta solamente los miembros de interés para la investigación.
- o Se caracteriza por tener un esfuerzo deliberado por obtener una muestra lo más representativa posible.

3.3.2 Herramientas estadísticas de exposición de resultados

Para denotar los resultados obtenidos se requiere de herramientas y técnicas que faciliten la comprensión, y con el uso de los resultados obtenidos realizar el debido análisis para la toma de decisiones dentro de la industria.

➤ *Diagrama de dispersión:*

Un diagrama de dispersión es una herramienta matemática basada en el plano de coordenadas cartesianas que se encarga de mostrar la relación existente entre dos variables o incluso la relación de más variables para obtener una tendencia de la información; para el control de calidad se puede utilizar como herramienta estadística para determinar la dispersión de una partícula u objeto referente a un parámetro de control definido, permitiendo tener un mapa general de las variaciones de los miembros de la muestra respecto al parámetro ideal.

Los diagramas de dispersión se emplean con los siguientes fines:

- o Observar la relación entre dos variables
- o Visualizar cambios anómalos en las características de la muestra.
- o Analizar resultados mediante el uso comparaciones.

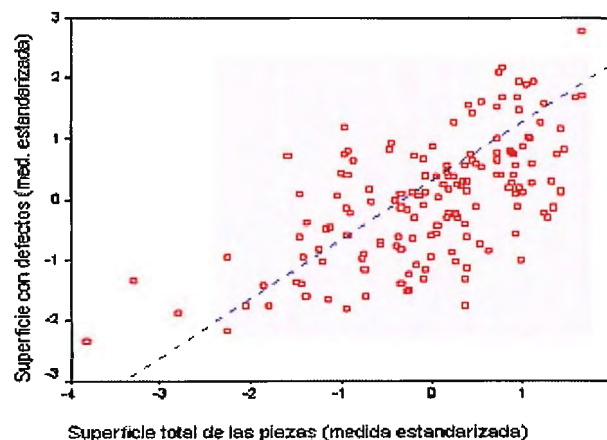


Imagen 3.1 - Ejemplo de diagrama de dispersión. (Tomada de <http://www.ub.edu/>)

➤ **Histograma:**

Consiste en un gráfico de barras que tratan de dar un panorama general de una característica específica de la población, en este caso se usa un histograma de distribución normal el cual debe ir orientado a obtener la forma de una campana de gauss.

Los histogramas se emplean con los siguientes fines:

- o Aumentar la calidad de los procesos realizados en la industria.
- o Visualizar el comportamiento de una variable de interés
- o Identificar anomalías examinando la tendencia de la gráfica.

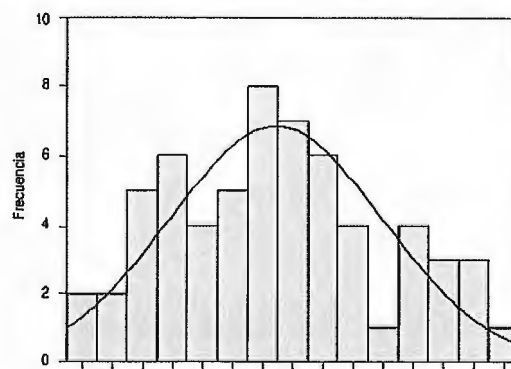


Imagen 3.2 - Ejemplo de histograma. (Tomada de <http://www.astropt.org>)

➤ **Curva de normalización:**

Consiste en un gráfico críticamente amortiguado que va disminuyendo su oscilación a medida que se aumenta el número de muestras tomadas, hasta llegar al punto crítico de homogenización donde los resultados se estabilizan y sin importar el número de datos que se agreguen no se cambiara la tendencia de la curva.

Las curvas de normalización se emplean con los siguientes fines:

- o Determinar la variación de los datos a medida que se aumenta la cantidad de muestras realizadas.

- o Representar gráficamente el teorema de límite central que busca transformar variables aleatorias en una distribución normal a medida que se aumenta el número de muestras.

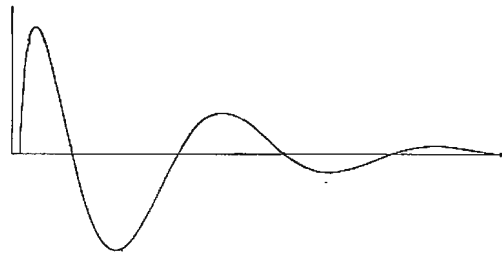


Imagen 3.3 - Ejemplo de curva de normalización. (Tomado de <http://rci.cujae.edu.cu/>)

3.3.3 Herramientas estadísticas utilizadas para crear la tabla de frecuencias

Una vez obtenidos todos los datos de la herramienta estadística de muestreo se deben organizar y ordenar los datos, de tal manera que sean fáciles de comprender y posteriormente analizar, por lo que a continuación se enunciarán y explicarán los términos necesarios para crear la tabla de frecuencias donde será organizada toda la información.

3.3.3.1. Medidas de dispersión.

- **Rango:** Es el valor existente entre la resta del valor máximo y el valor mínimo presentes en una serie de datos, y denota el barrido de valores que se pueden encontrar dentro de la serie de datos, el valor del rango puede ser muy grande como muy pequeño así como también puede ser decimal.
- **Varianza:** Es una medida estadística de dispersión que mide la variación de los valores de la serie respecto a la media aritmética y es el resultado del cuadrado de la desviación típica.

- Coeficiente de variación de Pearson: Es una medida estadística de dispersión cuya función es denotar la homogeneidad de los valores de una serie de datos, y se obtiene de la división entre la desviación típica y la media aritmética, entre el valor del coeficiente de variación de Pearson sea más grande significa que los datos están muy distantes entre sí, sin embargo si el coeficiente de variación de Pearson es pequeño significa que los datos están compactados , entre más pequeño sea el coeficiente de variación mayor será la calidad de la muestra.
- Diámetro medio como medida de dispersión: Es una medida estadística cuya función es determinar el diámetro representativo de una serie de datos, mostrando el diámetro que tenderá cada miembro de la serie de datos si todos los datos fueran iguales.
- Desviación típica: Es una medida estadística de dispersión que mide la cercanía de los valores de la serie respecto a la media aritmética esperando una tendencia normal, y es el resultado de la raíz cuadrada de la varianza.

3.3.3.2. Medidas de tendencia central

- Media aritmética: La media aritmética es el valor promedio resultante de la suma de todos los valores perteneciente a una serie de datos dividido entre el número de datos, y es el valor característico de la serie de datos.
- Mediana: La mediana representa el valor central de todos los valores de una serie de datos ordenados de menor a mayor, y ejemplifica el valor medio de la serie de datos.

3.3.3.3. Elementos de la tabla de frecuencia

- Clase: La tabla de frecuencia tiene una serie de clases, clases que representan a cada una de los datos que se encuentra dentro de sus límites, el número de clases presentes en la tabla de frecuencia se obtiene mediante la regla de Sturges.
- Punto medio de clase: El punto medio de la clase es el valor resultante entre la suma de los límites superior e inferior de la clase divididos entre dos, la finalidad de obtener el punto medio de clase es facilitar los cálculos de varianza y desviación típica.
- Ancho de clase: Es el rango de valores que abarca una clase, y se obtiene mediante la división entre el rango y el número de clases, también se puede calcular restando el valor del límite superior menos el valor del límite inferior.
- Límite inferior de clase: Es el valor mínimo perteneciente a una clase específica y sirve para marcar las fronteras entre la clase a la que pertenece y la clase anterior.
- Límite superior de clase: Es el valor máximo perteneciente a una clase específica y sirve para marcar las fronteras entre la clase a la que pertenece y la clase siguiente.
- Frecuencia absoluta: Es el número de datos que se encuentran en una clase de datos agrupados específica dentro de una serie de datos.
- Frecuencia absoluta acumulada: Es la suma progresiva de las frecuencias absolutas hasta que se hayan contado a cada uno de los miembros de la muestra o población.

CLASE 1		X_i	F_i	$X_i \cdot F_i$
0.360591911	0.440583286	0.400587598	2	0.801175197
CLASE 2				
0.440583286	0.520574662	0.480578974	5	2.40289487
CLASE 3				
0.520574662	0.600566037	0.56057035	2	1.121140699
TOTAL			25	15.61408625

Imagen 3.4 - Ejemplo de tabla de frecuencia para datos agrupados realizada en Excel

3.4 Análisis cualitativo y descriptivo del proceso.

3.4.1 Desarrollo de selección de la muestra a analizar

Los cristales de azúcar son un producto alimenticio que se genera en grandes cantidades, comúnmente lotes que pueden llegar fácilmente a decenas de quintales, por lo que se debe determinar cuál es el mejor procedimiento para obtener la muestra de toda la población.

Debido a que los cristales al ser diminutos y obtenidos en grandes proporciones sus poblaciones son extensas, y tomar una muestra altamente representativa es muy difícil, se propuso el siguiente plan de muestreo que permitirá tomar de manera efectiva la mejor muestra, el cual se divide en tres partes: selección del método de muestreo, asignación del nivel de inspección y tamaño de la muestra.

3.4.1.1 Selección del método de muestreo a utilizar

El control de calidad se realiza mediante una serie de pasos que van otorgando gradualmente la información necesaria para el análisis de calidad, el cual empieza con la selección del método estadística de muestreo; la selección de la herramienta se disputo entre los métodos de muestreo estratificado y el muestreo opinático aleatorio, por lo que se explicaran las ventajas de cada uno de los métodos y se explicara la razón de su elección.

➤ *Muestreo estratificado*

Método de muestreo seleccionado preliminarmente debido a que los cristales de azúcar deben tener una geometría y un tamaño específico, y este método de muestro nos permite tomar una muestra aleatoria de azúcar y posteriormente dividirla en tres estratos principales los cuales podrían ser cristales de baja calidad, cristales de media calidad y cristales de alta calidad u otra clasificación de los estratos podrían ser: cristales pequeños, cristales medianos y cristales grandes.

De igual manera permite analizar de forma individual cada estrato, Pudiendo identificar los rasgos distintivos de cada sub población o estrato y posteriormente reunir toda la información de cada estrato y obtener un resultado de toda la muestra tomada en un principio, permitiendo determinar posibles razones de la presencia de cristales de baja calidad y en el otro caso de clasificación presencia de muchos cristales grandes.

➤ *Muestreo opinático aleatorio*

Método de muestreo seleccionado preliminarmente debido a que los cristales deben cumplir solamente con una característica especial de interés, dividiendo la muestra en dos sub poblaciones los cristales que poseen la característica especial seleccionada y los que no la poseen, pudiendo dividir los cristales en cristales de alta calidad y cristales de baja calidad, otra característica de división podría ser cristales de tamaño regular y cristales de tamaño irregular.

De igual manera permite analizar de forma efectiva únicamente la sección de la muestra que es de mayor interés y así poder encontrar posibles fallas dentro del proceso de producción que hacen bajar la calidad de los cristales y determinar la calidad de la sub población que

cumple con la característica seleccionada; posibles razones de la presencia de anomalías en los tamaños de los cristales y la proporción de cristales que obtienen la dimensión esperada para determinar la eficiencia del proceso.

Luego de la puesta en común de ambos métodos de muestreo se seleccionó el método de muestro opinático aleatorio debido a su capacidad de extraer de la muestra los elementos que son de interés para el estudio y el desarrollo del control de calidad.

3.4.1.2 Asignación de nivel de inspección

La inspección de los cristales de azúcar es parte importante del aseguramiento de calidad, gracias a que permite estandarizar los procedimientos para obtener una muestra confiable a partir del lote de producción, ya que nos dictan la rigurosidad que se debe tomar con la muestra a la hora de analizar dependiendo de las características obtenidas.

El nivel de inspección se basará en el *military standard* (MIL-STD) 105E, donde existen tres niveles de inspección primarios o generales:

- o Nivel I: Se utiliza cuando se busca reducir desechos en la producción.
- o Nivel II: Se utiliza siempre que se busque una producción normal.
- o Nivel III: Se utiliza cuando puede desechar una mayor cantidad de producto.

De igual manera existen cuatro niveles de inspección secundaria o especial los cuales son S-1, S-2, S-3 y S-4 cuya función es reducir el tamaño de la muestra cuando sea necesario, que van desde la reducción a una muestra muy pequeña S-1 hasta una muestra considerable S-4.

Debido a que la primera etapa del muestreo es aleatoria se deben tomar en cuenta a todos los cristales por igual, por lo que se utilizara el

Nivel II de inspección, ya que el Nivel I no discrimina para nada a la muestra y el Nivel III discrimina demasiado a la muestra.

Posteriormente para el muestreo opinático que se realiza mediante fotos obtenidas del muestreo aleatorio, el nivel de inspección será el S-2, que es uno de los cuatro niveles especiales de inspección según el MIL-STD, ya que para la toma de fotos se requiere una muestra de tamaño menor que la obtenida mediante el muestreo aleatorio, cabe recalcar que para que la muestra sea representativa la calidad de la foto debe ser lo más clara posible.

3.4.1.2 Tamaño de la muestra.

Para el tamaño de la muestra se parte de los resultados obtenidos en la inspección realizada basada en el MIL-STD, por lo que se recurre a las siguientes tablas que nos denotan el tamaño de la muestra dependiendo del tamaño del lote del que provenga.

La primera tabla entregara un código alfabético dependiendo del tamaño del lote seleccionado, tal como se muestra a continuación.

Tamaño de lote	Niveles de inspección especial				Inspección de niveles generales		
	S-1	S-2	S-3	S-4	I	II	III
2 a 8	A	A	A	A	A	A	B
9 a 15	A	A	A	A	A	B	C
16 a 25	A	A	B	B	B	C	D
26 a 50	A	B	B	C	C	D	E
51 a 90	B	B	C	C	C	E	F
91 a 150	B	B	C	D	D	F	G
151 a 280	B	C	D	E	E	G	H
281 a 500	B	C	D	E	F	H	J
501 a 1200	C	C	E	F	G	J	K
1201 a 3200	C	D	E	G	H	K	L
3201 a 10000	C	D	F	G	J	L	M
10001 a 35000	C	D	F	H	K	M	N
35001 a 150000	D	E	G	J	L	N	P
150001 a 500000	D	E	G	J	M	P	Q
500001 y mas	D	E	H	K	N	Q	R

Imagen 3.5 - Tabla de relación entre tamaño de muestra y el tamaño del lote. (Tomado de "Military Estándar 10")

El código alfabético entregado se lleva a otra tabla donde se observara el número de muestras que se deben tomar para que sea representativa y cumpla con el control de calidad, cabe mencionar que la tabla varía entre los niveles primarios y los niveles secundarios.

Codigo de muestra	Codigo de muestra	Niveles de calidad aceptables (inspeccion normal)																											
		0.01	0.015	0.025	0.04	0.065	0.1	0.15	0.25	0.4	0.65	1	1.5	2.5	4	6.5	10	15	25	40	65	100	150	250	400	650	1000		
		Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc		
A	2	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↓	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31		
B	3	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↓	↓	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31		
C	5	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑	↑	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31		
D	8	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑	↑	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31		
E	13	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑	↑	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31		
F	20	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑	↑	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31		
G	32	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑	↑	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31		
H	50	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑	↑	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31		
J	80	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑	↑	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31		
K	125	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑	↑	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31		
L	200	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑	↑	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31		
N	315	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑	↑	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31		
M	500	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑	↑	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31		
P	800	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑	↑	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31		
Q	1250	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	0 1	↑	↑	1 2	2 3	3 4	5 6	7 8	10 11	14 15	21 22	30 31		
R	2000	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑	↑		

Imagen 3.6 - Tabla de nivel de aceptabilidad de calidad para inspección de nivel primario. (Tomado de "Military Estándar 10")

Codigo de muestra	Codigo de muestra	Niveles de calidad aceptables (inspeccion reducida)																											
		0.01	0.015	0.025	0.04	0.065	0.1	0.15	0.25	0.4	0.65	1	1.5	2.5	4	6.5	10	15	25	40	65	100	150	250	400	650	1000		
		Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	Ac Rc	
A	2	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		
B	2	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		
C	2	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		
D	3	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		
E	5	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		
F	8	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		
G	13	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		
H	20	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		
J	32	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		
K	50	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		
L	60	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		
N	125	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		
M	200	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		
P	315	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		
Q	500	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		
R	800	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓		

Imagen 3.7 - Tabla de nivel de aceptabilidad de calidad para inspección de nivel secundario. (Tomado de "Military Estándar 10")

Con los resultados obtenidos de las tablas se puede determinar fácilmente el tamaño de la muestra que se debe escoger, y es algo que se

debe tener a consideración debido a que no existen dos lotes iguales, y la muestra que debe ser tomada de cada uno de los lotes pese a que es escogida mediante el mismo procedimiento su tamaño puede ser variar; cabe mencionar que el tamaño de la muestra incide en número de fotos tomadas para el muestreo opinático.

3.4.2 Análisis de resultados de la muestra

Definido el plan de muestreo, se procede a tomar la muestra que cumple con las condiciones especificadas y empezar el proceso de control de calidad estadístico, el procesos de control de calidad se divide en una serie de análisis que definirán si el lote pasa el control de calidad, que van desde el ordenamiento de los resultados en una tabla de frecuencia, la obtención de medidas de tendencia central y finalmente las variables de interés las medidas de dispersión.

3.4.2.1 Construcción de tabla de frecuencias y cálculo de medidas de dispersión

Obtenida la muestra de cristales de azúcar se empieza a realizar una serie de cálculos que permitirán construir una tabla de frecuencias, cuya función es ordenar y clasificar toda la información obtenida de nuestra muestra para posteriormente analizarla y obtener la información que determinara la calidad del lote de azúcar y realizar el debido control de calidad.

Para generar la tabla de frecuencias se requiere de los siguientes pasos:

➤ Paso 1: Obtener el rango.

El rango consiste en la diferencia existente entre el número más grande de la muestra y el número más pequeño de la muestra y se obtiene mediante la siguiente ecuación.

$$\text{Rango} = \# \text{ mayor} - \# \text{ menor} \quad (3.1)$$

➤ Paso 2: Determinar el número de clases.

El número de clase se determina mediante la regla de Sturges y se redondea al entero superior, se calcula de la siguiente manera.

$$\# \text{ de Clases} = 1 + \log_2(N) \quad (3.2)$$

Cabe destacar en la ecuación anterior el valor de N representa el número de cristales presentes en la muestra, el número de clases aumenta a razón de que más cristales sean muestreados.

➤ Paso 3: Determinar el ancho de clase.

Se procede a obtener el abanico de valores que poseerá cada una de las clases, por lo cual se define el ancho de cada clase como la división entre el valor del rango entre el número de clases, tal como se muestra en la siguiente ecuación.

$$\text{Ancho de clase} = \frac{\text{Rango}}{\# \text{ de clases}} \quad (3.3)$$

➤ Paso 4: Determinar los límites superiores e inferiores de cada clase.

Los límites superiores e inferiores de cada clase se determinan mediante la suma progresiva del ancho de clase al valor mínimo muestreado hasta obtener como límite superior el valor máximo muestreado, esto se ejemplificará con las siguientes ecuaciones.

$$\text{Limite inferior} = \text{Valmin} + \text{ancho} * (\text{clase} - 1) \quad (3.4)$$

$$\text{Limite superior} = \text{Valmin} + \text{ancho} * (\text{clase}) \quad (3.5)$$

Cabe destacar en las ecuaciones anteriores que "Valmin" representa el valor mínimo de cristal muestreado, "ancho" representa el valor del ancho de clase, y "clase" el dígito de la clase que está representada por ejemplo para la clase uno el número 1.

➤ Paso 5: Determinar el punto medio de la clase.

Debido a que la tabla de frecuencias a generar será una tabla de datos agrupados se necesita un parámetro que permita puntualizar a cada una de las clases, este elemento es el punto medio de clase, el cual es usado para generar algunos cálculos dentro de la tabla y se calcula de la siguiente manera para cada clase.

$$\text{Punto medio} = \frac{\text{Limite superior} + \text{Limite inferior}}{2} \quad (3.6)$$

Donde el límite superior y el límite inferior representan los límites de clase para cada una de las clases, por lo que hay un punto medio por clase.

➤ Paso 6: Determinar elementos que conformaran la tabla de frecuencias

Definido los parámetros de la categoría de datos se procede a determinar las columnas que se necesitaran para obtener la información ocupada por el control de calidad, se destaca por ser una tabla de frecuencias de datos agrupados.

A continuación se mostrara una tabla con todos los elementos indispensables para la obtención de información.

Clases		f_i	X_i	$X_i * f_i$	$(X_i - \bar{X})^2$	$(X_i - \bar{X})^2 * f_i$
Lim inf.1	Lim sup.1					
Lim inf.2	Lim sup.2					
Lim inf.3	Lim sup.3					
Lim inf.4	Lim sup.4					
	Σ		Σ		Σ	

Imagen 3.8 – Tabla de frecuencias

Donde:

f_i es la frecuencia absoluta de cada una de las clases.

X_i es el valor del punto medio de cada una de las clases.

\bar{X} es el valor de la media aritmética

➤ Paso 7: Determinar la media aritmética.

La media aritmética al ser el elemento que representa el valor característico de todos los cristales muestreados y permite observar la tendencia de los cristales, se obtiene mediante la siguiente ecuación para datos agrupados.

$$\text{media aritmetica } (\bar{X}) = \frac{\sum X_i * f_i}{N} \quad (3.7)$$

Donde la sumatoria $X_i * f_i$ es la multiplicación del punto medio por el valor de la frecuencia de la clase respectiva, valor que se calcula en la tabla de frecuencias, y N representa el número total de cristales muestreados.

➤ Paso 8: Determinar la varianza

El siguiente elemento a calcular es la varianza (var), la cual consiste en una medida de dispersión capaz de determinar picos de desviación entre los datos y se calcula mediante la siguiente ecuación.

$$\text{Varianza (var)} = \frac{\sum (X_i - \bar{X})^2 * f_i}{N} \quad (3.8)$$

Donde la sumatoria $(X_i - \bar{X})^2 * f_i$ es la multiplicación del cuadrado de la resta entre el punto medio y la media aritmética por el valor de la frecuencia de la clase respectiva valor que se calcula en la tabla de frecuencias, y N representa el número total de cristales muestreados.

➤ *Paso 9: Determinar la desviación estándar*

Finalmente se obtiene la desviación estándar (σ), medida de dispersión que regula la relación existente entre los datos y la media aritmética, se calcula de mediante la siguiente ecuación.

$$\text{desviación estándar } (\sigma) = \sqrt{\text{var}} = \sqrt{\frac{\sum (X_i - \bar{X})^2 * f_i}{N}} \quad (3.9)$$

➤ *Paso 10: Calculo del diámetro medio*

El diámetro medio ponderado por partículas (AMP) es una medida de referencia utilizada en la industria alimenticia para determinar el tamaño ideal del cristal de azúcar dentro de la muestra y se calcula mediante la suma de todos los diámetros equivalentes dividido entre el número de cristales, como se muestra en la siguiente ecuación.

$$\text{AMP} = \frac{\sum \text{AM}}{N} \quad (3.10)$$

Donde $\sum \text{AM}$ representa a cada uno de los diámetros medio y el valor de N el número de cristales de azúcar presentes en la muestra.

➤ *Paso 11: Calculo del coeficiente de variación de Pearson ponderado por partículas*

Luego de haber realizados todos los cálculos de la tabla de frecuencia, se procede a obtener el coeficiente de variación de Pearson (C.V) el cual entrega uno de los resultados más importantes dentro del análisis de control de calidad estadístico.

El coeficiente de variación de Pearson permite determinar de manera exacta la cercanía o distanciamiento existente entre cada uno de los cristales de azúcar muestreados, y puede ser ponderado por peso o por partículas

Gracias que el método de obtención de información son imágenes los cristales muestreados poseen dos dimensiones, gracias a ello la variable característica en este método es el área, por lo que todos los pasos anteriores deben realizarse con el área de los cristales como variable de interés.

El Coeficiente de variación ponderado por partícula (CVP) basado en el lado más largo permite determinar de manera exacta la relación morfológica existente entre cada uno de los cristales de azúcar muestreados y expresarlo con un valor numérico.

El coeficiente de variación de Pearson ponderado por partículas se calcula mediante la siguiente ecuación.

$$CVP = \frac{\sigma}{\bar{X}} * 100\% \quad (3.11)$$

Donde σ representa a la desviación estándar y \bar{X} a la media aritmética, la ecuación se multiplica por el 100%, ya que el C.V nunca será mayor a 1 pero siempre mayor que 0, por lo que puede representares su resultado de manera decimal o porcentual.

3.4.2.2 Coeficiente de variación ponderado por volumen

Debido a que el resultado del C.V. puede ser exigido de las dos formas el análisis de control de calidad estadístico contendrá ambas ponderaciones por volumen y por partícula.

Pese a que el método de obtención de información son imágenes y se poseen dos dimensiones, se prefiere trabajar con una sola dimensión en este caso el diámetro equivalente de cada cristal, debido a que este se calcula de manera diferente los pasos que sigue no son los mismos, toda la información que utiliza viene directamente del software.

El Coeficiente de variación ponderado por volumen (CVV) es el método tradicional utilizado en la industria alimenticia para mostrar la variación morfológica existente entre los cristales por lo que su inclusión es imprescindible en el proceso de control de calidad.

A diferencia del coeficiente ponderado por partículas el coeficiente ponderado por peso se basa en el modelaje de los cristales como esferas perfectas, por lo que primero hay que calcular el diámetro medio ponderado por volumen (AMV) mediante la siguiente ecuación.

$$AMV = \frac{\sum MA_i * V_i}{\sum V_i} \quad (3.12)$$

Donde $\sum MA_i * V_i$ representa la sumatoria de la multiplicación del volumen del cristal por su respectivo eje mayor y $\sum V_i$ la sumatoria de todos los volúmenes de los cristales.

Obtenido el diámetro equivalente ponderado por volumen se procede a obtener el coeficiente de variación ponderado por volumen, mediante la siguiente ecuación.

$$CVV = \frac{\sqrt{\frac{\sum (MA_i - AMV)^2 * V_i}{\sum V_i}}}{AMV} * 100\% \quad (3.13)$$

Donde $\sum (MA_i - AMV)^2 * V_i$ representa la sumatoria de la multiplicación del volumen del cristal por el cuadrado de la resta entre el diámetro equivalente de los cristales menos el diámetro equivalente ponderado por volumen y $\sum V_i$ la sumatoria de todos los volúmenes de los cristales.

Ambos coeficientes de variación permiten determinar el comportamiento de la muestra obtenida del lote, la diferencia entre uno y el otro reside en el tipo de información que analiza pero cabe resaltar que los valores obtenidos por lo mismo que son ponderados distintos no necesariamente son iguales pero si cercanos.

3.4.3 Visualización de resultados

La visualización de resultados no es más que la exposición de toda la información obtenida a lo largo del proceso de control de calidad y saber expresarla de tal manera que los resultados sean comprensibles para cualquier lector con conocimientos básicos en estadística, por lo que se procede a exponer las razones de selección de los gráficos y medidas a visualizar.

➤ *Histograma*

El histograma de distribución normal permite representar las concentraciones de cristales de azúcar a razón de la geometría esperada, obteniendo la forma de una campana de gauss, este método es bastante útil debido a que no requiere de muchas imágenes para realizar un análisis efectivo y de igual manera denota el comportamiento de la muestra y las concentraciones que puede llegar a tener.

Por lo que si luego del análisis estadístico el histograma denota una forma fuera de la tendencia, representaría una serie de problemas en el procesos de extracción elaboración del azúcar, o en peor de los casos la mala cosecha de la caña de azúcar, sin embargo si llegara a estar dentro de los estándares reflejaría el buen desempeño de la elaboración del azúcar.

➤ *Diagrama de dispersión*

El diagrama de dispersión permitirá reconocer la forma geométrica que posee cada cristal y la variación que posee con el cristal esperado, gracias a ello se puede identificar si los cristales han crecido con la misma proporción durante su producción.

Si la variación de tamaños de los cristales excede un límite podría denotar malformaciones en los cristales, cristales a medio formar, o mayor formación de cristales diminutos debido a la absorción de sacarosa por parte de otros cristales, denotando la necesidad de verificar el proceso de calentamiento y formación de cristales.

➤ *Resultados de las medidas de dispersión.*

Las medidas de dispersión plasmadas en el análisis de resultados permiten identificar las variaciones presentadas en la homogeneidad de los cristales de azúcar a nivel de partículas y de peso, con la finalidad de controlar las variaciones tanto en tamaño, geometría y peso de la muestra de cristales.

Estos valores son importantes ya que representan alarmas que denotan si una muestra se está saliendo de control o necesita atención especial.

De igual manera mediante registros guardados en una base de datos se podría observar si la información está variando y de esa manera encontrar el origen de la falla y proceder rápidamente de la mejor manera posible.

Por otro lado si la información no varía y se mantiene dentro de los parámetros se puede analizar a detalle y encontrar formas de optimizar el sistema, aumentando el número de protecciones que reduzcan el número de fallas y realizar las maniobras de mantenimiento preventivo necesarias.

➤ *Curva de normalización (descartado).*

La curva de normalización en primera instancia permite determinar el número de fotos necesarias para tener un resultado bastante aceptable, ya que a medida se aumenta el número de fotos a analizar se aumenta más la cercanía entre los datos, reduciendo la dispersión hasta valores muy cercanos de la media aritmética (\bar{X}).

Pese a ser una buena forma de representar los resultados requiere de muchas imágenes volviéndolo poco práctico, sin embargo resulta ser un método muy efectivo.

3.4.4 Análisis de resultados.

El control de calidad estadístico entrega una serie de resultados posterior a los análisis realizados, sin embargo esta información no tiene ninguna utilidad si no se sabe interpretar, por lo cual existen diversas formas que orientan a comprender los resultados.

➤ *Estándares internacionales.*

Existen distintos estándares internacionales que permiten determinar si un lote de azúcar es aceptable o no y estos estándares varían dependiendo de la región, el país o incluso la empresa que solicita el azúcar; un ejemplo de estos estándares es la Norma ISO 22000.2005.

➤ *Pruebas constantes.*

Mediante la realización de análisis a distintas muestras se puede determinar la tendencia entre la variación de los cristales y determinar parámetros de regularidad presente en los cristales, cabe mencionar que los cristales de azúcar pueden tener distintos tamaños por lo que la serie de análisis a realizar son extensos si se desea obtener buenos resultados.

Capítulo IV

Desarrollo de Software

4.1 ¿Qué es MATLAB?

4.1.1 Descripción general

MatLab es un acrónimo para *Matrix Laboratory* (Laboratorio de Matrices); es utilizado para referirnos tanto a un lenguaje de programación de alto nivel diseñado para el manejo de matrices, como al software que se utiliza como interfaz de programación y compilador de este lenguaje.

Entre las funciones principales podemos encontrar:

- Manipulación de matrices
- Análisis de datos
- Desarrollo de algoritmos
- Creación de modelos y aplicaciones
- Creación de interfaces de usuario
- Importación y exportación de datos
- Interconexión con programas desarrollados en otros lenguajes como Java, C/C++, Fortran y Python.
- Interacción con hardware de terceros

4.1.2 Historia

- 1954 - 1957: Desarrollo del primer compilador Fortran por un equipo de IBM.

Fortran fue el primer lenguaje de programación de alto nivel y fue diseñado específicamente para el cálculo numérico. Es el principal sucesor de MatLab.

- 1958 - 1961: Se desarrollan las versiones posteriores como el Fortran II, Fortran III y Fortran IV las cuales implican mejoras significativas a este.
- 1962: Se estandariza Fortran para su comercio.
- 1960s-1970s: Se realizan las últimas actualizaciones del Fortran “pre-MatLab”.
- Década de 1970:

A pesar de la gran cantidad de modificaciones que sufrió Fortran, seguía presentando serios inconvenientes como la pérdida de precisión, codificación y algoritmos ineficientes; así como lo extenso y complejo que se tornaba un código para una solución sencilla.

Debido a esto el gobierno de los Estados Unidos impulsó un proyecto para la creación de algoritmos de buena calidad para su aplicación en tareas comunes, que cualquiera pudiera ocupar en sus desarrollos. Como resultado de este surgieron librerías libres para cálculo numérico como BLAS, LINPACK y EISPACK.

A finales de esta década, uno de los autores de LINPACK y EISPACK, Cleve Moler, un programador especializado en el análisis numérico, empezó a desarrollar MatLab como una herramienta para los alumnos de cursos de álgebra lineal y análisis numéricos de la Universidad de Nuevo México, para facilitar el uso de las librerías antes mencionadas sin tener que conocer Fortran.

MatLab tuvo tal impacto y fama en la comunidad educativa que se esparció rápidamente hacia las demás universidades, y tuvo una gran aceptación entre los conocedores y aprendices de las matemáticas.

- 1983: El cofundador y actual presidente de MathWorks, Jack Little, coincide con Cleve Moler durante una visita a la Universidad de Stanford, donde conoció por primera vez su trabajo; reconociendo su potencial decide formar parte de este proyecto y se une al equipo de desarrollo.
- 1984: Jack Little y Steve Bangert reprograman MatLab en lenguaje C, con nuevas librerías denominadas JACKPAC y añaden los archivos M, *toolboxes* y otras herramientas para graficar.

Este año es uno de los más importantes debido a que se funda MathWorks y se lanza, el que ahora es su principal producto, MatLab.

- 2000: Se reescribe MatLab para que sea capaz de utilizar nuevas librerías desarrolladas en Fortran 90 llamada LAPACK - *Linear Algebra Package* (Paquete de Álgebra Lineal).

4.2 Herramientas de visión artificial de MATLAB

4.2.1 Generalidades sobre las *toolbox*

MatLab, además de poseer las estructuras comunes de cualquier lenguaje de programación, como son las de decisión y las de repetición, posee grupos de herramientas llamados *toolbox*.

Dentro de los *toolbox* se agrupan funciones, algoritmos, aplicaciones y comandos para tareas específicas que se pueden separar en 11 categorías:

1. Matemáticas, Estadística y Optimización
2. Sistemas de control
3. Procesamiento de señales y comunicaciones
4. Procesamiento de imágenes y Visión Artificial
5. Prueba y medición
6. Finanzas computacionales
7. Biología computacional

8. Generación de Código y Verificación
9. Implementación de aplicaciones
10. Acceso Base de Datos e Informes
11. Cálculo Paralelo

4.2.2 Toolbox: Procesamiento de Imágenes y Visión Artificial

Conocida como *Image Processing Toolbox*, es el conjunto de herramientas especializadas para el procesamiento, análisis y visualización de imágenes; está compuesta por algoritmos, funciones y aplicaciones que facilitan las tareas antes mencionadas.

Dentro de las principales funciones encontramos:

1. Análisis de imágenes
 - a. Segmentación
 - b. Morfología
 - c. Estadística
 - d. Medición
2. Mejora de imágenes
 - a. Filtrado
 - b. Enfoque de imágenes borrosas
3. Transformaciones geométricas y métodos de registro de imágenes
4. Transformaciones de imágenes
 - a. FFT - *Fast Fourier Transform* (Transformada Rápida de Fourier)
 - b. DCT - *Discrete Cosine Transform* (Transformada de Coseno Discreta)
 - c. Radón
 - d. Proyección de haz de rayos de abanico
5. Flujos de trabajo para imágenes grandes
 - a. Procesamiento por bloques
 - b. Creación de mosaicos
 - c. Visualización en varias resoluciones

6. Aplicaciones de visualización

- a. *Image Viewer*
- b. *Video Viewer*

Dentro de estas series de herramientas hay mucho que destacar y definir; dentro de los parámetros más importantes están los que nos darán la flexibilidad y alcance en nuestros desarrollos, dentro de los cuales hemos identificado:

- Formatos de escritura - lectura: Los formatos son especificados en MatLab como variables o parámetros del tipo *string*, y están regidos por una lista de compatibilidad, guardados en un registro al cual se puede acceder por medio del comando `imformats`. Dentro de los formatos más conocidos podemos encontrar:
 - BMP - *Windows Bitmap*
 - JPEG - *Joint Photographic Experts Group*
 - GIF - *Graphics Interchange Format*
 - PNG - *Portable Network Graphics*
 - TIFF - *Tagged Image File Format*
 - Entre otros formatos de imágenes...
- Tipo de color: En MatLab el tipo de color no solo determinan los colores que se representarán en la visualización; sino que también influye en la dimensión de la matriz que representa a la imagen y en profundidad de bits de cada una.

Al cargar una imagen en MatLab mediante el comando `imread`, se realiza la conversión de esta a una matriz, cuyas dimensiones están determinadas por características de la imagen como su largo, ancho y cantidad de capas.

Los datos que componen esta matriz dependen a su vez de la profundidad de bits que posee cada imagen. En MatLab, dependiendo del formato original, se realiza la conversión a cualquiera de las 3 profundidades de bits en las cuales importa MatLab; estas pueden ser imágenes de 1 bit (`logical`), de 8-bits (`unit8`) o de 16-bits (`unit16`).

Cabe mencionar que cada formato de imagen maneja rangos de profundidades de bits diferentes; por ejemplo, el formato CUR - *Cursor File* es capaz de soportar 1,4 y 8 bits; mientras que el formato JPEG - *Joint Photographic Experts Group* soporta 8,12 y 16 bits. Debido a esto, MatLab clasifica por rangos las profundidades para "decidir" a cuál de las que soporta MatLab deben ser convertidas, de la siguiente manera:

Tabla 4.1

Conversión de la profundidad de bits en MatLab

Profundidad de Bits del archivo original	Clase de matriz generada en MatLab
1 bit	Logical (0-1)
De 2 a 8 bits	unit8 (0-255)
De 9 a 16 bits	unit16 (0-65535)

En la primera columna se muestran los rangos de profundidades de bits en los que se encuentran la mayoría de formatos de imágenes, en la segunda columna se muestran la profundidad a la cual MatLab las convierte.

Hay que aclarar que esta es una regla que se aplica a la mayoría de casos; algunos formatos como BMP - *Windows Bitmap*, PNG - *Portable Network Graphics*, entre otros, son casos particulares y se recomienda revisar la información que brinda MatLab al respecto.

Estos son las conversiones que realiza MatLab cuando se importan imágenes, sin embargo existe otra profundidad de bits que se utiliza y es la denominada "de precisión doble", con 64 bits y es representado por

números de punto flotante entre 0 y 1. Esta es la forma nativa en que MatLab representa los datos numéricos, aunque por optimización del uso de memoria se acostumbra a trabajar con las otras profundidades.

- Tipo de imágenes: A continuación se explican los tipos de imágenes con las que se trabajan en MatLab.

- *Imágenes Indexadas*

Una imagen indexada consiste en dos matrices; la primera es una matriz bidimensional de apuntadores cuyas dimensiones son el largo y ancho de la imagen. Los apuntadores son números enteros, los cuales hacen referencia a una posición en la segunda matriz llamada *colormap* de n-por-3; esta contiene valores de punto flotante para especificar la proporción de los colores rojo, verde y azul para definir el color del pixel al cual el apuntador representa.

- *RGB o Color Real*

Las imágenes RGB son procesadas como matrices tridimensionales, cuyas primeras dos dimensiones están definidas por la altura y ancho de la imagen, y la tercera es el número tres, representando el número de capas las cuales corresponden a los colores rojo, verde y azul. Cada pixel posee el color determinado por la combinación de la proporción indicada de los tres colores; además, normalmente, cada pixel posee una profundidad de 8 bits en cada capa, aunque se encontraran casos con las demás profundidades.

- *Escala de Grises o Imagen de Intensidad*

Es llamada Escala de Grises debido a que los colores de visualización son diferentes intensidades de gris; el tono de gris representado depende del valor de cada pixel que contiene una matriz bidimensional, cuyas únicas dos dimensiones son las mismas dimensiones de la imagen. Esta matriz puede ser de las clases `double`, `uint8` o `uint16` donde, entre más profundidad de bits posea, mayor cantidad de tonos de grises se harán presente en la imagen. Entre más se acerque el valor de cada pixel a cero, el color será más oscuro; mientras que más se acerque a uno, será más claro.

También se le conoce como Imagen de Intensidad, porque de esta forma se representa la intensidad de un color en una imagen; si nosotros extraemos una de las capas de una imagen RGB, por ejemplo la capa Roja, y la mostramos a través del comando `imshow`, observaremos una imagen en escala de grises determinando la intensidad de rojo de esa imagen; donde, un color oscuro representa la ausencia de este color, y un color claro su presencia.

- *Imágenes Binarias*

Las imágenes binarias en realidad es el resultado de un proceso de conversión de una imagen en escala de grises, ya sea de 8, 16 o 64 bits a una representación de 2 bits. En la cual se determina un valor de umbral entre 0 y 1 a partir del cual los valores menores a este se vuelven 0, es decir, todos los tonos de grises que representan un porcentaje menor al indicado se volverán negros; de igual manera los valores mayores al umbral se volverán blancos, obteniendo así una imagen constituida sólo por blanco y negro. Estas imágenes son sumamente útiles para la extracción de características, entre otras cosas, debido a que MatLab posee varias herramientas que solo se pueden aplicar a imágenes binarias.

4.3 Definición del Software

4.3.1 ¿Qué es SugarVision?

SugarVision es un software desarrollado para facilitar los procesos de control de calidad de las azucareras. Hace uso de herramientas de visión artificial para extraer características geométricas de los cristales de azúcar a partir de fotos; los datos son analizados estadísticamente para obtener información útil y poder realizar conclusiones acerca de la muestra estudiada y los factores a los que estuvo sometido durante su producción. SugarVision se ha desarrollado con fines didácticos de investigación y desarrollo en el área de la visión por computador enfocado en la industria alimenticia.

4.3.2 Consideraciones de desarrollo del software

Dentro de esta sección explicaremos la secuencia en la que se fue desarrollando el software, empezando por los requerimientos iniciales, luego los pasos que se tomaron para el desarrollo y por último las funciones con las que se terminó el software.

4.3.2.1 Requerimientos del software

Se solicitó un software que, por medio de visión artificial, fuera capaz de extraer las medidas de los cristales de azúcar, a partir de una serie de fotos captadas desde un estereoscopio. Debía poseer suficiente flexibilidad en cuanto a las variables que afectan el proceso, como el umbral, calibración, etc... Y a la vez tener un flujo de trabajo bien definido. Además, estos datos debían ser analizados por métodos estadísticos para poder obtener, principalmente, valores como la media y el coeficiente de variabilidad, los cuales son valores estándares en este tipo de análisis, y mostrar comparaciones a través de gráficos.

4.3.2.2 Pasos

➤ *Definir el software de desarrollo*

Al iniciar el proyecto se dejó la libertad de desarrollar el software en cualquier programa que tuviera la facilidad del manejo, análisis y visualización de datos; así como, que tuviera la capacidad de soportar herramientas de visión por computador y procesamiento digital de las imágenes. Por lo cual, considerando la accesibilidad al programa y el conocimiento previo que se tenía de este, se eligió MatLab; ya que nos brinda las herramientas necesarias con las *toolboxes* antes mencionadas y la capacidad de desarrollar aplicaciones ejecutables a partir del desarrollo de interfaces.

➤ *Identificar los pasos necesarios para realizar de manera completa y óptima el análisis*

Dentro del proyecto podemos identificar dos pasos importantes, el procesamiento de las imágenes y el análisis estadístico. Pero son pasos demasiado generales, en especial el relacionado con visión artificial; por lo cual, nos encontramos en la necesidad de detallar los pasos necesarios para lograr cada uno de ellos:

1. Procesamiento digital de imágenes
 - 1.1. Importación de imagen
 - 1.2. Conversión a formato binario
 - 1.3. Ajuste de valor de umbral
 - 1.4. Especificación de factor de calibración
 - 1.5. Limpieza de ruido de la imagen
 - 1.6. Análisis de coincidencia
2. Estadística
 - 2.1. Análisis
 - 2.2. Muestra de resultados

➤ *Identificar cantidad y tipo de variables a utilizar*

Las variables utilizadas las podemos clasificar según dos diferentes criterios:

- Tipo
 - Numéricas - Principalmente para cálculos y contadores.
 - Cadenas de texto - Nombres de archivos.
 - Matrices - Las imágenes son representadas de esta manera, también se utilizarán para guardar datos de cristales.
 - Estructuras - De esta forma se muestran características obtenidas para cada cristal.
- Función que desempeñan
 - Variables de entrada - Son los datos que el usuario ingresará por medio de los cuadros de texto y otros métodos de introducción de datos.
 - Variables de salida - Todos los resultados obtenidos a partir de los cálculos y que son de interés para el usuario, entre estos se encuentran las imágenes resultado de cada proceso.
 - Variables de cálculo - Son todos aquellos datos que sirven como entrada o salida de cada cálculo y se mantienen de forma interna ya que no son de interés para el usuario o cambian repetitivamente y de manera instantánea.

➤ *Definir la distribución de la interfaz*

Se espera que la interfaz sea intuitiva y de fácil acceso a cualquier función; que no se sature de contenido y sea enfática, en primera instancia, la visualización de cada paso del procesamiento de las imágenes.

4.4 Interfaz y funcionamiento de SugarVision

4.4.1 Interfaz

➤ *Pantalla para procesamiento de imágenes*

Esta pantalla está condicionada para procesar las imágenes y especificar las condiciones bajo las cuales se extraerán las características.



Imagen 4.1 - Pantalla para procesamiento de imagen. Se muestran las partes que la conforman.

En esta pantalla podemos encontrar las siguientes partes:

1. Barra de herramientas. (Se detallará más adelante)
2. Secuencia de pasos, indica en qué paso del proceso se encuentra.
3. Información del paso actual
4. Ventana de visualización
5. Ingreso de parámetro para el proceso actual
6. Botón de aplicación
7. Navegador de funciones

➤ *Pantalla para visualización de estadística*

Esta pantalla está dedicada a mostrar las gráficas y datos resultados de los análisis estadísticos a los cuales fueron sometidos los datos adquiridos en el paso anterior.

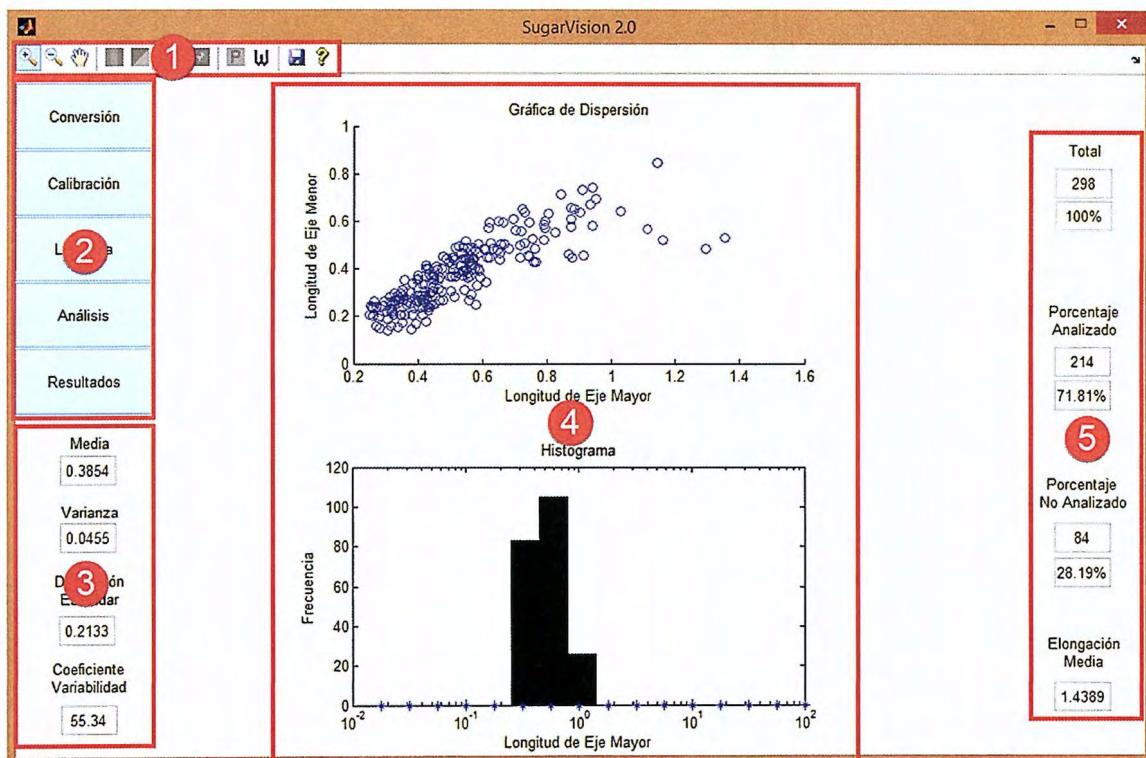


Imagen 4.2 - Pantalla para visualización de estadística. Se muestran las partes que la conforman.

En esta pantalla podemos encontrar las siguientes partes:

1. Barra de herramientas
2. Secuencia de pasos
3. Resultados estadísticos
4. Ventana de visualización
5. Porcentajes de cristales analizados y no analizados

4.4.2 Funciones y pasos

4.4.2.1 Importar archivos de imagen

El primer paso de nuestro Software es la importación de imágenes, la cual se realiza por medio de una ventana de exploración, esta tiene capacidad para mostrar archivos en formatos JPEG y PNG, además de ser multiselectiva debido a que el programa está diseñado para funcionar con una o más imágenes de la misma muestra.

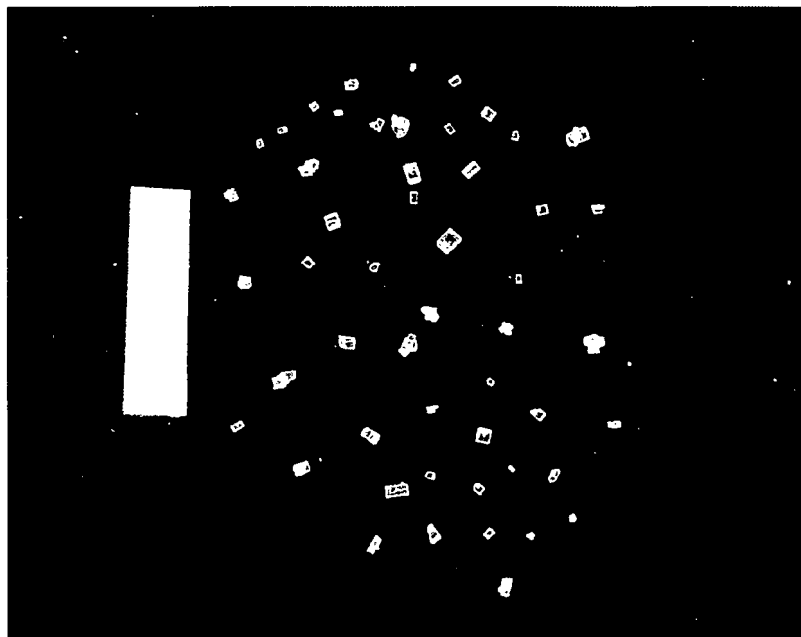


Imagen 4.3 - Muestra de 50 cristales como ejemplo del primer paso.

4.4.2.2 Conversión de imagen a matriz

El paso inmediato a la selección de las imágenes, es su conversión a matrices por parte de MatLab, tomando en cuenta criterios como el formato y la profundidad de bits.

Además, en este paso se realizan cambios del tipo de color: de RGB a escala de grises y luego a binario; debido a que cada tipo nos brinda información necesaria para el proceso, como por ejemplo, el “auto-umbral”, definido de manera automática por MatLab a través del comando `graythresh`, por medio del estudio del histograma representativo de la imagen en escala de grises.

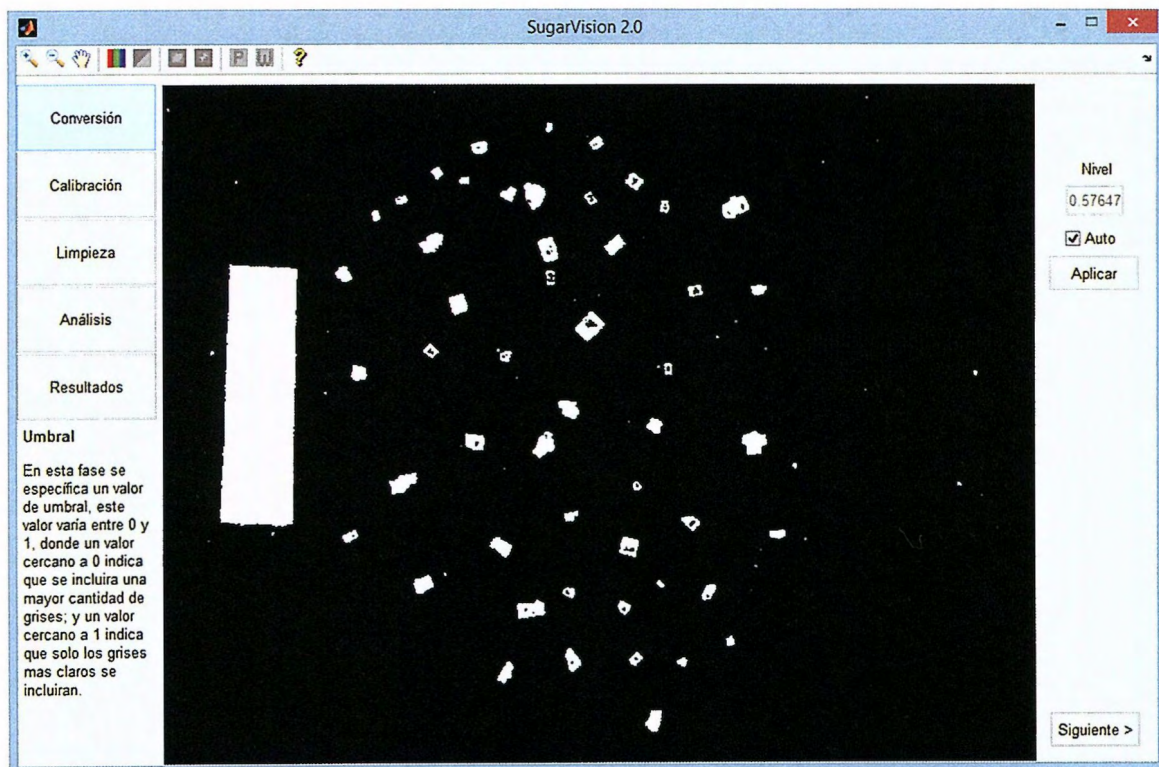


Imagen 4.4 - Primera visualización del programa con umbral automático

4.4.2.3 Cambio de Umbral

Del paso anterior obtenemos nuestra primera visualización, la cual ya está convertida en una imagen binaria con un umbral definido por MatLab; en este paso tenemos la capacidad de cambiar el valor del umbral.

Este paso es opcional y se puede editar el umbral definido por MatLab desmarcando la casilla "Auto"; luego se modifica ingresando un valor entre 0 y 1, donde un valor más cercano a 0 provocará que se tomen en cuenta mayor cantidad de tonos de gris, por lo cual aparecerán más regiones blancas en la imagen. Por el contrario, un valor más cercano a 1 reducirá la cantidad de regiones.

Este cambio es bastante subjetivo, debido a que depende de la iluminación de la imagen y se debe especificar intentando mantener un equilibrio entre los pro y los contras de este cambio; ya que aumentar el umbral reducirá la cantidad de áreas no deseadas dentro del análisis, pero también reducirá las áreas de los cristales modificando sus dimensiones; por el contrario, un valor muy bajo, podría definir de mejor manera los bordes de los cristales adquiriendo así medidas más precisas, pero también nos aumentará la cantidad de ruido y áreas no deseadas dentro del análisis.

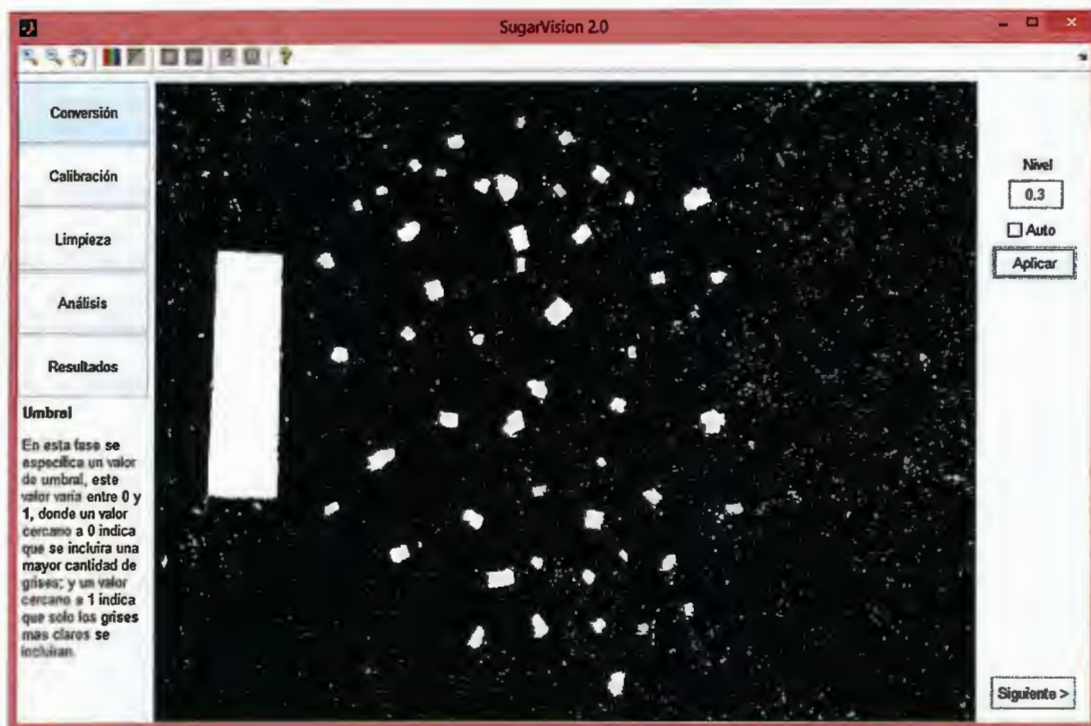


Imagen 4.5 - Ejemplo de umbral muy bajo aplicado a la misma imagen

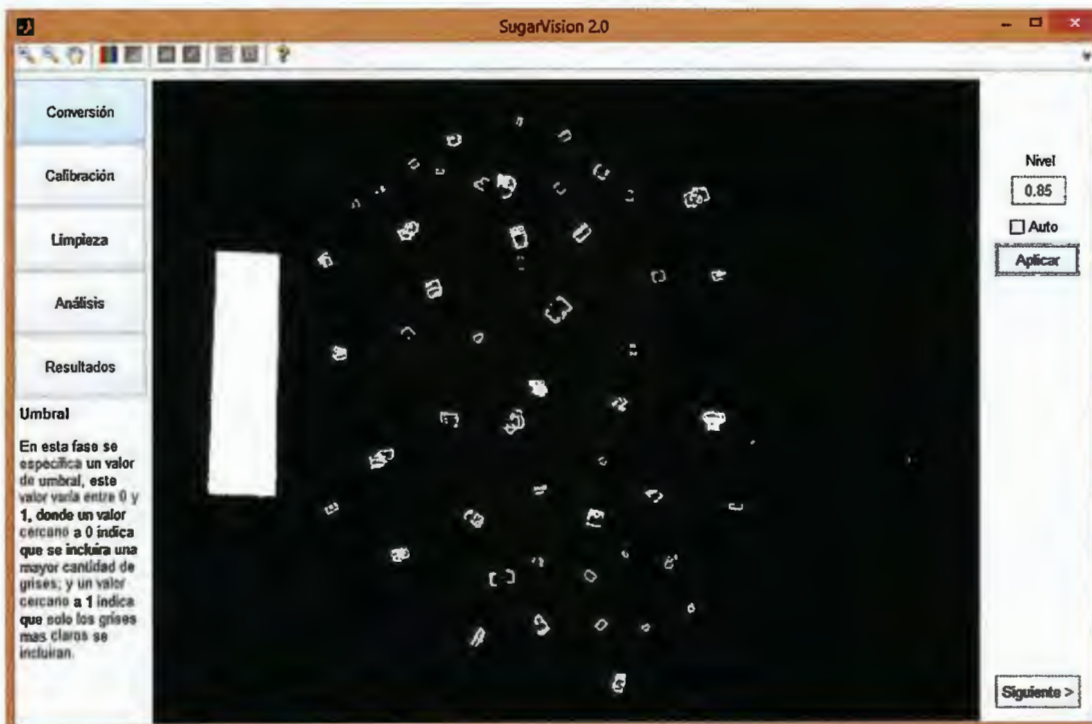


Imagen 4.6 - Ejemplo de umbral muy alto aplicado a la misma imagen

4.4.2.4 Transiciones en Navegador de Funciones

Al presionar “Siguiente” en el Navegador de funciones, la transición entre paso y paso trae consigo la adquisición y establecimiento de una serie de parámetros necesarios para el paso siguiente. Estos parámetros pueden ser visibles como cambios en la interfaz para la recepción y despliegue de información; o la adquisición y procesamiento de datos provenientes de las características geométricas de las regiones detectadas en la imagen.

La gran mayoría de estos datos se obtiene por medio del comando `regionprops`, el cual regresa una estructura, donde se etiqueta cada una de las regiones identificadas y se muestran todas las características necesarias. A continuación se especifican las características más utilizadas en SugarVision con la terminología utilizada en MatLab:

- MajorAxisLength - *Longitud de eje mayor*: como su nombre lo indica, es la medida en pixeles del lado más largo de una región.
- MinorAxisLength - *Longitud de eje menor*: es la medida en pixeles del lado más corto de una región; es perpendicular al eje mayor.
- FilledArea - *Área llena*: Indica la cantidad de pixeles que contiene una región cerrada sin importar si esta posee agujeros internos.
- Orientation - *Orientación*: es el ángulo que posee el eje mayor con respecto a la horizontal.
- BoundingBox - *Caja Delimitadora*: Determina la posición y tamaño de la caja mínima que encierra a la región.
- EquivDiameter - *Diámetro equivalente*: Calcula el diámetro que poseería un círculo con la misma área

4.4.2.5 Definición de línea de referencia para calibración

Este es un paso obligatorio dentro del proceso, debido a que gracias a este, toda dimensión adquirida es trabajada en unidades métricas y no en pixeles; por lo cual, el navegador de funciones para pasar al siguiente paso se activará sólo si este ya fue realizado.

Solo se necesita especificar la medida en milímetros de la línea de referencia, que debe ser parte de la imagen; luego presionar el botón de aplicación "Referencia" y SugarVision determinará cuál es la línea de referencia debido a sus dimensiones y la identificará con un borde verde. Hay que recalcar que esta línea siempre debe ser el área más grande dentro de la imagen.

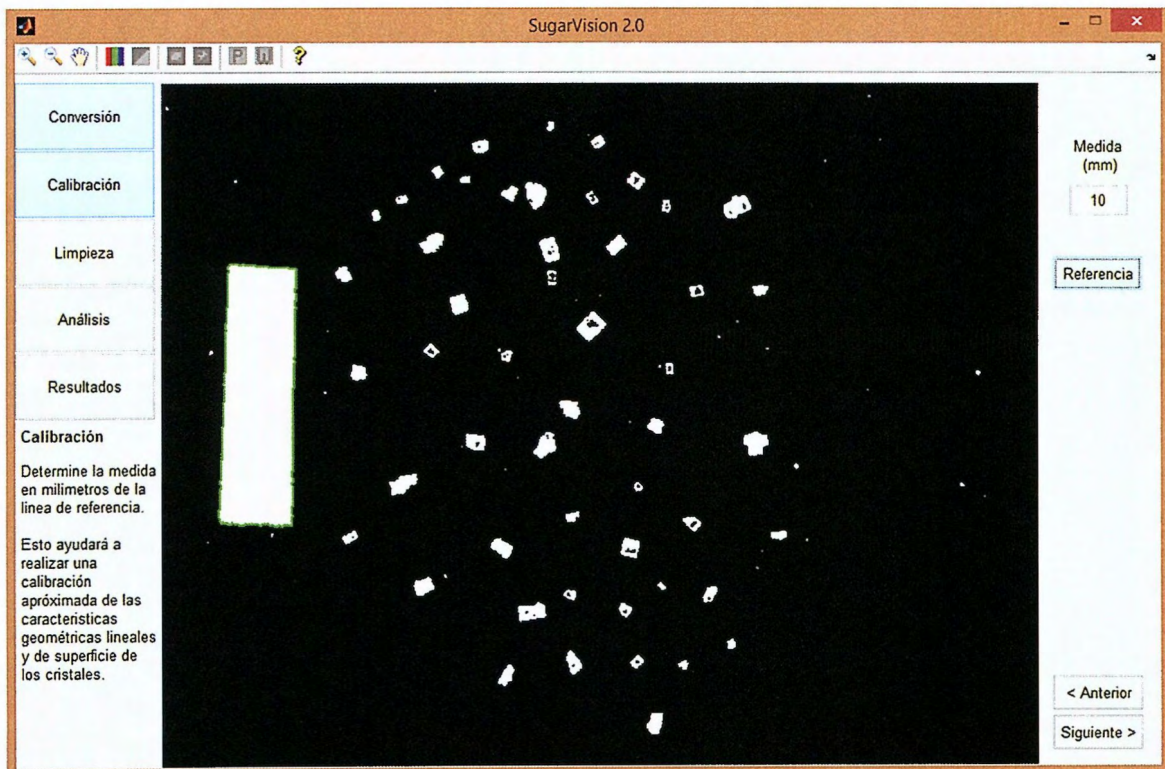


Imagen 4.7 - Detección de línea de referencia

La precisión de las medidas de los cristales dependerá de aspectos como la resolución de la cámara, distancia a la que es tomada la foto, resolución de la foto, tamaño de los cristales y tamaño de la línea de referencia.

4.4.2.6 Limpieza de imagen

En este paso somos capaces de establecer un filtro de longitud, definiendo una medida de longitud de eje mayor, a partir de la cual todas las menores a esta serán eliminadas. Esto sirve para deshacernos de regiones constituidas generalmente por ruido de la imagen, pequeños reflejos, imperfecciones en el fondo y polvo de los mismos cristales.



Imagen 4.8 - Resultado de la limpieza

Como podemos observar, al aplicar un filtro para longitudes menores de 250 μm se eliminan todas las regiones más pequeñas y nos deja solo las correspondientes a cristales; una de las ventajas que el filtro sea por longitud, es que es independiente de la distancia desde la que se tomó la foto.

Aunque este paso es opcional, es un paso bastante útil, debido a que nos permite depurar la imagen para que el análisis sea aplicado solo a cristales de azúcar y resulten datos más precisos.

4.4.2.7 Análisis de regiones

Este paso es una transición entre el procesamiento de imágenes y el análisis estadístico; es el filtro que determinará cuales son los cristales que poseen las características de forma, necesarias para ser sometidos al análisis estadístico.

Para esto, cada cristal es comparado con la caja mínima que encierra su región cuando el eje mayor del cristal se encuentra completamente horizontal.

Como podemos observar en la imagen, el software analiza cada región con su eje mayor completamente horizontal, independientemente de la orientación que este posea en la imagen; luego cada región se rodea con su `BoundingBox` para realizar la comparación de áreas y determinar qué proporción del `BoundingBox` ocupa cada cristal.

Con esta representación, podemos identificar fácilmente, de manera visual, cuales cristales son más rectangulares; para el software, la forma de determinarlo es comparando las áreas del `BoundingBox` con el `FilledArea` de cada cristal, determinando un coeficiente de coincidencia.

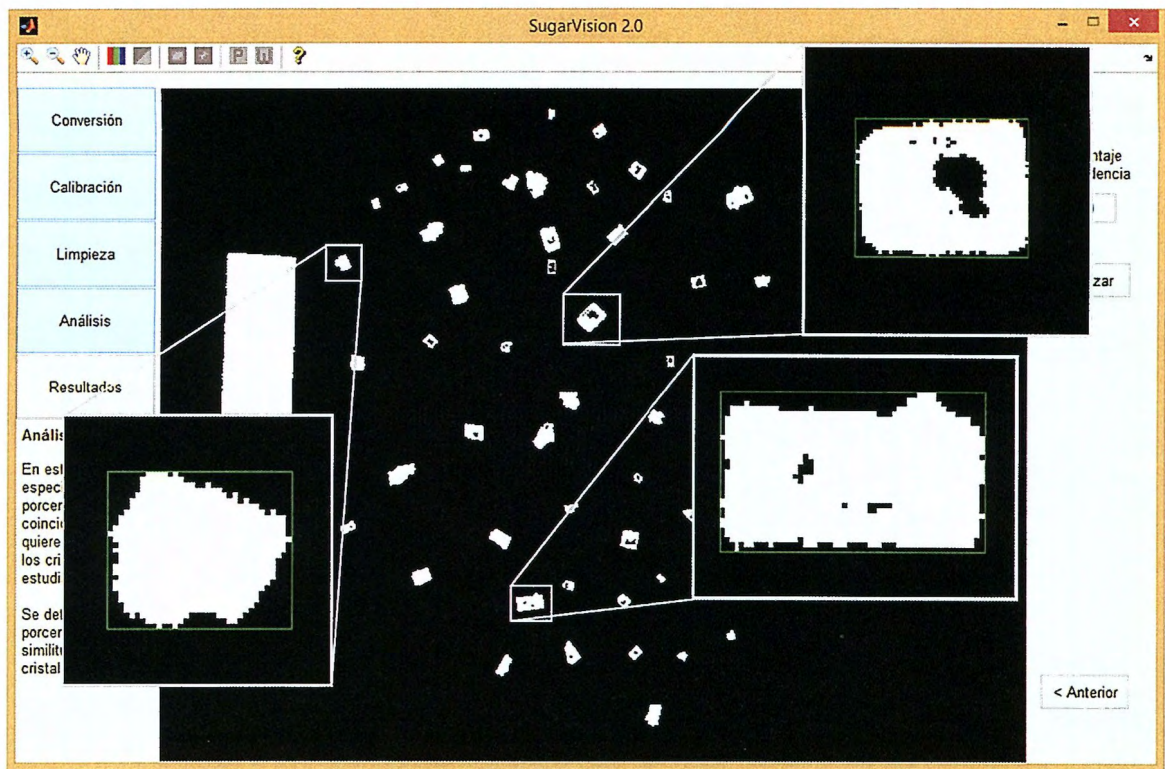


Imagen 4.9 - Análisis de regiones. Representación de cómo el software realiza el análisis.

La fórmula del coeficiente de coincidencia está determinada por:

$$CC = FA/BB \times 100\% \quad (5.1)$$

Donde:

- CC es el Coeficiente de Coincidencia
- FA es el Área llena de la región - *FilledArea*
- BB es el Área de la caja que lo encierra - Área del *BoundingBox*

El coeficiente de coincidencia dará como resultado un porcentaje para cada cristal, el cual servirá como nuevo filtro para determinar cuáles de los cristales serán sometidos al análisis estadístico. El porcentaje mínimo a ser tomado en cuenta se determinará en el texto para ingreso de parámetro de esta pantalla.

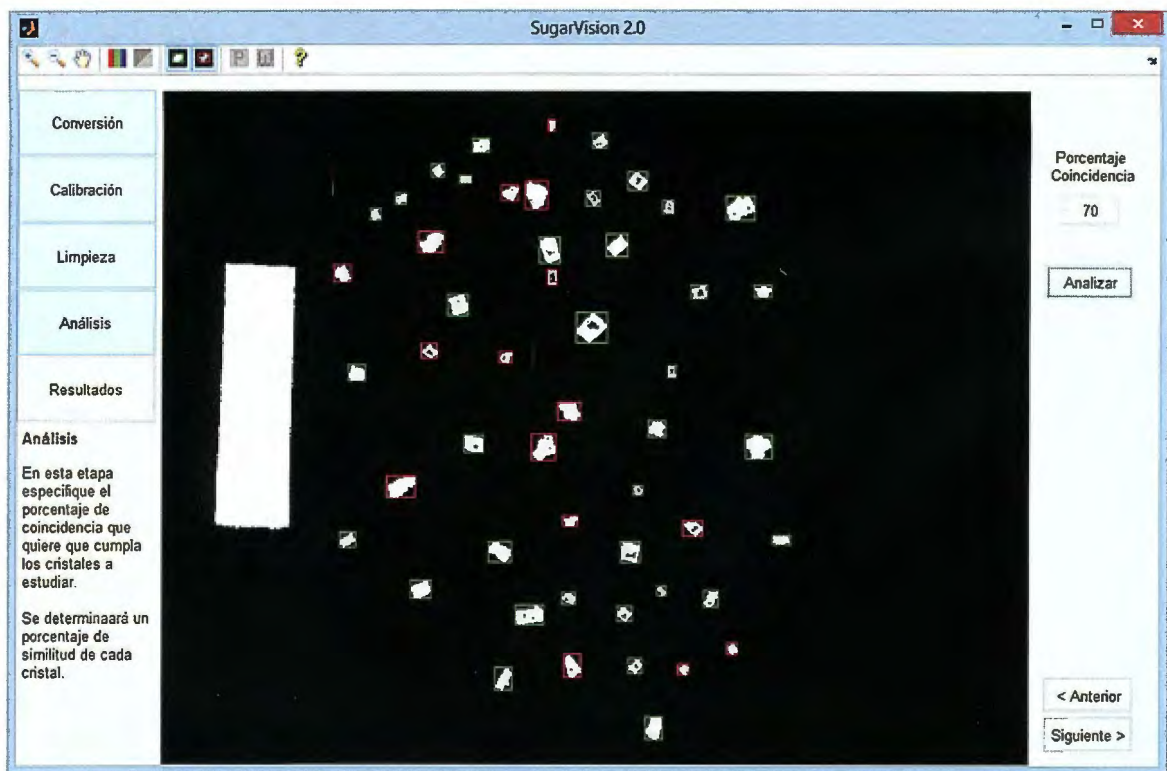


Imagen 4.10 - Resultado de análisis. Para un coeficiente de coincidencia del 70%

Al realizar el análisis, este dará como resultado la señalización de las regiones que cumplen el coeficiente definido, encerrándolas con un rectángulo verde, y las que no, en un rectángulo rojo. Además se activará en la barra de herramientas los botones de activación/desactivación de indicadores de cristales incluidos y excluidos del análisis estadístico.

La determinación del coeficiente de coincidencia depende de la rigurosidad del análisis que sea desea realizar, porque, como es de esperarse, en cuanto mayor sea nuestro criterio, menos cristales serán sometidos al análisis y viceversa. Es importante definir un coeficiente intermedio que satisfaga nuestras necesidades de estudio, debido a que extremadamente pocos cristales podrían cumplir con criterios superiores al 90%.

De manera empírica se ha determinado que los cristales mejor formados cumplen criterios a partir del 70%, pero el software se ha desarrollado para que éste criterio pueda ser definido según las necesidades del usuario; ya que podría ser necesario un estudio con coeficiente alto, donde solo se requiere estudiar cristales estrictamente rectangulares, o estudios con coeficientes muy bajos donde se quiera estudiar toda la muestra, incluyendo los cristales llamados conglomerados.

Este es el último paso modificable, a partir de aquí no se podrán realizar cambios a la imagen, por lo cual las características extraídas a las regiones que pasaron el filtro de coincidencia serán definitivas, así que es recomendable asegurarse de que cumple nuestros requisitos, si no todavía tenemos la capacidad de volver aplicar cualquiera de los pasos anteriores.

Debido a que SugarVision permite seleccionar varias imágenes de la misma muestra, al presionar siguiente se verificará si ya se han analizado todas las imágenes para empezar el análisis estadístico, sino se procederá a intervenir la siguiente imagen para su procesamiento y análisis, y así sucesivamente hasta que se haya realizado el mismo proceso a todas las imágenes seleccionadas.

4.4.2.8 Visualización de Estadística

En esta etapa se muestran los resultados del análisis estadístico al que fueron sometidos los datos extraídos de los cristales que pasaron el filtro anterior. Podemos encontrar valores estadísticos como la media y coeficiente de variabilidad, los cuales son valores estándar para este tipo de análisis; además encontramos varianza, desviación media, elongación media, la cual es la media aritmética de la relación del eje mayor entre el eje menor de cada cristal, y las cantidades del total de cristales, los analizados y los no analizados. Todos los resultados son calculados por el

método de estratificación ponderado por partícula, puede revisar estos conceptos en el Capítulo 3.

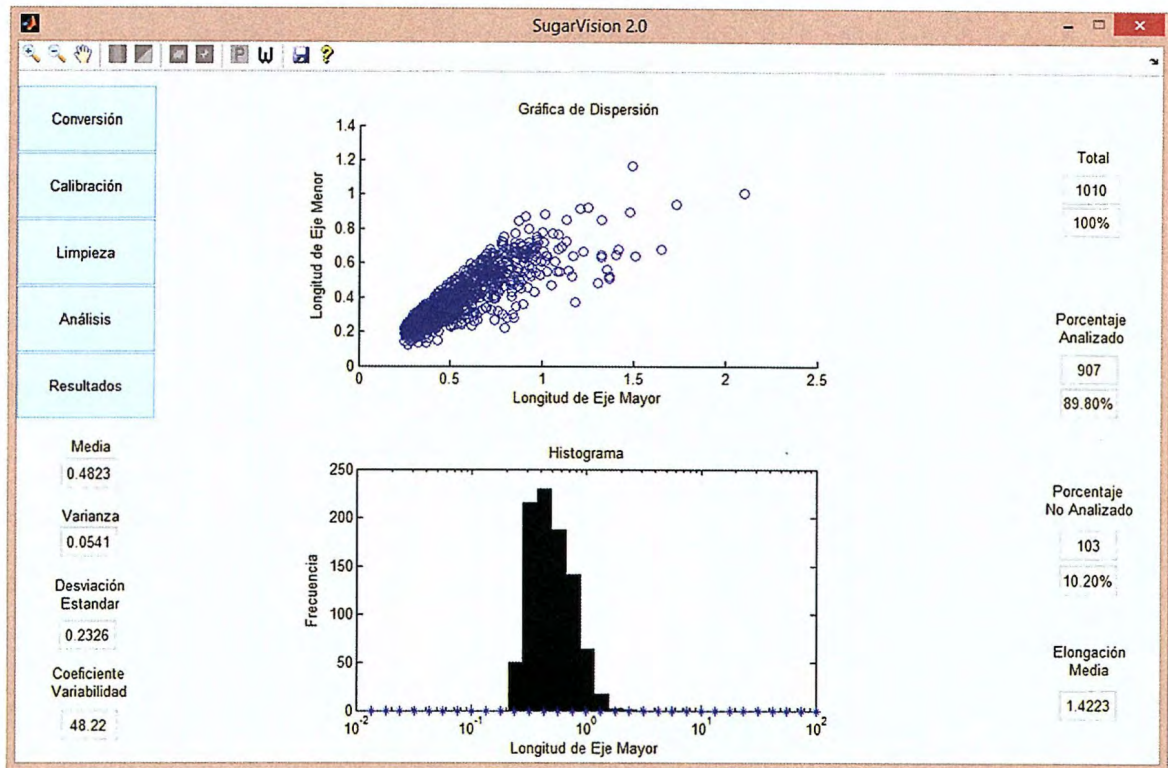


Imagen 4.11 - Visualización de estadística.

Además posee un área de gráficos donde podemos encontrar, en la parte superior, una gráfica de dispersión en el cual se muestra la relación entre el eje mayor y menor de cada cristal; gracias a esto podemos identificar si los cristales han crecido con la misma proporción durante su producción.

En la parte inferior encontramos un histograma semi logarítmico en el cual se evalúa la cantidad de cristales que se encuentra en cada rango de longitudes de eje mayor; en este se espera una tendencia normal, debido a que idealmente el proceso de producción es ajustado para que los cristales posean ciertas características geométricas; pero sabemos que en la realidad siempre existen deficiencias y errores durante el proceso,

por lo cual no todos los cristales poseen estas características, pero se espera que la mayoría sí, o que se encuentren cercanos a estas, y en cuanto más nos alejemos de esta medida será menor la cantidad de cristales; de otra forma, esto podría representar una alteración de las condiciones a las que es sometido el proceso.

4.4.3 Barra de herramientas

La barra de herramientas está compuesta por funciones de visualización y exportación de los datos, tanto gráficos como numéricos. Posee tanto botones y funciones estándar, como botones y funciones propios del software, los cuales pueden ser agrupados en los siguientes conjuntos:

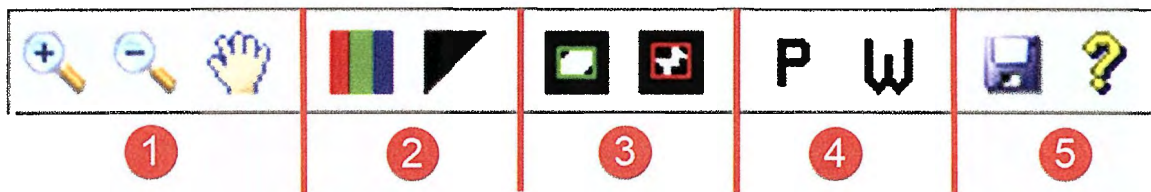


Imagen 4.12 - Barra de herramientas. Se resaltan las divisiones según su función.

1. Herramientas estándar de visualización - Posee herramientas estándar que podemos encontrar en cualquier software de edición de imágenes.
 - a. Acercar
 - b. Alejar
 - c. Movimiento de paneo
2. Herramientas de cambio de tipo de color - Posee herramientas propias del software dedicadas al cambio en la visualización del tipo de color, se utiliza principalmente para la comparación entre la imagen real y la imagen a ser procesada.

- a. Tipo de color RGB: Además de hacer visible la imagen original, en los pasos de Calibración y Limpieza incluye un borde en los cristales que están siendo tomados en cuenta en el proceso; por ejemplo, en el paso de Calibración se mostrarán cristales que cumplen el Umbral definido, mientras que después de ser aplicada la Limpieza, sólo se mostrarán las regiones que cumplen con el criterio de longitud. Además se puede observar en qué medida están siendo tomado en cuenta los cristales porque, debido a la iluminación, algunos cristales no son tomados en su totalidad.



Imagen 4.13 - Visualización en Tipo de Color RGB. Se observa como las regiones blancas pequeñas no son resaltadas debido a que no cumplen con el requisito de longitud, además se observa que algunos cristales no son tomados en su totalidad.

- b. Imagen Binaria: Nos permite regresar al tipo de visualización predeterminada de SugarVision.

3. Indicador de resultado de análisis - Este conjunto solo se habilita después de haber realizado un análisis, y posee herramientas para la activación y desactivación de los indicadores de resultado, los cuales son visibles tanto en RGB como en Binario. Ambos están activados de manera predeterminada.
- a. Activación/desactivación de cristales incluidos: Permite mostrar u ocultar los indicadores verdes que identifican los cristales que serán incluidos en el análisis estadístico.
 - b. Activación/desactivación de cristales excluidos: Permite mostrar u ocultar los indicadores rojos que identifican los cristales que serán excluidos en el análisis estadístico.

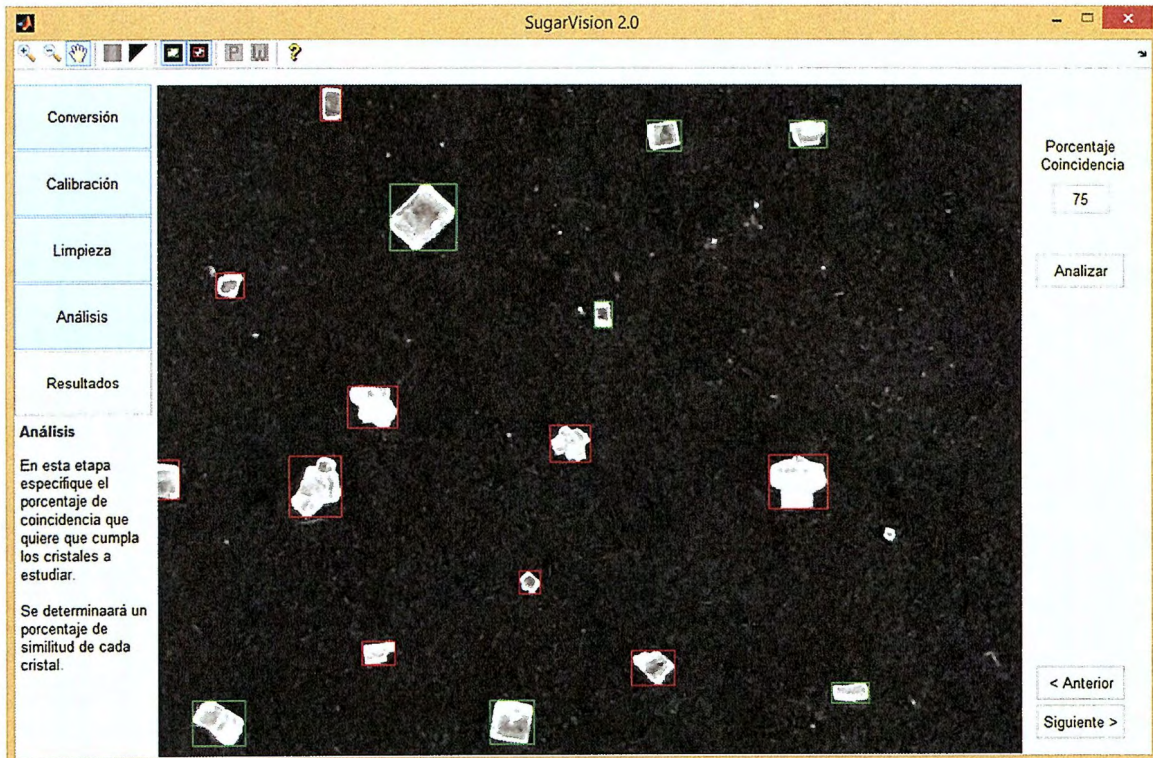


Imagen 4.14 - Detalle de indicadores en RGB. Indicadores rojos y verdes pueden ser activados o desactivados de manera independiente.

4. Cambio de tipo de ponderación en análisis estadístico - Habilitados solo durante la visualización de estadística, sirven para mostrar los resultados con diferentes ponderaciones del análisis estadístico.

- a. Mostrar datos ponderados por partícula: Se muestran media, coeficiente de variabilidad, varianza, desviación estándar y elongación media calculadas en base a partícula. Opción Predeterminada.
- b. Mostrar datos ponderados por peso: Se muestran media, coeficiente de variabilidad, varianza, desviación estándar y elongación media calculadas en base a peso.

5. Herramientas generales

- a. Exportar datos: Se genera una nueva carpeta en un ubicación especificada por el usuario; dentro de esta se guardan tres archivos:
 - i. *Documento de Excel* - Contiene matriz con datos extraídos de los cristales, cantidad de estos, resultados estadísticos y la resolución de calibración de medidas.

Reporte SugarVision 2.0					
Area	Eje Mayor	Diametro Eq.	Datos Ponderados Por Particula		
0.036479713	0.2739412386	0.2155166192	ELONGACION	AM	CV
0.0367915054	0.2503236381	0.2164356708	1.4270645547	0.4027470679	50.82%
0.1446716823	0.5708597177	0.4287245275	Datos Ponderados Por Peso		
0.0801306516	0.3578879198	0.3194143303	ELONGACION	AM	CV
0.3105452491	0.6710380403	0.6288071975	1.3397912079	0.7310315749	31.16%
0.1954938466	0.5402475901	0.4933079019	Numero de particulas		
0.1989235632	0.6145297564	0.4638605923	ANALIZADAS	NO ANALIZADAS	TOTAL
0.0991499892	0.3987312773	0.3445271321	232	57	289
0.1088155542	0.4649621746	0.372220723	RESOLUCION		
0.0770127274	0.3439089666	0.313138388	0.0176576448		
0.0464570704	0.2691736286	0.2432097432			
0.2862254405	0.6659847876	0.6030253532			
0.163067435	0.9235224674	0.4503998995			
0.059552352	0.3240758139	0.2753623241			

Imagen 4.15 - Ejemplo de reporte generado en Excel.

- i. *Archivos de imagen* - Se exportan ambas gráficas de manera independiente en formato PNG.
- b. Acerca de SugarVision: Despliega una pantalla donde se brinda información del software en cuanto al motivo de desarrollo y los nombres de los desarrolladores y del docente responsable del proyecto.

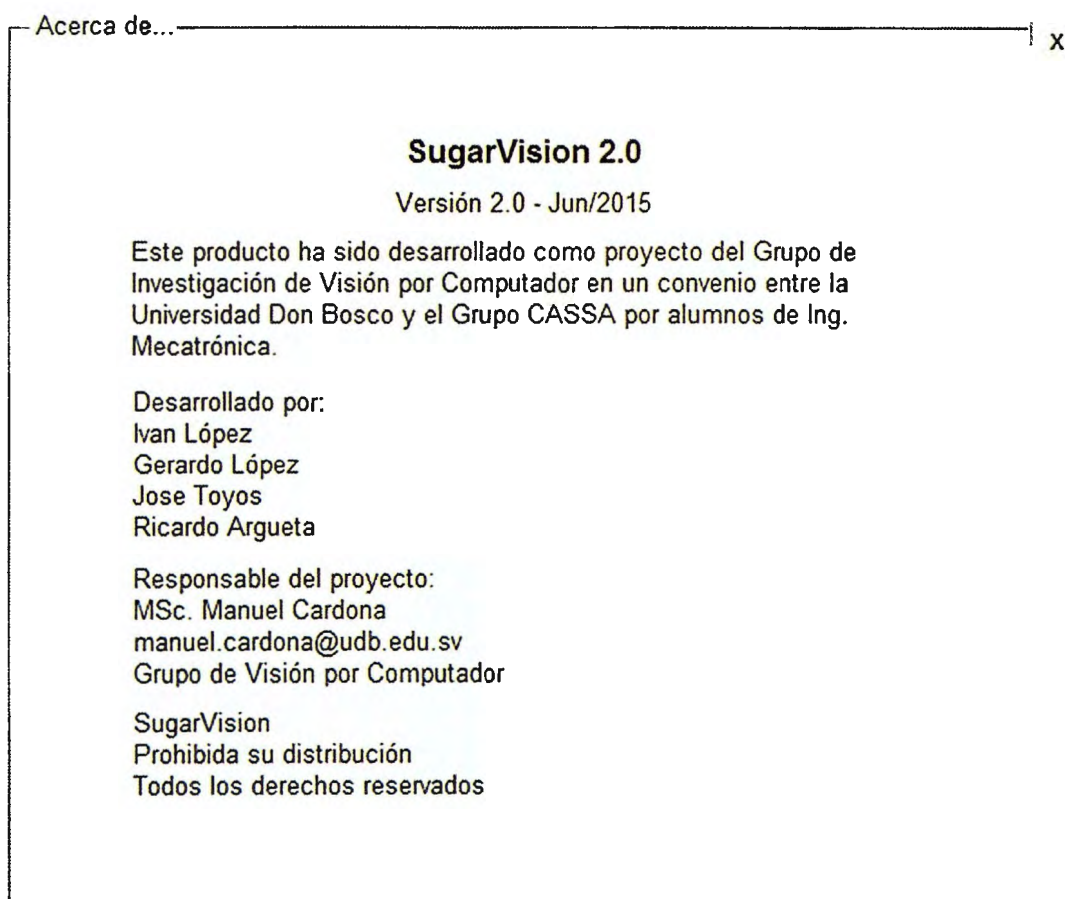


Imagen 4.16 - Ventana de Acerca de SugarVision.

4.5 Consideraciones de Hardware

Se han establecido recomendaciones en cuanto al hardware y las condiciones bajo las cuales realizar la captura de las fotos, con el afán de que el proceso se realice de manera más simple. Una foto realizada bajo las siguientes consideraciones nos brindará beneficios como datos más precisos, un procesamiento más rápido, menos pérdida de características y valores más cercanos a los brindados automáticamente por SugarVision.

A continuación se especifican las condiciones con las cuales se estableció una imagen funcional para el software:

- Fondo - Se necesita un fondo contrastante para facilitar el proceso de segmentación de los cristales en la imagen, además no debe ser reflejante porque afectaría la incidencia de la luz. Debido a estas características se escoge cartón para montaje color negro, este fue sometido a pruebas con diferentes luces y tiene la bondad que no produce reflejos de luz que afecten la imagen.
- Línea de referencia - Esta línea no es opcional y puede ser o no parte del fondo, debe poseer una medida conocida y tener una longitud mayor que cualquier cristal para poder calibrar de manera correcta las medidas dentro del procesos.
- Distancia de la cámara - La distancia entre la cámara y el fondo estará definida por la cantidad de cristales de la muestra, entre más cristales, se necesitará más distancia para abarcar la misma área, por lo cual se tendrá menos resolución y menos precisión en la calibración que los cristales; entre menos cristales, todo lo contrario. Por esta razón SugarVision posee multiselección de imágenes al inicio del proceso, para que se puedan tomar mayor cantidad de fotos concernientes a la misma muestra. Además

se debe optimizar el área de la foto evitando dejar mucho espacio entre la muestra de azúcar y los márgenes de la foto.

- Resolución de cámara - SugarVision es capaz de utilizar fotos provenientes de cualquier dispositivo ya sea profesional o no, siempre y cuando cumpla con las condiciones necesarias como son el fondo y la línea de referencia. Se realizaron pruebas con instrumentos especializados como estereoscopio; cámaras digitales semi y profesionales con zooms ópticos y digitales; y cámara de celulares inteligentes de alta gama. Como se mencionó con anterioridad, la resolución influye mucho en la precisión de la calibración de datos, una mayor resolución de la imagen significa que cada cristal estará formado por mayor cantidad de píxeles, lo cual nos brindará medidas más precisas.
- Iluminación - Es uno de los factores más importantes que determinan la facilidad con que se tratan las imágenes; un exceso de iluminación puede sobreexponer la imagen en cuanto al balance de blancos y no nos permitiría diferenciar de manera correcta los cristales, debido que no existirá mucho contraste con el fondo; el caso contrario, una iluminación pobre nos dará como resultado una imagen muy oscura, en la cual no se podrán distinguir los cristales. Después de realizar pruebas con diferentes tipos y direccionamientos de iluminación natural y artificial se escogieron las siguientes:
 - Iluminación artificial enfocada de manera diagonal a los cristales. Se utilizan dos fuentes de luz LED dispuestas de manera diagonal y paralelas. El hecho que estén contrapuestas nos genera pequeñas sombras laterales que nos ayudan a definir los bordes de los cristales, además de ayudarnos a diferenciar cristales que estén muy cercanos.

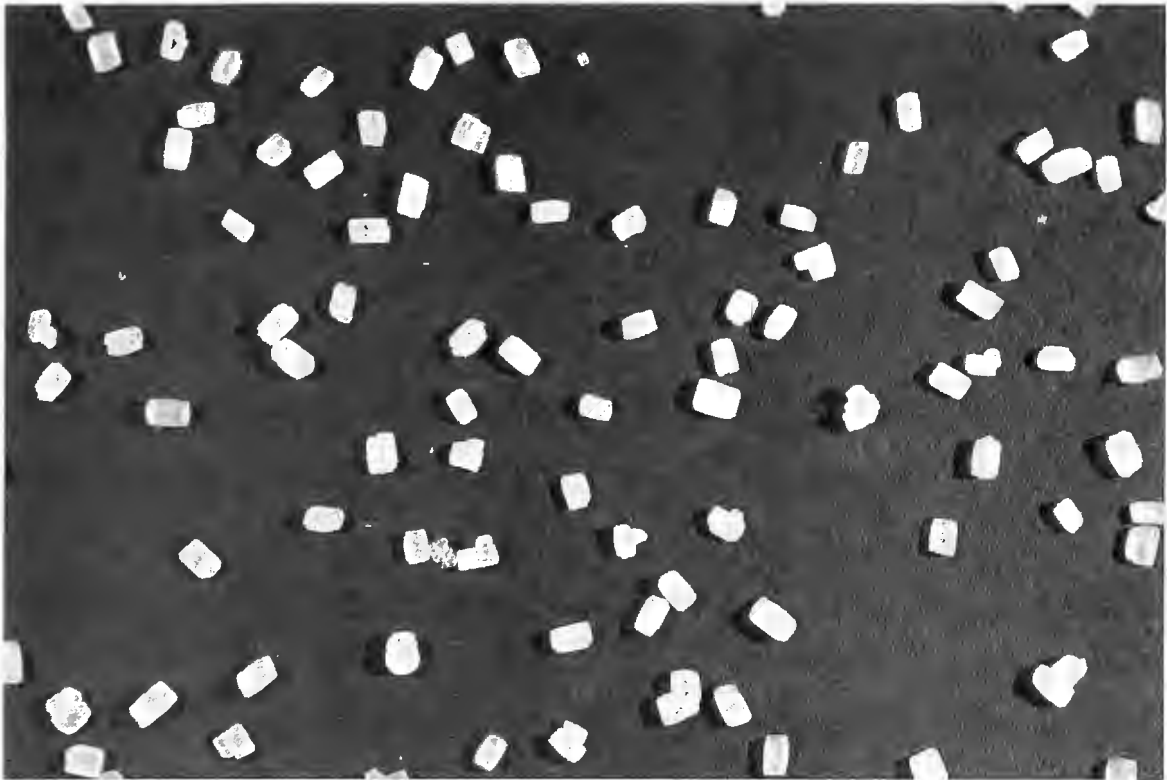


Imagen 4.17 - Detalle de imagen obtenida con la iluminación LED.

- Iluminación natural controlada. Es una buena opción a falta de recursos para realizar el sistema anterior. Una de las grandes desventajas es que las capturas solo se pueden realizar de día. Hay que recalcar la palabra controlada, debido a que una iluminación muy fuerte del sol nos daría como resultado una imagen sobreexpuesta; por lo cual se recomienda que se realice en el interior de un local utilizando solamente la luz natural proveniente de las ventanas del local, teniendo el cuidado que no sea muy fuerte. Este tipo de iluminación, al nivel adecuado, nos da como resultado una imagen con tonos más suaves y menos definición de bordes y contraste. A continuación se muestra un ejemplo

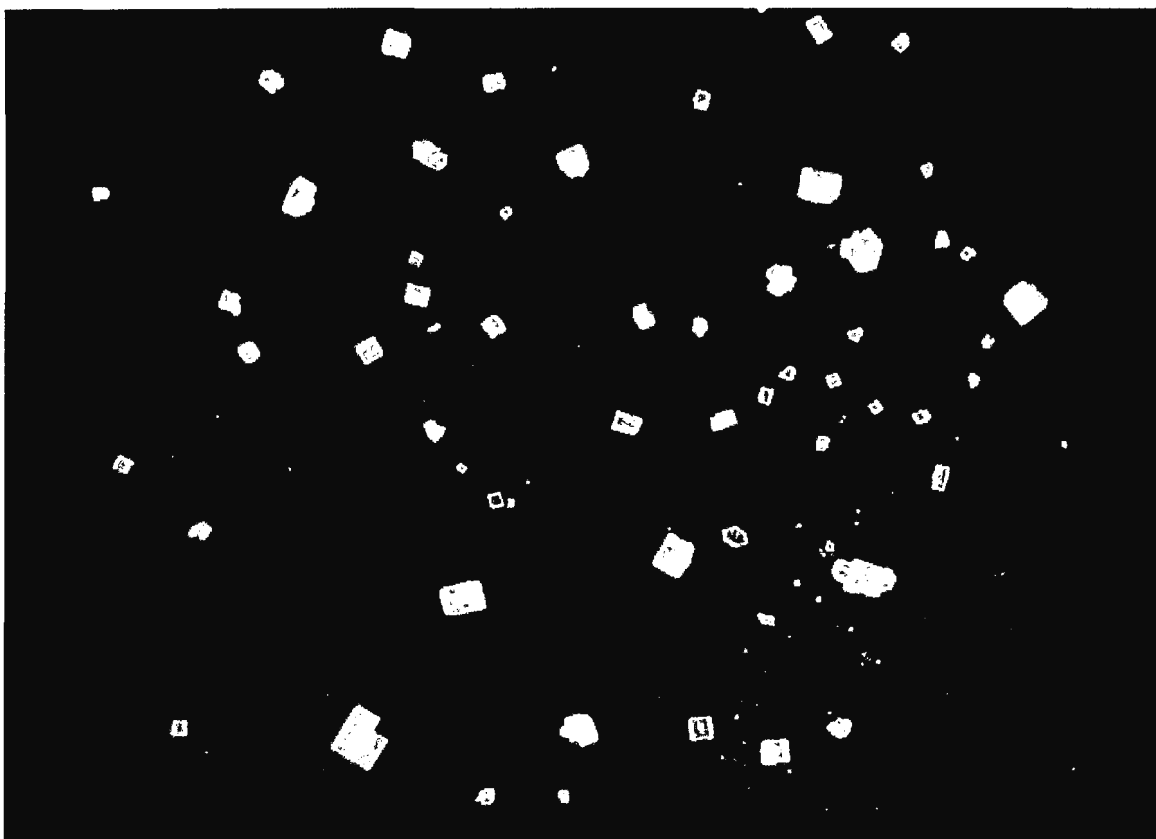


Imagen 4.18 - Detalle de imagen obtenida con iluminación natural.

Capítulo V

Validación de Resultados

5.1 Descripción de métodos utilizados

Para realizar la validación de los resultados que devuelve el software SugarVision se utilizaron dos métodos manuales consecutivos que hacían uso de dos herramientas diferentes; los cuales se detallan a continuación:

5.1.1 Medición manual utilizando software NX

La primera medición manual se realizó utilizando el software de Siemens para aplicaciones CAD/CAM/CAE NX 9. Para llevarlo a cabo, primero se tomaron fotografías de dos muestras de cristales de azúcar de aproximadamente 50 cristales cada una y con valores diferentes de CV y AM, con su respectiva línea de referencia como la siguiente:

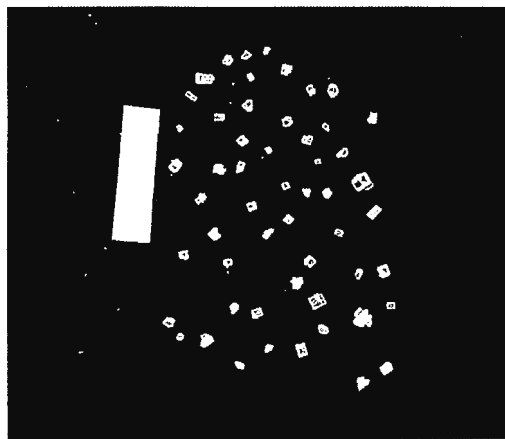


Figura 5.1 - Imagen para validación de 50 cristales de azúcar

Luego estas fotografías se importaron en formato .tiff al programa NX; donde a cada cristal se le dibujo su respectivo *Bounding Box* utilizando la herramienta Rectángulo, tomando en cuenta la orientación del cristal en base al

lado más largo de este a consideración personal. Al realizar esta acción el software automáticamente mostraba las dimensiones de largo y ancho del rectángulo con un escalado desconocido.

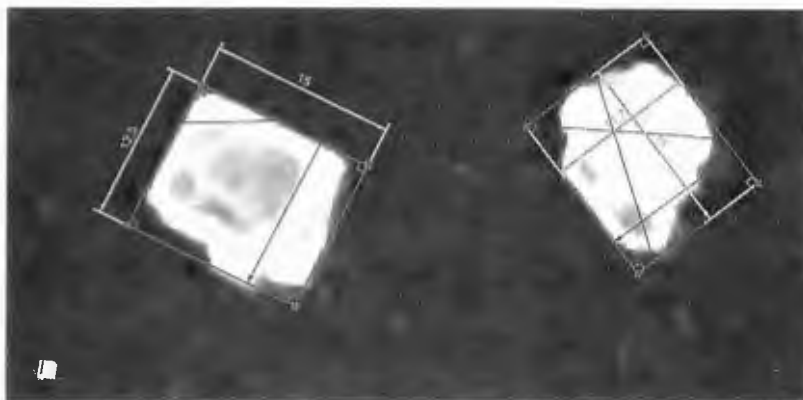


Figura 5.2 - Imagen con Bounding Box manual y sus respectivas dimensiones.

Los datos obtenidos a partir del *Bounding Box* de los cristales se colocaron en una tabla de Excel, donde teniendo también el dato del largo de la línea de referencia y su medida real, que es de 1 cm, se realizó el escalado de las medias de los cristales de la misma manera que se realizan en el software SugarVision a partir de una regla de tres. Además en esta tabla también se colocaron el diámetro equivalente de cada cristal y el valor de AM o diámetro medio de la muestra en milímetros, el cual es el dato de referencia que se compara con los otros métodos posteriormente.

5.1.2 Medición manual utilizando estereoscopio

La medición manual de los cristales de azúcar se realizó mediante un método visual en el cual se utilizó el estereoscopio LEICA EZ4D debido a la capacidad de realizar aumentos desde x8 hasta x35. Para llevarlo a cabo, primero se tomaron las dos mismas muestras de cristales de azúcar del método anterior y se colocan en el portaobjetos el cual posteriormente es posicionado en la platina del estereoscopio.



Imagen 5.3 - Estereoscopio LEICA EZ4D

Luego de que la muestra es colocada en la platina, se procede a realizar el aumento necesario con el objetivo para obtener una imagen de la muestra lo más clara posible y tomarle una foto con su escalado respectivo para determinar las características reales de cada uno de los cristales, cabe resaltar que el estereoscopio está equipado con una cámara, una foto aumentada de los cristales utilizando el estereoscopio se muestra a continuación.



Imagen 5.4 - Muestra de azúcar con aumento del estereoscopio.

Se procede a imprimir las fotos y mediante escuadras graduadas obtener el lado más largo y lado más cortos de cada uno de los cristales; cabe mencionar que las mediciones deben ser multiplicadas por el factor de aumento del estereoscopio, y a partir de los resultados se obtiene el diámetro equivalente el cual será utilizado para encontrar el diámetro medio y el coeficiente de variación ponderado por partículas.

5.2 Comparación de resultados

Definidos los procesos que realizan los otros métodos para la obtención y procesamiento de la información se procede a hacer la comparación entre los resultados de los tres métodos.

En la siguiente tabla se muestran los resultados obtenidos en los tres métodos:

Tabla 5.1

Comparación de resultados de AM y CV de los diferentes métodos de validación

	Medición con NX		Estereoscopio		SugarVision	
	AM	CV	AM	CV	AM	CV
Muestra 1	0.6459	-	0.632	33.99%	0.6017	33.8%
Muestra 2	0.862	-	0.783	27.97%	0.731	28.65%

Para tener una medida de comparación de los métodos utilizados, se procederá a calcular el porcentaje de error que existe entre los resultados de las mediciones manuales y los resultados del software SugarVision:

5.2.1 Comparación NX vs. SugarVision

Debido que la medición con NX se realizó principalmente para verificar si el software SugarVision tomaba de forma correcta las dimensiones de los cristales de azúcar, únicamente se calculó el valor de diámetro medio (AM).

Muestra 1:

$$\frac{0.6459 - 0.6017}{0.6459} \times 100\% = 6.84\%$$

Muestra 2:

$$\frac{0.802 - 0.731}{0.802} \times 100\% = 8.85\%$$

5.2.2 Comparación Estereoscopio vs. SugarVision

Muestra 1:

AM

$$\frac{0.632 - 0.6017}{0.632} \times 100\% = 4.79\%$$

CV

$$\frac{0.3399 - 0.338}{0.3399} \times 100\% = 0.558\%$$

Muestra 2:

AM

$$\frac{0.783 - 0.731}{0.783} \times 100\% = 6.64\%$$

CV

$$\frac{0.2797 - 0.2865}{0.2797} \times 100\% = 2.43\%$$

Debido a que los datos de AM con los que se trabaja están en el orden de los milímetros, es posible que un pequeño cambio de decimas de milímetro cause un aumento considerable en el porcentaje de error. Por tanto se toma como porcentaje aceptable un máximo de 10%; por lo que se puede decir que al comparar los datos de las mediciones realizadas por medio del software NX y la utilización de un estereoscopio con los datos obtenidos por el software SugarVision, estos son aceptables.

5.2.3 Resultado de las comparaciones

Con estos datos comparativos se puede decir que el software SugarVision realiza correctamente la toma de dimensiones de la mayoría de granos de una muestra de azúcar, de tal forma que este proceso puede ser comparado con la toma de dimensiones que realizaría una persona de forma manual; resultando así, en la devolución de datos bastante confiables y útiles en la industria.

Conclusiones

El software SugarVision es una herramienta útil y confiable que puede ser utilizada en aplicaciones de control de calidad dentro de la industria alimenticia azucarera; ya que se ha demostrado que es capaz de tomar de forma precisa las dimensiones de los cristales de azúcar e ingresarlas dentro de un modelo estadístico robusto que devuelve como resultado datos de mucha utilidad para el desarrollo del control de calidad de dichos productos; reduciendo de forma considerable tanto el tiempo como el grado de esfuerzo y fatiga de los encargados de realizar dicha actividad.

En la validación de los datos que devuelve el software se utilizaron dos métodos de comparación que usualmente se utilizan en la industria; los cuales se basan en la toma de las dimensiones de los granos de azúcar de forma manual por parte de una persona. Cuyos resultados fueron similares a los entregados por el software SugarVision; presentando un porcentaje de error menor al 10%, el cual se considera aceptable para este tipo de proceso ya que se toma en cuenta que los métodos de comparación que se utilizaron son más sensibles al error humano; garantizando así, la confiabilidad de los datos que devuelve dicho software.

Mediante la prueba de distintas herramientas estadísticas se determinó que la elección de las herramienta, tanto de muestreo como de análisis de información, deben escogerse de la manera más cautelosa posible; ya que fácilmente puede perderse información, por lo que al software SugarVision se hicieron las pruebas necesarias hasta obtener los resultados deseados.

Para asegurar el correcto funcionamiento del software SugarVision así como el grado de exactitud de los resultados que devuelve, se deben tomar las fotografías tomando en cuenta las consideraciones de hardware de fondo, línea de referencia, distancia y resolución de cámara digital y tipo de iluminación.

Bibliografía

- Cuevas, E., Zaldivar, D., & Pérez, M. (2010). *Procesamiento digital de imágenes con MatLab & Simulink* (Primera ed.). Alfaomega Grupo Editor, S.A. de C.V.
- España, Ministerio de Educación. (s.f.). Educación Inclusiva: Discapacidad Visual. Instituto de Tencologías Educativas. Obtenido de http://www.ite.educacion.es/formacion/materiales/129/cd/pdf/m1_dv.pdf
- Esqueda Elizondo, J. J., & Palafox Maestre, L. E. (2005). *Fundamentos de procesamiento de imágenes* (Primera ed.). Departamento Editorial Universitaria de Baja California.
- Gavves, E. (11 de Diciembre de 2014). *A brief history of Computer Vision (and not time)*. Obtenido de Efstratios Gavves - Computer Vision & Machine Learning: <http://www.egavves.com/a-brief-history-of-computer-vision/#sthash.IFptEIUi.gvnzwiaZ.dpbs>
- Gomez-Alende, D. M. (1993). *Reconocimiento de formas y visión artificial* (Primera ed.). Editorial RA-MA y Addison-Wesley Iberoamericana, S.A.
- Kenett, R., & Zacks, S. (1998). *Estadística industrial moderna. Diseño y control de la calidad y la confiabilidad* (Primera ed.). Matemáticas Thomson.
- MathWorks. (2015). *Documentation*. Obtenido de MathWorks: <http://es.mathworks.com/help/index.html>
- Moler, C. (2004). *The Origins of MATLAB*. Obtenido de MathWorks: <http://es.mathworks.com/company/newsletters/articles/the-origins-of-matlab.html>

- *Muestreo Estadístico*. (s.f.). Obtenido de La gran enciclopedia de economía: <http://www.economia48.com/spa/d/muestreo-estadistico/muestreo-estadistico.htm>
- *Muestreo Indexado según AQL: MIL-STD-105E*. (s.f.). Obtenido de Gestion de Calidad Total: http://www.gestiondecalidadtotal.com/mil_std_105e.html
- Pajares Martisanz, G. (2008). *Visión por computador: Imágenes digitales y aplicaciones*. RA-MA. Obtenido de http://dmi.uib.es/~ygonzalez/VI/Material_del_Curso/Teoria/Aplicaciones_VC.PDF
- Passariello, G., & Mora, F. (1995). *Imágenes Médicas: Adquisición, análisis, procesamiento e interpretación* (Primera ed.). Ediciones de la Universidad Simón Bolívar.
- Prieto, M., Mouwen, J. M., López Puente, S., & Cerdeño Sánchez, A. (2008). *Concepto de calidad en la industria Agroalimentaria*. Obtenido de Scielo: http://www.scielo.org.ve/scielo.php?pid=s0378-18442008000400006&script=sci_arttext
- *Short history of FORTRAN/Fortran*. (1998). Obtenido de User notes on Fortran Programming: <http://www.ibiblio.org/pub/languages/fortran/ch1-1.html>
- Somolinos Sánchez, J. A. (2002). *Avances en robótica y visión artificial* (Primera ed.). Ediciones de la Universidad de Castilla-La Mancha.
- Tadiou, K. M. (2013). *Artificial Intelligence: Computer Vision*. Obtenido de The future of human evolution: <http://futurehumanevolution.com/artificial-intelligence-future-human-evolution/artificial-intelligence-computer-vision>