



EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO®

**INSTITUTO TECNOLÓGICO DE
DURANGO**

**INSTITUTO TECNOLÓGICO DEL
VALLE DEL GUADIANA**

**DIVISIÓN DE ESTUDIOS DE
POSGRADO E INVESTIGACIÓN**



**Análisis por software de la calidad de semen bovino crio-
preservado destinado a la Inseminación**

TESIS

Que como parte de los requisitos para obtener el grado de

Maestro en Ingeniería

Presenta: Ing. Miguel Mijares Chairez

Director(a) de tesis: MC. Aurelio Castillo Liñán

Codirector(a): Dr. Rubén Guerrero Rivera

Durango, Dgo. México

Mayo 2023





**Análisis por software de la calidad de semen bovino crio-preservado
destinado a la Inseminación**

Presenta: Ing. Miguel Mijares Chairez

COMITÉ TUTORIAL

MC. Aurelio Castillo Liñán Director(a)	 Firma
Dr. Rubén Guerrero Rivera Codirector(a) o asesor(a)	 Firma
Dra. Yolocuauhtli Salazar Muñoz Asesor(a)	 Firma
MC. Norma A. García Vidaña Asesor(a)	 Firma

MC. Norma A. García Vidaña

**Coordinador(a) del programa del
posgrado correspondiente**

MC Adriana Eréndira Murillo

**Jefe(a) de la División de Estudios de
Posgrado e Investigación**

Durango, Dgo. México

noviembre 2022



AGRADECIMIENTOS

Comité tutorial:

MC. Aurelio Castillo Liñán

Dr. Rubén Guerrero Rivera

Dra. Yolocuahutli Salazar Muñoz

MC. Norma A. García Vidaña

Especialmente:

Al Consejo Nacional de Ciencia y Tecnología (Conacyt)

Publicaciones:

Mijares, M. M. C. (2022). Sistema de análisis de esperma asistido por computadora de código abierto. Revista del centro de graduados e investigación. Instituto Tecnológico de Mérida, 37(93), 0185-6294. <https://www.revistadelcentrodegraduados.com/2022/06/>



Victoria de Durango, Dgo., a 29 / Mayo / 2023

DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN
DEPI / C / 192 / 23.

ASUNTO: Autorización de Tema de Tesis de Maestría.

C. MIGUEL MIJARES CHAIREZ
No. DE CONTROL G20040549
PRESENTE.

Con base en el Reglamento en vigor y teniendo en cuenta el dictamen emitido por el Jurado que le fue asignado, se le autoriza a desarrollar el tema de tesis para obtener el **Grado de Maestro en Ingeniería** cuyo título es:

"ANÁLISIS POR SOFTWARE DE LA CALIDAD DE SEMEN BOVINO CRIO-PRESERVADO DESTINADO A LA INSEMINACIÓN"

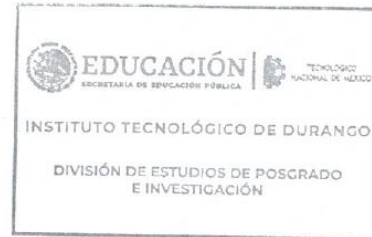
CONTENIDO:

	RESUMEN
CAPÍTULO I	INTRODUCCIÓN
CAPÍTULO II	MARCO TEÓRICO
CAPÍTULO III	METODOLOGÍA
CAPÍTULO IV	RESULTADOS
CAPÍTULO V	CONCLUSIONES
CAPÍTULO VI	REFERENCIAS BIBLIOGRÁFICAS

ATENTAMENTE.

Excelencia en Educación Tecnológica®
"La Técnica al Servicio de la Patria"

C. ADRIANA ERÉNDIRA MURILLO
JEFA DE LA DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN



AEM'ammc.



Fecha de Emisión: 09/05/2023
Fecha de Validación: 09/05/2023
Fecha de Expiración: 09/05/2023
ASIS 987





EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO

Instituto Tecnológico de Durango
División de Estudios de Posgrado e Investigación

Victoria de Durango, Dgo., a **29 / Mayo / 2023.**

DIVISIÓN DE ESTUDIOS DE POSGRADO E INVESTIGACIÓN
DEPI / C / 193 / 23.

ASUNTO: Autorización de Impresión de Tesis de Maestría.

C. MIGUEL MIJARES CHAIREZ
No. DE CONTROL G20040549
PRESENTE.

De acuerdo al reglamento en vigor y tomando en cuenta el dictamen emitido por el jurado que le fue asignado para la revisión de su trabajo de tesis para obtener el **Grado de Maestro en Ingeniería**, esta División de Estudios de Posgrado e Investigación le autoriza la impresión del mismo, cuyo título es:

“ANÁLISIS POR SOFTWARE DE LA CALIDAD DE SEMEN BOVINO CRIO-PRESERVADO DESTINADO A LA INSEMINACIÓN”

Sin otro particular de momento, quedo de Usted.

ATENTAMENTE.

Excelencia en Educación Tecnológica®
"La Técnica al Servicio de la Patria"

C. ADRIANA ERÉNDIRA MURILLO
JEFA DE LA DIVISIÓN DE ESTUDIOS DE
POSGRADO E INVESTIGACIÓN



AEM/ammc



Av. Felipe Pescador #1830 Ote. Col. Nueva Vizcaya C.P.34060 Durango, Durango.
Tel. (618) 2290900 e-mail: dir_itdurango@tecnm.mx; tecnm.mx | itdurango.edu.mx



2023
Francisco VILLA
100 años del Centenario del México Libre

Contenido

1	Introducción.	1
1.1	Antecedentes.	3
1.1.1	Inseminación Artificial.....	3
1.1.2	Sistemas CASA.....	3
1.2	Planteamiento del problema.....	6
1.3	Objetivo.	8
1.3.1	Objetivo General.	8
1.3.2	Objetivos específicos.....	8
1.4	Justificación.....	9
2	Marco Teórico.....	11
2.1	Imágenes Microscópicas.....	11
2.1.1	Resolución, magnificación y contraste en microscopio.	11
2.1.2	Microscopio de contraste de fase.....	12
2.1.3	Imagen óptica e imagen digital.....	13
2.2	Parámetros característicos de una imagen.....	14
2.2.1	Sensibilidad.....	14
2.2.2	Ruido.....	14
2.2.3	Resolución espacial.....	15
2.2.4	Linealidad y uniformidad.....	15

2.2.5	Tiempo de respuesta.....	16
2.2.7	Rango dinámico.....	16
2.2.8	Ganancia.....	16
2.2.9	Velocidad de fotogramas (frame rate).....	16
2.3	Adquisición de imagen.....	18
2.4	Introducción al procesamiento de imágenes.....	19
2.4.1	Digitalización de imágenes.....	19
2.4.2	Píxeles.....	20
2.5	Métodos computacionales.....	21
2.6	Fuentes de imágenes.....	23
2.7	Métodos de procesamiento.....	24
2.7.1	Reducción de ruido.....	24
2.7.2	Segmentación.....	25
2.7.3	Umbrales y discriminación.....	27
2.7.4	Imagen binaria.....	28
2.7.5	Erosión y dilatación.....	28
2.8	Medición de imágenes.....	31
2.8.1	Mediciones globales.....	31
2.8.2	Medición de características.....	31
2.8.3	Longitud y amplitud.....	32
2.8.4	Posición.....	33

2.9	Seguimiento.....	35
3	Metodología.....	37
3.1	Características de la imagen/video a analizar.....	37
3.2	Segmentación y detección de partículas.....	38
3.3	Estructura del software de análisis.....	40
3.4	Algoritmo de seguimiento.....	42
3.5	Análisis de datos.....	44
3.6	Presentación de los datos.....	46
4	Resultados.....	49
5	Conclusiones.....	51
6	Referencias.....	53

Lista de figuras.

Fig. 1.	Diagrama de flujo del proceso de análisis.....	40
Fig. 2.	Diagrama de flujo del algoritmo de seguimiento.....	43
Fig. 3.	Interfaz del software CASA.....	46
Fig. 4.	Trayectorias dibujadas sobre el video original.....	47

RESUMEN

Actualmente existe una gran variedad de sistemas de análisis de esperma que cumplen con su propósito de manera satisfactoria, los más conocidos están sujetos a una marca comercial lo que los convierte generalmente en una inversión difícil para laboratorios con presupuestos limitados debido a su elevado costo, normalmente es el caso de veterinarios asociados a pequeños productores. En este estudio se describe el procedimiento de desarrollo de un software de análisis de esperma asistido por computadora, las características principales que debe contener para cumplir con su objetivo, así como las diferentes rutas de desarrollo que se pueden seguir con la tecnología actual para integrarlo en un sistema de bajo costo accesible incluso para productores independientes. El programa se implementa en lenguaje Python con el objetivo de convertirse en un programa completamente libre para cualquier persona además queda abierto a la posibilidad de añadir mejoras constantes por parte de la comunidad. Por último, se hacen algunas recomendaciones para mejoras futuras como la posible aplicación de un modelo de machine learning anticipando el surgimiento de las bases de datos necesarias.

Palabras clave: CASA, Código abierto, esperma bovino, Python.

ABSTRACT

Currently there is a huge variety of sperm analysis systems that fulfill satisfactorily well its purpose, the most renowned ones are linked to a trademark which makes them unaffordable to small and mid-size laboratories due to high cost and limited budgets, this is the case of veterinaries associated to independent farmers. This study describes the different stages of the development of a Computer Aided Sperm Analysis software, including the requirements necessary to reach its objective, as well as the possible developing alternatives within the current state of the art of the technologies, which could leverage it to make it fully integrated into a low-cost product affordable to most of the independent stockbreeders. The software was implemented in python to keep it open and free to the users. The software is still susceptible to continuous improvements as well as modifications from the users. Finally, the paper contains some suggestions for future work, such as the implementation of a learning algorithm to enhance the performance, this will be important as more data become available.

Keywords— Bull Sperm, CASA, Open Source, Python.

Capítulo I

1 Introducción.

Desde hace mucho tiempo la comunidad científica ha trabajado en el desarrollo de métodos de captura de imágenes y formas de describir las características cinemáticas de espermatozoides. Los primeros intentos de captura y seguimiento se realizaron mediante la proyección de imágenes microscópicas de células espermáticas en películas fotográficas, estas cinemicrografías lograron capturar secuencias de imágenes en un rango de 50 a 200 cuadros por segundo (Accot, Katz, Hoskins, 1983) (Mortimer, 1997). Con esta capacidad de captura fue posible describir por primera vez incluso algunos patrones de movimiento del flagelo. Estos primeros intentos dependían de la habilidad del investigador para interpretar las imágenes y dibujar manualmente las trayectorias basándose en las secuencias de las películas fotográficas (Mortimer, 1997), (Beresford-Smith & Van Helden, 1994), (Gray, 1958). Con el paso del tiempo y la adición de las computadoras a la fórmula el análisis de imágenes de esperma se convirtió en un tema de investigación que ha llevado al desarrollo de softwares de análisis cada vez más sofisticados y precisos capaces de determinar cada vez más características dinámicas y morfológicas de forma automática (Gray, 1958), (Amann & Wabersi, 2014), (Sørensen, 2008).

Un sistema de análisis CASA consiste típicamente de un microscopio óptico equipado con lentes objetivo de 10X-20X y capacidad de contraste de fases, una cámara digital adaptada al microscopio con una frecuencia de captura de 60 fps o mayor que graba las imágenes microscópicas, las cuales son procesadas y analizadas por medio de un software especializado (Daloglu & Ozcan, 2017). Los análisis de esperma se llevan a cabo en cámaras de 20 μ m de profundidad lo cual limita el movimiento vertical de las células y se analizan campos de un área típicamente de 250 μ m por 250 μ m en las cuales pueden verse hasta 200 espermatozoides (Jaqaman, 2008). El software de análisis hace uso de técnicas de filtrado y segmentación de imágenes para identificar las células, y complejos algoritmos de seguimiento conectan las posiciones espacio-temporales para dibujar las trayectorias (Aitken, 2008). Posteriormente cuando se tienen los datos de cada trayectoria ordenados se aplica un análisis que sirve para calcular las variables que ayudan a definir la calidad de la muestra tales como la concentración de espermatozoides por mililitro, el porcentaje de movilidad progresiva y forma normal (Mortimer, 2000). Y otras más complejas como la velocidad en línea recta (VSL), velocidad curvilínea (VCL), linealidad (LIN), amplitud del desplazamiento lateral de cabeza (ALH) (Katz, 1985), las cuales han probado ser un factor determinante en algunos estudios donde la concentración y movilidad progresiva no son suficientes para explicar variaciones en resultados experimentales (Mortimer, 2000), (Katz, 1985), (Sørensen, 2008). Los avances en técnicas de detección y seguimiento de esperma son un tema de estudio actual el cual añade técnicas cada vez más complejas como un filtro de partículas, filtro de Kalman, agrupación en clústeres por

medio de secuencias de coordenadas espacio-temporales (Urbano, 2016), el seguidor multi-hipotesis (MHT) o el filtro de asociación de datos probabilísticos conjuntos (JPDAF) (Yang, 2014), o algoritmos especializados como el de seguimiento de cabeza y rastreo del flagelo (García, 2012).

En el presente artículo se describe el proceso de desarrollo de un software de análisis de esperma implementado en Python haciendo uso de librerías recientes dedicadas al procesamiento de imágenes, implementación de interfaces y manipulación de matrices, además de crear un algoritmo de seguimiento para múltiples objetivos.

1.1 Antecedentes.

1.1.1 Inseminación Artificial.

Los orígenes de la inseminación artificial se remontan al siglo XIV cuando caciques árabes emplearon por primera vez técnicas para cruzar valiosos caballos de manera artificial (Foote, 2005). En el siglo XVII en Holanda Leeuwenhoek y su asistente hicieron las primeras observaciones de espermatozoides por medio de lentes amplificadores. En el año 1784 un profesor de historia natural en Pavía, Spallanzani llevo a cabo la primera inseminación exitosa de una perra la cual produjo tres cachorros 62 días después (Foote 2002).

En Rusia bajo la dirección de Ivanov se logró establecer la inseminación artificial como un procedimiento practico donde fue desarrollada para la reproducción de animales de granja, zorros, conejos, y aves de corral. Ivanov extendió las técnicas para ser utilizadas en vacas y ovejas con excelentes resultados, en 1909 estableció el primer laboratorio donde se capacito a personal veterinario en las técnicas de IA. Gran parte del desarrollo de la IA en Rusia fue después continuada por Milovanov quien estableció centros especializados de reproducción de ovejas y bovinos (Milovanov, 1964).

En 1906 en estados unidos se creó una técnica donde se colocó semen bovino en una capsula de gelatina y se implanto en una vaca en celo resultando preñada en el proceso. Más tarde se utilizó esta misma técnica para preñar 800 yeguas con el semen de un solo semental (Herman, 1981).

1.1.2 Sistemas CASA.

Computer assisted sperm analysis (CASA) traducido a sistemas de análisis de esperma asistido por computadora, son sistemas computarizados desarrollados desde hace más de 40 años para el análisis de movilidad de esperma que utiliza técnicas de alto grado de sofisticación y provee datos de otras características de movimiento (basándose en la observación de células individuales) muy relevantes en la predicción de fertilización exitosa. Además, provee información con datos útiles en el estudio de efectos de varios procedimientos in-vitro, también ayuda en el análisis del fenómeno de hiper-activación de esperma.

Desde las primeras observaciones de espermatozoides por Leeuwenhoek en 1678 el microscopio se convirtió en una de las herramientas principales en el análisis de esperma, pero no fue hasta mediados de los 1950 cuando se utilizó por primera vez un microscopio de contraste de fases en un laboratorio de andrología, que es el mismo instrumento que se utiliza hoy en día en los sistemas casa actuales, esto debido a su capacidad de aumentar el contraste de las imágenes y detectar el

contorno de cada célula translúcida lo cual permite una visualización más clara (Amann & Katz, 2004).

Durante las décadas de 1960 y 1970 se extendió el uso de microscopios ópticos adaptados con cámaras digitales dando paso así al surgimiento de las micrografías las cuales son fotografías capaces de mostrar objetos microscópicos imposibles de ver a simple vista, durante esta época el análisis se realizaba aun manualmente pero las micrografías permitieron una comprensión más profunda sobre las trayectorias y velocidades de células mejorando así la clasificación de poblaciones dentro de las muestras, sin embargo hubo aquellos que se percataron de las limitaciones que suponía depender únicamente de la habilidad humana para detectar los datos complejos de movilidad y comportamiento del espermatozoide a simple vista, esta preocupación impulso en gran medida el desarrollo posterior de los sistemas CASA hacia un enfoque más computarizado (Amann & Katz, 2004).

La capacidad de una computadora de realizar cálculos matemáticos complejos en fracciones de segundo ha permitido su integración en infinidad de aplicaciones, fue a mediados de 1980 cuando surgió por primera vez la combinación de un microscopio de contraste de fases con una computadora capaz de analizar imágenes digitales, este siendo el primer sistema CASA desarrollado (Mortimer, 2000). La demanda creciente por un sistema de análisis automático impulso el desarrollo de sistemas más sofisticados, originalmente el primer sistema capturaba imágenes a una frecuencia de 4Hz a 5Hz, actualmente los sistemas CASA más comunes utilizan frecuencias de captura de 30Hz-60Hz gracias a los estándares de captura de video de la mayoría de cámaras digitales, sin embargo, para algunas aplicaciones es posible encontrar equipos capaces de lograr frecuencias de captura mayores a 100Hz. Sistemas de análisis actuales son capaces de ofrecer intervalos de confianza de hasta 95%.

Los sistemas CASA tienen un gran potencial de desarrollo, en los últimos años han surgido sistemas especializados categorizando a los sistemas dependiendo de su enfoque de análisis, sistemas muy sofisticados dedicados al análisis de movilidad son capaces de detectar incluso el movimiento del flagelo ofreciendo así una gran cantidad de información útil en el análisis de células individuales. Por otro lado, están aquellos sistemas enfocados en el análisis morfológico que ofrecen capturas de células con una definición de muy alta calidad con las que es posible detectar anomalías antes invisibles (Lu, 2012).

Aun con todos los avances que ha habido en el desarrollo de los sistemas CASA, todavía es necesaria la intervención humana para llevar a cabo un análisis de espermatozoide. Enormes incrementos en la potencia de las computadoras, así como la disminución de su tamaño, costo y el desarrollo de más y más lenguajes de programación contrastan enormemente con los conceptos básicos del análisis de

esperma y sus patrones de movimiento, que permanecen iguales a través de los años. Por lo que sistemas antiguos y más lentos siguen en uso actualmente, por lo que el grado de confiabilidad varía enormemente de un sistema a otro (Lu, 2014).

1.2 Planteamiento del problema.

Semen crio-preservado

Uno de los métodos más comunes para el almacenamiento de muestras de semen es la crio-preservación, esta técnica consiste en la congelación de las muestras para ralentizar su deterioro natural, luego se descongela al momento de su utilización. Una muestra de semen almacenada en dióxido de carbono en alcohol a -79°C permanece fértil durante de 8 años, Alternativamente, si se almacena en nitrógeno líquido a -196°C la viabilidad del esperma se extiende hasta 12 años (Salisbury, 1978).

El proceso de congelar y descongelar el semen resulta en un decremento significativo del porcentaje de células funcionales dentro de las muestras sometidas al proceso, aun cuando este se realiza correctamente (Holt, 1997). Entre las causas de dicho decremento de células funcionales se encuentran la formación de cristales de hielo dentro de la célula, estrés osmótico, daño oxidativo, envejecimiento prematuro, y cambios de tipo capacitivo (Watson, 2000). Al descongelar las muestras, la permeabilidad de la membrana se ve disminuida lo cual permite un flujo iónico que daña la integridad de la membrana plasmática (Holt & North, 1984).

Es un hecho innegable que la calidad del semen en una muestra sometida al proceso de crio-preservación se ve seriamente disminuida, por lo cual es necesario un método para medir la calidad de dichas muestras. Entre los métodos de análisis de semen conocidos están el examen visual por medio de un microscopio, en el cual un técnico cuenta visualmente las células en buenas condiciones y calcula matemáticamente los valores determinantes en la calidad de dicha muestra, otro método es la utilización de un sistema de análisis de esperma asistido por computadora (CASA), el cual por medio de una cámara adaptada a un microscopio es capaz de generar imágenes digitales las cuales son analizadas en un software especializado que cuenta y calcula los datos necesarios en cuestión de segundos.

Cada uno de los métodos mencionados tiene ventajas y desventajas frente al otro. en la inspección visual está presente la posibilidad de errores humanos y la necesidad de invertir mucho tiempo, mientras que con el uso del sistema CASA la inversión para adquirir la licencia del uso de software puede estar fuera de las posibilidades de algunos productores, más aún si se suma la inversión del equipo de análisis y la capacitación en el uso del sistema (Lu, 2014).

En el presente documento se describe el desarrollo de un sistema de análisis de esperma basado en lenguaje Python el cual es capaz de calcular los valores de las características más importantes que determinan la calidad de una muestra de

esperma. Esto con el fin de dar solución a los problemas que surgen mediante el uso de las técnicas antes mencionadas.

1.3 Objetivo.

1.3.1 Objetivo General.

Determinar la calidad del semen de ganado bovino para la validación de un sistema de crio-preservación, por medio de técnicas computarizadas de procesamiento y análisis de imágenes digitales tomadas con microscopio.

1.3.2 Objetivos específicos.

Determinar los valores de las diferentes características que influyen en la calidad del esperma (morfológicas, de movilidad, concentración).

Seleccionar las técnicas de análisis basadas en OpenCV óptimas que se aplicaran para obtener predicciones de calidad avanzadas.

Crear un algoritmo de análisis que aplique las técnicas seleccionadas, recopile e interprete los datos obtenidos de cada imagen/video.

Integrar todo el método en un programa basado en Python, con una interfaz intuitiva, confiable y capaz.

Validación del software comparando los resultados con los de un técnico experto.

1.4 Justificación.

La Producción de bovinos en el estado de Durango.

El estado de Durango aporta el 11.5% de la producción nacional de leche de vaca, según el reporte de la Secretaría de Agricultura y Desarrollo Rural (Tercer lugar nacional). Además de la leche el estado de Durango aporta el 5.5% a la producción nacional, de carne de bovino en canal (Cuarto lugar nacional). Y es el tercer principal exportador de ganado bovino en pie a Estados Unidos, en el ciclo (2020-2021) se exportaron 169,570 cabezas de Ganado (Secretaría de Agricultura, Ganadería y Desarrollo Rural, 2021).

Estos datos ubican al estado de Durango como uno de los principales productores de bovinos del país, por lo que es necesario la implementación de tecnologías que faciliten el desarrollo de este sector. Un dato poco alentador sobre la producción de bovinos en Durango y en México en general es que la producción por inseminación artificial es utilizada casi exclusivamente en la industria lechera la cual esta monopolizada por un par de empresas millonarias, el 80% de bovinos destinados a la producción de leche es producido mediante la inseminación artificial, mientras que por el lado de los bovinos productores de carne, únicamente el 5% es producido mediante inseminación, esto puede deberse a varios factores como la gran inversión necesaria para implementar un laboratorio de reproducción, la cual únicamente está al alcance de grandes productores.

Un nuevo modelo de análisis de calidad de esperma de ganado bovino, permitirá corroborar la viabilidad del esperma crio-preservado y así aumentar la probabilidad de éxito en el proceso de inseminación artificial al evitar el uso de unidades defectuosas, sin la necesidad de invertir grandes capitales en un equipo de análisis convencional.

Siendo también una tecnología pensada para promover el desarrollo ganadero local, se ubicará como una herramienta al alcance incluso de pequeños productores, propiciando un crecimiento del sector en todos sus niveles.

Desarrollado mediante un software de código abierto abre las posibilidades para implementar nuevos métodos continuamente, mantener la aplicación constantemente actualizada y estar disponible para nuevos desarrolladores interesados en aportar innovaciones.

Capítulo II

2 Marco Teórico.

2.1 Imágenes Microscópicas.

El ojo humano, aunque profundo en su capacidad como órgano sensorial, está limitado en muchos aspectos. Por sí mismo puede detectar únicamente longitudes de onda en el rango visible, esto es de 400 a 700 nm, y puede percibir únicamente contrastes formados por diferencias de intensidades específicas. La resolución del ojo humano está limitada a aproximadamente 0.1mm, además de ser insensible al estado de polarización o diferencias fases en la luz (Rittcher, 2008).

El microscopio es un dispositivo que ayuda a sobrepasar estas limitaciones, permitiéndonos visualizar objetos que son invisibles a simple vista. Un buen microscopio logra esto aumentando la magnificación, el contraste y la resolución en la imagen del objeto formada.

Hoy día existe un gran número de técnicas para llevar a cabo la microscopia, tales como ondas de luz, ondas de electrones, técnicas físicas como la microscopia de barrido. Estas técnicas se utilizan para escanear la muestra en caso de la microscopia de barrido (escaneo por láser, escaneo mediante el uso de electrones) o se transmiten a través de la muestra como en el caso de las microscopias de transmisión de luz o electrones. Estas técnicas nos dan la posibilidad de visualizar y estudiar una gran gama de objetos a micro y nano-escala, que van desde secciones de tejido, células, hasta moléculas individuales y nano-partículas. En las últimas décadas hemos presenciado un agudo aumento en los avances y aplicaciones de técnicas de microscopia para el estudio de muestras biológicas, con innovaciones en el diseño de microscopio, técnicas de preparación de muestras y adquisición de imágenes (Rittcher, 2008).

2.1.1 Resolución, magnificación y contraste en microscopio.

El microscopio original hace más de 500 años consistía de una sola lente amplificadora, alrededor del 1600 fue desarrollado el microscopio compuesto en Holanda e Italia. El microscopio compuesto actual dista mucho del diseño de estos dispositivos, hoy en día un microscopio compuesto consta entre otros elementos de un sistema de iluminación, un acondicionador, un condensador, un espécimen, un objetivo, un filtro de imagen, un ocular, y un sistema de detección. El microscopio está definido por dos conjuntos de planos conjugados conocidos como el plano de campo y el plano de difracción. Cada conjunto de planos debe estar correctamente alineado sobre los ejes del microscopio para asegurar una iluminación, formación de la imagen, magnificación y contraste de la muestra apropiada.

Una característica importante en una lente es la apertura numérica (NA), esta característica es la capacidad de la lente para capturar los patrones de difracción de orden mayor. La NA es esencial para generar un patrón de difracción tan real como sea posible, además gobierna también la resolución y el contraste en un microscopio.

La selección de lentes para una microscopia depende de la magnificación deseada, la resolución, el brillo y la técnica de microscopia utilizada. Normalmente para exámenes de rutina en células o secciones de tejido, un lente objetivo 20x es adecuado, por otro lado, para estudios relacionados con características subcelulares o polimerización de proteínas, una lente con una apertura alta como 63x o incluso 100x sería una elección objetiva (Rittcher, 2008).

El contraste en microscopios de transmisión de luz se logra de dos maneras: una es explotando la capacidad del ojo humano para distinguir entre diferentes longitudes de onda (diferentes colores) añadiendo sustancias como tintas o anticuerpos cromóforos que tiñan ciertos objetos dentro de la muestra, otra forma de lograr un buen contraste es generar suficiente modulación en las intensidades de la imagen formada, mediante la observación de diferentes fenómenos ópticos como la absorción de luz, reflexión, variación espacial en índices refractivos, dispersión, difracción, y fluorescencia los cuales están presentes en diferentes componentes de muestras biológicas. Aunque una lente con una apertura numérica alta es benéfica para conseguir una alta resolución esto puede suponer un decremento en el contraste debido a una mayor absorción de luz externa lo cual puede causar fenómenos como astigmatismo (Murphy, 2001). Por lo cual al elegir una lente es necesario tener en cuenta un equilibrio entre el contraste y la resolución.

2.1.2 Microscopio de contraste de fase.

El microscopio de contraste de fases es el método más popular para estudiar muestras biológicas sin teñir que son transparentes. El contraste de fases se basa en el principio descubierto por Zernike que dice que la luz difractada por un objeto no absorbente sufre un cambio de fase. Así en un microscopio de contraste de fases, mediante el uso de un condensador especial y su anillo de fase complementario, la amplitud de la luz central no difractada que normalmente pasa a entre y alrededor del espécimen es reducida. Estos filtros introducen un cambio de fase adicional en la luz no difractada. La luz difractada o dispersa que pasa mayormente fuera del anillo de fase no sufre ningún cambio de fase. Por lo tanto, debido a la presencia de un plano de cambio de fase, se crea una diferencia de fase entra la luz difractada y no difractada, los rayos de luz con diferentes fases interfieren en el plano de la imagen formando máximos y mínimos bien definidos en

las intensidades de la imagen que resultan en un incremento del contraste de la imagen formada (Zernike, 1955).

Imágenes formadas por contraste de fases se caracterizan por halos y sombras, los cuales son imágenes de baja intensidad de contraste opuesto formadas debido a cierta cantidad de luz difractada que logra pasar a través del anillo de fases. El sistema de fases limita en cierto grado la capacidad de la apertura numérica de una lente, reduciendo así la resolución. El contraste de fases no funciona bien con especímenes robustos (como cultivos de células de varias capas) debido a que ocurren cambios de fase fuera del plano enfocado, estos cambios de fase fuera de foco producen distorsiones en la imagen. El contraste de fases funciona bien con luz monocromática (verde) (Rittcher, 2008).

2.1.3 Imagen óptica e imagen digital.

Una cámara es un dispositivo que convierte una imagen óptica en su superficie foto sensible a una secuencia de señales eléctricas. Esto se logra mediante el escaneo de la superficie fotosensible de la cámara la cual está compuesta por secciones, en cada escaneo la intensidad de la luz incidente es convertida en voltaje el cual es proporcional dentro de un rango a la intensidad de luz. Un ajuste adecuado es de vital importancia debido a que la información perdida a nivel de cámara no puede ser recuperada posteriormente. Los controles que modifican el ajuste de la cámara son la ganancia, linealidad de la ganancia gamma, intensidad de iluminación y nivel de negro (Wolf & Sluder, 1998).

A grandes rasgos una cámara debe estar ajustada de forma que todos los valores de intensidad sean representados en su señal de salida usando la amplitud total de la señal para incluir los valores máximos y mínimos de intensidad de la imagen.

2.2 Parámetros característicos de una imagen.

El propósito fundamental de una cámara electrónica es generar una señal eléctrica la cual codifique la imagen óptica proyectada en su superficie. Cuando la intensidad de luz es razonablemente alta, existe una gran variedad de cámaras con un desempeño muy bueno. Conforme la intensidad de la imagen disminuye se vuelve cada vez más difícil obtener imágenes de alta calidad, comprometiendo la resolución espacial, la resolución temporal y el nivel de ruido. Sin embargo, gracias a los avances en electrónica y a que cada día surgen cámaras más sofisticadas las limitaciones son menos severas. Idealmente una cámara registraría la llegada de cada fotón con una alta precisión espacial y temporal sin introducir ningún tipo de señal externa o ruido. Un detector ideal no existe por lo cual es importante tomar en cuenta varios parámetros que determinan la respuesta característica de una cámara (Wolf & Sluder, 1998).

2.2.1 Sensibilidad.

La sensibilidad de una cámara se refiere a la respuesta de la cámara dada cierta intensidad de luz. Una medida conveniente de esto es la eficiencia cuántica la cual se refiere a la fracción de fotones incidentes en la superficie del detector que generan una respuesta. Diferentes materiales y diseños pueden tener muy diferentes eficiencias cuánticas, desde porcentajes bajos hasta más de 80% de eficiencia en algunos dispositivos. Entre mayor sea la eficiencia cuántica la cámara es generalmente más sensible, aunque un mínimo nivel de ruido debe ser también considerado para determinar la sensibilidad.

2.2.2 Ruido.

Una imagen electrónica es complicada debido a la presencia de ruido, el cual degrada la calidad de la imagen. Si la respuesta de una cámara a cierto nivel de luz dado está por debajo del nivel de ruido, la imagen no estará resuelta. Así pues, la medida más importante de la respuesta de una cámara es por lo tanto la relación señal-ruido. La relación señal-ruido es definida como el nivel de señal generado por la respuesta de la cámara a la luz, dividido por el nivel de ruido promedio. Entre mayor sea la relación señal-ruido de una cámara, mejor será la calidad de la imagen.

Entre las fuentes de ruido más comunes se encuentran el ruido de lectura, que es ruido generado durante la transferencia de la imagen grabada en el detector hacia un componente electrónico en donde se almacena. El ruido de disparo, al utilizar una cámara electrónica para detectar una imagen se está haciendo esencialmente un experimento en contar el número de fotoelectrones almacenados en un pixel. La cuenta de electrones es un proceso aleatorio que sigue las estadísticas de Poisson, la incertidumbre en la medida resultante es llamada ruido

de disparo o ruido de fotón, el ruido de disparo es intrínseco a las estadísticas del fotón de la imagen y no puede ser reducido. Ruido de fondo, el ruido de fondo o ruido oscuro es generado por varios motivos como electrones generados térmicamente los cuales son detectados como fotoelectrones. Otros factores como la iluminación dispersa de Rayleigh y Raman, o rayos de luz que se filtran a través de la cámara pueden contribuir al ruido de fondo. Generalmente para aplicaciones de video el ruido oscuro no es tan importante como el ruido de lectura o el ruido de disparo. Por otro lado, para tiempos de adquisición de imágenes más largos como en obturación lenta, el ruido oscuro puede ser significativo.

2.2.3 Resolución espacial.

Las cámaras electrónicas forman imágenes muestreando una imagen óptica en píxeles discretos. La resolución espacial de un sistema de cámara es determinada por el espaciado entre píxeles y es independiente de la resolución óptica del microscopio, aunque limitada por este (Castleman, 1993). Con el fin de resolver características de una imagen de cierta frecuencia espacial, la cámara debe muestrear la imagen con al menos el doble de esa frecuencia. Este requerimiento puede ser entendido simplemente declarando que los detalles de la imagen serán resueltos únicamente cuando el espaciado entre píxeles sea menor a la escala sobre la cual las características de la imagen cambian. Al dividir el detector en píxeles la imagen se resuelve, y al incrementar el número de píxeles, niveles más altos de detalles llegan a ser aparentes, la resolución de una cámara puede aumentarse hasta igualar la resolución óptica del microscopio un aumento más allá de este punto no produciría ningún beneficio adicional.

2.2.4 Linealidad y uniformidad.

Generalmente las cámaras digitales están diseñadas para ser lineales, es decir, que el detector tiene una relación lineal entre la intensidad de la señal de entrada y la señal de salida, esta característica esencial para imágenes cuantitativas. La linealidad se puede ver comprometida en casos de sobreexposición del detector en donde se produce un fenómeno llamado blooming, también la respuesta del detector no a lo largo de su superficie no es siempre uniforme lo que causa distorsiones y sombreados. Otras fuentes de pérdida de linealidad en cámaras digitales son acoplamientos en la etapa de salida de señal como el control de brillo automático y la corrección de gama, las cuales alteran la señal de salida del detector original perdiendo por lo tanto cierta cantidad de información.

2.2.5 Tiempo de respuesta.

Las cámaras no siempre responden inmediatamente a un cambio en la intensidad de entrada, esto introduce un retraso y disminuye la resolución temporal. En las cámaras digitales el retraso no suele ser significativo, sin embargo, es importante asegurar que el retraso de la cámara no afecte las mediciones.

2.2.7 Rango dinámico.

Hay un número finito de niveles de intensidades o niveles de grises que pueden ser detectados por un dispositivo, lo cual puede limitar la detección de los detalles más finos. El rango de niveles es llamado rango dinámico. Un rango dinámico alto es necesario en aplicaciones donde se requiere grabar un amplio rango de niveles de intensidades. El rango dinámico puede ser especificado para una sola imagen (frame) y para secuencias de imágenes. El rango dinámico es determinado por dos factores el rango operacional total o profundidad del pixel, y el nivel de ruido promedio, y es calculado dividiendo el nivel de señal máximo entre el nivel de ruido promedio mínimo excluyendo el ruido de disparo.

2.2.8 Ganancia.

La ganancia se refiere a la amplificación de señal por la cámara para crear una señal eléctrica la cual pueda ser medida precisamente. Mientras la sensibilidad describe la respuesta de entrada del dispositivo, la ganancia determina la respuesta total de salida. La etapa de ganancia siempre aumenta el ruido en la señal, pero cierto nivel de ganancia es necesario para transmitir la señal entre componentes electrónicos.

2.2.9 Velocidad de fotogramas (frame rate).

Diferentes aplicaciones requieren adquisición de imágenes a diferentes velocidades o velocidad de fotogramas. La velocidad de fotogramas es determinada por la velocidad de lectura del pixel y el número de pixeles en la imagen. Para un determinado número de pixeles, la principal desventaja a velocidades de fotogramas altas es el incremento de ruido debido a velocidades de lectura mayores, una mayor velocidad de lectura también tiene menos luz por pixel por fotograma. La digitalización es normalmente la etapa que limita la velocidad de lectura. Para cierto nivel de ruido y profundidad de digitalización (número de bits), la velocidad máxima de digitalización es consistente entre fabricantes de cámaras. La velocidad de fotogramas es pues determinada por el número de pixeles. La mayoría de las cámaras tienen una velocidad de fotogramas fijo de 30 o 60 cuadros por segundo (fps).

Una cámara electrónica operada apropiadamente es esencialmente un arreglo de foto detectores independientes. Cuando la relación señal-ruido es razonablemente alta para la mayoría de los píxeles, el resultado serán imágenes de alta calidad. Varios diseños de cámaras emplean diferentes estrategias para maximizar la relación señal-ruido, usualmente comprometiendo la resolución espacial y la velocidad de fotogramas.

Actualmente hay muchas opciones disponibles en cámaras electrónicas y dispositivos con características de desempeño mejorados. La mayoría de las cámaras son fabricadas con diseños para aplicaciones específicas lo que resulta en un gran número de posibilidades para la captura de imágenes biológicas (Wolf & Sluder, 1998).

2.3 Adquisición de imagen.

No todas las cámaras electrónicas tienen el mismo formato de salida, y evaluar un sistema de adquisición de imágenes es una parte importante al elegir una cámara. El objetivo básico es obtener una imagen en formato digital y leerla en una computadora. Para muchas cámaras especializadas el fabricante incorpora un digitalizador dedicado como parte del sistema de la cámara. Este incorpora la elección de desempeño más deseable y debería ser fácil llevar la imagen digital a una computadora para realizar los procesos que sean necesarios. Una vez en la computadora las imágenes digitales deben ser almacenadas. Muchos formatos personalizados son utilizados por varios programas de adquisición de imágenes.

Hay numerosas formas de procesar una imagen una vez que esta almacenada en formato digital. En ocasiones procesamientos adicionales son necesarios únicamente para visualizar los datos. Finalmente, los datos más útiles no son siempre la intensidad de imágenes iniciales sino una combinación de imágenes. Mientras que el almacenamiento y procesamiento básico de imágenes es normalmente muy sencillo y directo, puede volverse una tarea muy complicada si los componentes de interface y de cámara no están correctamente emparejados (Shotton, 1993).

2.4 Introducción al procesamiento de imágenes.

2.4.1 Digitalización de imágenes.

En un sistema de video por microscopio, una cámara toma el lugar del ojo humano como detector de luz. Una cámara digital escanea la imagen y genera una salida en forma de niveles de voltaje que representa la imagen. Los datos de voltaje son entonces enviados a un procesador de imagen. El procesador de imagen convierte entonces este voltaje en valores numéricos que representan el nivel de brillo o intensidad de esta forma digitalizando dicha imagen. Una vez que la imagen ha sido convertida a forma digital, esta puede ser tratada por una computadora.

Los componentes que hacen posible que un procesador de imagen tome una imagen, la convierta en valores numéricos, manipule y almacene esos datos, para luego convertirlos de nuevo en una imagen para visualizarla son:

Un convertidor análogo-digital (A/D), el cual asigna un valor numérico a cada nivel de voltaje dentro de un rango dado. El conversor A/D muestrea la señal de video a una frecuencia establecida, que resulta en el número deseado de muestras (píxeles) por renglón de la imagen. Una vez que la señal de video se ha convertido a números, los datos pasan a otro dispositivo llamado unidad aritmética lógica (ALU).

La unidad aritmética lógica (ALU). La velocidad a la cual una cámara de video genera imágenes es llamada tasa de video. Usualmente en países que adoptaron frecuencias de 60 Hz en sus líneas de energía, la tasa de video es de 30 fps. Para una imagen típica con resolución de 640X480 píxeles la imagen contiene un total de 307, 200 píxeles esto se traduce a 9,216,000 píxeles por segundo pasando a través del procesador de imagen. Procesar 9 millones de datos por segundo requiere una capacidad de procesamiento considerable incluso para la tecnología actual además de esta puede haber otros factores que causen una disminución en la tasa de video. La unidad aritmética lógica puede llevar a cabo otras operaciones sobre y entre imágenes. Cuando ocurre una disminución en la velocidad de procesamiento por cualquier razón, la ALU actúa como un componente de desahogo, esta combina un píxel de cada imagen fuente para computar el resultado. Las computaciones típicas de la ALU incluyen: adición, sustracción, lógica OR, AND y XOR. La mayoría de las ALU también tienen la capacidad de dividir la imagen por una constante.

Búfer de fotogramas, este componente es una memoria dentro del procesador de imágenes usado para almacenar una o varias imágenes que han sido adquiridas, procesadas o van a ser mostradas. Algunos sistemas de procesamiento de imágenes no cuentan con búfer de fotogramas, en su lugar utilizan la memoria de la computadora que los controla. Sin embargo, los procesadores de alta velocidad tienen búferes integrados en el hardware del mismo

procesador debido que el acceso a la memoria de una computadora puede ser mucho más lento que a un buffer de fotogramas integrado.

Look-Up tables (LUT) o tablas de búsqueda. Después del procesamiento y almacenamiento una imagen usualmente pasa a través de una tabla de búsqueda. La LUT es simplemente una tabla de valores, uno por nivel de gris en la imagen, que convierte un valor de la imagen a un valor correspondiente en la tabla. En una imagen de 640X480 en la que se desea realizar una operación compleja a cada pixel que resultaría en más de un cuarto de millón de operaciones, en su lugar se realiza la operación a cada nivel de gris posible de un pixel, 256 operaciones en un pixel de 8 bits, los resultados de la operación y su nivel de gris correspondiente se almacenan en la LUT y están disponibles para ser referenciados por el procesador (operación rápida) en lugar de hacer la operación en cada pixel (operación lenta). La LUT es especialmente útil para modificar rápidamente el contraste de una imagen mientras se visualiza en un monitor.

Convertor digital-análogo. Los valores convertidos que pasan a través de la LUT usualmente entran a un convertidor análogo digital (D/A). este dispositivo recibe valores numéricos y los convierte en una señal de video análoga. La señal puede ser usada por un monitor de video para mostrar la imagen o por una grabadora VCR para grabar la imagen (Wolf & Sluder, 1998), (Gonzales, 1987).

2.4.2 Pixeles.

Una vez que la imagen ha sido convertida a forma digital es fundamentalmente descompuesta en una malla finita de puntos llamados pixeles.

El pixel y el voxel (en el caso de una imagen tridimensional) son la unidad más pequeña que componen una imagen. Un pixel representa un objeto muestreado en una locación específica en el espacio. Cada pixel en una imagen tiene unas coordenadas únicas X y Y (en el caso del voxel tiene X , Y y Z). Por razones históricas, los pixeles en una imagen son medidos usualmente desde la esquina superior izquierda de la imagen, comenzando en la posición $(0, 0)$.

En una imagen digital cada pixel tiene su propia locación y un valor, su valor gris. Típicamente el valor de un pixel es un número de 8 o 16 bits que la computadora puede usar para representar 256 o 65,536 distintos niveles de brillo. Los pixeles en imágenes digitales a color son representados por tres valores, valor de rojo, verde y azul. Los pixeles están distribuidos en la imagen de forma uniforme, cada uno está muestreado en una locación específica. Debido a que un pixel tiene un tamaño específico el valor del pixel se puede pensar como el promedio de los valores dentro del área cubierta por ese pixel (Wolf & Sluder, 1998).

2.5 Métodos computacionales.

Muchos de los métodos computacionales utilizan algoritmos que consciente o inconscientemente copian muchos aspectos de la visión humana. Por conveniencia es usual separar los procesos visuales humanos en visión temprana y visión tardía. La visión temprana corresponde al procesamiento de la información en la retina y las redes neuronales cerca de ella, antes de que datos de alto nivel sean transmitidos al cerebro. La visión tardía se refiere al manejo de información en el cerebro donde los hechos “aprendidos” acerca del mundo entran en juego. Esta distinción también tiene una correspondencia aproximada a la distinción entre extraer información de bajo nivel de una imagen como la presencia, locación, y orientación de bordes, límites y quizás objetos (visión temprana), comparada al uso de información para “entender” una escena y la relación entre los objetos presentes en ella (visión tardía).

La mayoría de los métodos computacionales para el análisis y medición de imágenes utiliza algoritmos relacionados a la visión temprana. Hay por otro lado, un gran campo de estudio en robótica e inteligencia artificial que busca entender y describir las escenas. Afortunadamente la gran mayoría de las imágenes que son importantes para análisis y mediciones son obtenidas en situaciones muy controladas donde mucha información acerca del espécimen y sus condiciones es conocida.

La reproducibilidad de un análisis de imagen computarizado y métodos de medición pueden ser mucho mejores que las de un observador humano, porque los algoritmos ignoran mucho del contenido de la imagen y los sensores y discriminadores responden únicamente a la imagen misma. Los humanos son influenciados por muchos otros factores. Mediciones computarizadas presentan variaciones que están más o menos directamente relacionadas a simples patrones estadísticos de fluctuación, de este modo, los errores pueden ser predichos y muchas veces controlados. Por otro lado, reproducibilidad no implica exactitud, por lo que es necesario un método de calibración contra alguna fuente externa, como estándares de especímenes conocidos o simulación matemática, para producir resultados exactos.

Es fácil distraerse del propósito real del análisis de imágenes por el hardware y software usados para su implementación. Para un problema en particular existe una gran variedad de aproximaciones, variando entre ellas por los métodos de procesamiento, la utilización de hardware o software para realizar los cálculos y los lenguajes de programación. Todas estas diferencias son de poca importancia comparadas con la elección del algoritmo apropiado para llevar a cabo el método deseado. El resultado obtenido al implementar el método por cualquiera de las aproximaciones disponibles debe ser el mismo. Lo anterior aplica para la elección

del método de procesamiento, la realización de los cálculos y el lenguaje de programación. Las diferencias entre hardware y lenguajes a menudo son pequeñas y hasta triviales, y cualquiera de ellos es capaz de expresar los algoritmos necesarios (Russ, 2012).

2.6 Fuentes de imágenes.

Un análisis de imágenes computarizado puede ser aplicado a casi cualquier tipo de imagen, pero las aplicaciones con mejores resultados son aquellas en las que las condiciones de iluminación y punto de vista son bien conocidos y controlados, y existe algún conocimiento previo acerca de los objetos que van a ser encontrados. En particular las imágenes son usualmente superficies, aunque pueden no necesariamente estar restringidas a superficies planas, o de objetos dentro de un volumen transparente de espesor conocida. Esta situación incluye la observación de partículas dentro de un sustrato.

La escala de la imagen es de poca importancia. Las mismas técnicas básicas aplican para imágenes obtenidas con microscopios o telescopios.

Cuando una imagen es almacenada en una computadora, el grado de cuantización llega a ser aparente. Cada locación de memoria en la computadora (usualmente un byte, 8 bits) almacena la información de brillo de un punto o pixel de la imagen. El número de bytes corresponde entonces al número de pixeles y define la resolución de la imagen. Usar muchos bytes para almacenar una imagen incrementa el costo del análisis de la imagen. El costo no es únicamente económico (Las memorias cuestan dinero), más importante aún, todos los procesos que operan en la imagen toman tiempo, que es aproximadamente proporcional al número de pixeles almacenados (Russ, 2012).

2.7 Métodos de procesamiento.

Por definición el procesamiento de imágenes incluye aquellos métodos que comienzan con una imagen (un arreglo de píxeles, cada uno con un valor de gris o color) y terminan con una imagen. Usualmente el tamaño de la imagen resultante será similar al original. El brillo de los píxeles resultantes en la mayoría de los casos habrá sido modificado usando reglas que toman en cuenta el valor original del píxel y sus vecinos, o quizás su posición o relación espacial con respecto a muchos otros píxeles. En ocasiones muchas imágenes son combinadas para producir una nueva imagen. Los varios tipos de reglas incluyen algunas basadas en principios teóricos, o algunas técnicas empíricas ad-hoc. Con ellas se pretende producir una imagen procesada modificada que enfatice algunos aspectos de la imagen original, a la expensa de otros. Esto es en efecto, una técnica de filtrado que selecciona cierto tipo de información. Seguido las características deseadas son objetos, bordes límites o algunas otras estructuras definidas. Dependiendo de aquello que se considere de interés es posible construir una operación de procesamiento adecuada para extraerlo de la información que no es de interés (fondo) de la imagen, para ser medido posteriormente. En ocasiones es necesario realizar operaciones varias veces en la misma imagen original, usando diferentes operaciones de procesado, para aislar diferentes tipos de información, como estructuras de diferentes tallas u orientaciones (Russ, 2012).

2.7.1 Reducción de ruido.

El mayor uso del procesamiento de imágenes en conjunción con la medición de imágenes es la corrección de defectos específicos en la imagen adquirida. Los problemas más comunes están asociados con el ruido en la imagen original, o una iluminación no uniforme, u otras variaciones sistemáticas en el brillo.

El ruido puede ser aleatorio (estocástico) o periódico. El ruido aleatorio es mejor conocido como nieve o estática en la imagen de un televisor. El valor de brillo de los píxeles es modificado aleatoriamente difiriendo de su valor apropiado, o puede estar abandonado. Píxeles abandonados o inactivos son usualmente mostrados con el valor máximo o mínimo disponible del píxel.

Dentro de una imagen cada píxel tiene un valor de brillo. Este tiene también 8 vecinos en los casos más comunes. En cualquier caso, la información en la imagen es usualmente de un tamaño significativamente mayor al tamaño de los píxeles individuales, por lo tanto, se espera que píxeles adyacentes usualmente sean parte de la misma superficie o el mismo objeto. En estos casos se puede llevar a cabo un promediado utilizando los píxeles vecinos. El valor original es reemplazado por el resultado obtenido, la operación se lleva a cabo simultáneamente en cada píxel individual.

Normalmente no se utiliza un promediado simple, en su lugar se añade una ponderación a cada pixel vecino dependiendo su posición respecto al pixel central para producir un suavizado gaussiano. Otros factores de ponderación pueden ser ideados, permitiendo darles un valor de mayor importancia al pixel central y valores de menos importancia a pixeles vecinos en proporción a su distancia respecto al pixel central.

El número de pixeles vecinos que se utiliza en la operación puede es determinado por el tamaño del kernel. El kernel más pequeño es de 3x3 que contiene al pixel central y a los ocho pixeles adyacentes, otros tamaños de kernel pueden ser 5x5 o 7x7 con ponderaciones personalizadas.

Es un hecho que suavizar una imagen con estas funciones ponderadas no es lo óptimo ya que degrada la nitidez de los bordes o discontinuidades de la imagen. En adición a promediar los niveles de ruido de la imagen, el suavizado promedia también valores que son realmente diferentes en la imagen. Esto significa que los límites de superficies y objetos serán atenuados y el cambio de brillo entre ellos será reducido en magnitud. Estas desventajas son exactamente contradictorias al propósito original el cual es encontrar y medir dichas características por lo que se debe estar consiente al utilizar esta aproximación.

El mejor método para remover ruido aleatorio de una imagen es aplicar un filtro de mediana. Este funciona de forma similar al filtro gaussiano, trabaja con cada pixel y sus vecinos usualmente en kernel de 3x3. Los valores de los pixeles se ordenan por valor, y el valor de la mediana (el quinto en un kernel de 3x3) es usado para reemplazar el valor original del pixel central. El proceso se lleva a cabo simultáneamente en todos los pixeles en la imagen de esta forma los valores utilizados sean los originales. Mientras que es una herramienta de procesamiento ideal en términos de sus resultados, el filtro de la mediana presenta una dificultad, la operación en si consume mucho tiempo por lo que es muy demandante en términos de capacidad de procesamiento. El procedimiento de ordenar los valores de los pixeles no es ameno al uso de arreglos de procesadores y la operación puede iterarse un numero impredecible de veces antes de que los valores permanezcan constantes. El tiempo necesario para aplicar un filtro de mediana puede variar de segundos hasta más de un minuto.

2.7.2 Segmentación.

Para reconocer y medir objetos dentro de una imagen es necesario distinguirlos del entorno. Esto es separar un objeto del fondo, o separar una característica de otra contigua. Una característica en una imagen se define como una región de pixeles continuos que comparten alguna propiedad.

El ojo humano tiene un enorme número de receptores de luz individuales en la retina. Si cada uno de estos simplemente transmitiera la información de la luz al cerebro sería necesario un gran número de conexiones y simplemente quedaría el problema de interpretar esa enorme cantidad de información. En su lugar la naturaleza ha desarrollado un esquema de proceso más distribuido, en el cual solo una pequeña parte de información de alto nivel debe ser transmitida. De hecho, dentro de la retina hay un cierto número de células adicionales dedicadas a la extracción de esta información de alto nivel. De esta forma se crean simples circuitos neuronales para detectar por ejemplo un borde con una orientación en particular en la cual la respuesta de varias células es combinada algunas positivas y algunas negativamente. Cuando una región uniforme o un borde con una orientación incorrecta es visto, no habrá una salida de este detector. Únicamente cuando se visualice un borde con la orientación adecuada el detector producirá la señal deseada. De esta forma otros detectores son usados para detectar bordes en diferentes direcciones. Otros patrones de detectores detectan esquinas, u otras características.

El principio de este método es la inhibición. Conexiones positivas o negativas en las células “lógicas” permiten que “píxeles” vecinos sean considerados. En el caso de las imágenes digitales se utiliza un kernel con ponderaciones que representen de forma numérica la característica que se está buscando, al utilizar ponderaciones negativas se logra inhibir las características de poca importancia y únicamente cuando se detecte la característica buscada (un borde, esquina, etc.) el operador producirá una respuesta.

Para detectar ciertas características como bordes en diferentes direcciones es necesario un conjunto o familia de kernels que cubran las posibles formas en las que dicha característica puede aparecer en la imagen.

Existen otros detectores de bordes, algunos están específicamente diseñados para copiar la forma en la que trabaja el sistema visual humano, basados en estudios fisiológicos. Otros grupos de técnicas de detección de bordes utilizan patrones de vecindad en formas que no son fácilmente configuradas como kernels. Estos incluyen máximos y mínimos, y métodos de entropía local.

Con la mayoría de las técnicas derivativas de detección de bordes, uno de los principales problemas que surgen es la continuidad de los bordes (o mejor dicho la falta de continuidad). Al utilizar cualquiera de estas técnicas no es raro encontrar segmentos que no conectan para formar una línea continua alrededor de la característica u objeto. Esto es debido a que los métodos utilizan únicamente comparaciones locales entre píxeles, y no hay garantía de que los puntos se unirán formando líneas conectadas. Existen aproximaciones para conectar segmentos de

línea incompletos en perímetros continuos. Estos incluyen métodos que pertenecen al entendimiento de la escena o visión tardía (Russ, 2012).

2.7.3 Umbrales y discriminación.

Discriminación mediante el uso umbrales, se trata de la discriminación por niveles de brillo para seleccionar pixeles que pertenecen a las características de interés. Este es un método ampliamente usado de convertir una imagen en escala de grises a una imagen binaria (blanco y negro). La discriminación mediante umbral es mucho más eficiente que cualquier método de seguimiento de bordes o crecimiento de regiones, debido a que trabaja en la imagen completa al mismo tiempo. También la imagen binaria resultante es una representación basada en pixeles de las características de interés, y es más sencilla para la mayoría de las operaciones de medición.

En la forma más simple de discriminación de brillo, dos niveles de umbral son establecidos manual o automáticamente. Todos los pixeles cuyo valor de brillo se encuentre entre estos umbrales es considerado como parte de la característica de interés, y todos los demás son considerados como fondo. En otras palabras, figura y fondo son separados.

La configuración manual del umbral de discriminación se realiza interactivamente, se modifica el valor del umbral mientras se observa la imagen resultante. Esto se hace con la premisa de que el usuario conoce cual característica es de interés.

Es muy importante entender que operaciones de discriminación con umbrales no se deben aplicar en la imagen original en escala de grises. Las operaciones de suavizado como método de reducción de ruido, producen una imagen en escala de grises derivada de la original, la cual puede ser discriminada para producir un mejor resultado del que produciría discriminar la imagen original. En imágenes suavizadas el filtro de suavizado modifica los valores de la escala de grises solo ligeramente y la uniformidad de regiones es mejor.

Otros operadores también pueden ser aplicados a la imagen antes de la discriminación para producir efectos particulares o extraer tipos particulares de información. Esto permite selecciones características particulares de la imagen.

Sin importar cual operador u operadores sean utilizados, el propósito clave de esta operación es reemplazar la imagen en escala de grises original con una nueva imagen en la cual el brillo represente algún parámetro diferente en la imagen que la característica del brillo (Russ, 2012).

2.7.4 Imagen binaria.

La representación binaria basada en píxeles que resulta de la discriminación de la imagen en escala de grises, puede no delinear perfectamente todas las características presentes. Esto puede deberse al ruido, o imperfecciones en la imagen original, o métodos de procesamiento inadecuados, o algo inherente a la imagen misma como objetos traslapados.

Normalmente este tipo de situaciones son manejadas mediante la manipulación de la imagen binaria. La edición de la imagen comienza con una o varias imágenes binarias y produce una imagen resultante usualmente binaria. La mayoría de las operaciones de edición se realizan directamente sobre las imágenes mismas donde el arreglo espacial de los píxeles es directamente analizado. Una vez que una secuencia de operaciones ha sido establecida, esta puede ser aplicada automáticamente a subsecuentes imágenes.

Una clase principal de operaciones con imágenes binarias es la combinación lógica de dos o más imágenes para producir una única como resultado. Los operadores lógicos booleanos que están disponibles son AND, OR, XOR, y NOT. Estos pueden ser usados con cualquier combinación para producir la declaración lógica deseada. Combinaciones lógicas o booleanas pueden ser usadas para propósitos determinados como restringir el campo de visión a una apertura o perímetro reducido.

Otra clase importante de operaciones con imágenes binarias son procedimientos que implican relaciones vecinales. Muchas de estas operaciones forman la base de un campo conocido como morfología matemática. Un ejemplo básico es conocido como erosión y dilatación. Estos sirven para un gran número de funciones útiles incluyendo el suavizado del perímetro de las características, unificación de características rotas o discontinuas y la separación de aquellas que se tocan (Russ, 2012).

2.7.5 Erosión y dilatación.

En su forma más simple la erosión consiste en examinar cada píxel binario y cambiar su valor de activo a inactivo si cualquiera de los píxeles adyacentes está inactivo. Como en otras operaciones que toman en cuenta píxeles vecinos, es necesario especificar cuáles vecinos han de ser considerados.

La operación de erosión reduce las características alrededor de toda su periferia. Es una operación global que se lleva a cabo en la imagen completa al

mismo tiempo. Características con protuberancias angostas las verán removidas por erosión, y características conectadas por hebras delgadas serán separadas.

La operación de dilatación es inversa a la erosión. Esta causa que las características crezcan o se dilaten en tamaño, lo cual rellenara pequeñas roturas en las características, vacíos internos o muescas a lo largo de la superficie de la característica. Matemáticamente esto se hace cambiando cada pixel inactivo a activo si este tiene al menos un pixel vecino activo.

En la práctica las operaciones de erosión y dilatación pueden ser más flexibles. En lugar de remover un pixel si cualquiera de los pixeles vecinos está inactivo, se puede contar el número de vecinos inactivos y únicamente remover el pixel analizado, si más de un numero preestablecido de vecinos están inactivos. Diferentes valores producen diferentes efectos. Típicamente valores de 1 a 4 son utilizados debido a que producen una erosión general y un suavizado del contorno de la característica.

Al utilizar las operaciones de erosión y dilatación en secuencia se recobra gran parte del tamaño original de la característica mientras se produce una forma suavizada, esta secuencia de erosión seguida de una dilatación se conoce como opening. La erosión inicial remueve pequeñas características que pueden representar ruido, o ligeras protuberancias en el contorno de la característica. La subsecuente dilatación no restaura las pequeñas protuberancias, las cuales desaparecen de forma permanente, además rellena pequeñas muescas en los contornos. El resultado final es que el tamaño de la característica es restaurado casi a su tamaño original, mientras la forma es modificada a una forma más redondeada y suave.

Cuando la secuencia es invertida, esto es cuanto la operación de dilatación es aplicada en primer lugar y luego se aplica la erosión, esto es referido como closing. En algunos respectos esto es lo opuesto a el opening debido a que pequeñas características no desaparecen, en su lugar pequeños huecos en la característica son rellenados, y roturas o espacios entre características son unidos. Debido a que la operación de dilatación tiende a aumentar el tamaño de la característica por lo que aumenta también el número de pixeles que son susceptibles a ser removidos en el paso subsecuente, el contorno resultante es también más suave y así la diferencia en el área de la característica después del ciclo de un ciclo de closing es mucho más pequeña que para un opening.

Las operaciones de closing y opening pueden ser aplicadas más de una iteración, y los coeficientes pueden ser los mismos para cada operación o pueden ser cambiados en cada iteración. Ambas operaciones opening y closing, al igual que las operaciones básicas de erosión y dilatación son usadas para limpiar imágenes

binarias removiendo pequeñas características o irregularidades las cuales se deben a ruido u otras imperfecciones del objeto o de la imagen.

Resulta que existe una forma muy conveniente de aplicar la lógica de erosión y dilatación que no envuelve contar el número de pixeles vecinos el rededor de cada punto, en su lugar se usa el estado de activo o inactivo de los vecinos para construir un numero binario. Con ocho pixeles vecinos esto produce un numero de ocho bits el cual tiene un valor numérico de 0 (00000000, o todos inactivos) a 255 (11111111, o todos activos) este número sirve como índice a una tabla llamada tabla de destino la cual indica si el pixel central se debe activar o desactivar. El uso de esta tabla hace posible responder a cualquier patrón posible de vecinos y hace posible otros tipos de operaciones (Russ, 2012).

2.8 Medición de imágenes.

Los tipos de medición que se pueden realizar en una imagen se clasifican en dos tipos, el primer tipo es determinar parámetros globales o parámetros de la escena, y el segundo se relaciona con la medición de valores para cada característica individual que está presente.

2.8.1 Mediciones globales.

El primero de los parámetros globales es la fracción del área la cual está cubierta por los píxeles seleccionados. Vale notar que esto solo es significativo cuando se trabaja con una representación basada en píxeles (imagen binaria) de las características seleccionadas, ya que la representación de los límites debe ser reconvertida a píxeles para obtener el valor. La fracción del área cubierta es determinada simplemente contando los píxeles. Es importante notar que la fracción no debe estar basada en el número total de píxeles en la imagen, sino que debe haber otra área de referencia que pueda ser usada.

2.8.2 Medición de características.

Aunque las mediciones globales son formas extremadamente rápidas y eficientes para obtener datos útiles acerca de la imagen, hay más información disponible al medir las características individuales presentes. Algunas veces esto es mucha más información de la que necesitamos, y se puede reducir simplemente a un valor promedio, pero en otros casos la distribución de valores puede en sí mismo ser caracterizado, o comparado a otro conjunto de datos de una muestra diferente.

Por conveniencia los parámetros se dividen en cuatro grupos. El primer grupo cubre las mediciones de tamaño, el segundo trata con varios aspectos de forma, el tercero trata de describir posición de características, y el cuarto incluye parámetros que utilizan la información de los valores de la escala de grises. Dentro de cada una de estas categorías hay una variedad de parámetros individuales que pueden ser medidos, o calculados de otros que son medidos directamente.

La medida más obvia de tamaño es el área de una característica, esto es simplemente el número de píxeles interiores en el caso de una imagen binaria, o el área dentro de los límites si se utiliza una imagen con contornos. El área es posteriormente convertida a las unidades de medida reales apropiadas. Cuando una imagen con bordes es utilizada en lugar de una imagen binaria, el cálculo del área de la característica es muy directo. Las coordenadas de los puntos en los bordes

son las esquinas de un polígono, y son usualmente expresadas en coordenadas cartesianas (X, Y). entonces el área total de la característica es:

$$Area = \frac{1}{2} \sum_{k=0}^{n-1} (X_k * Y_{k+1} - X_{k+1} * Y_k) \quad (1)$$

Una medida linear del tamaño es muchas veces más útil que el área total de una característica. Y es común convertir el área medida ya sea a su diámetro circular equivalente, o diámetro esférico equivalente, estos cálculos son simplemente como:

$$D_{Circ} = \sqrt{\frac{4}{\pi} * Area} \quad (2)$$

$$D_{Sph} = \sqrt[3]{\frac{6}{\pi} * Area} \quad (3)$$

Ambas de estas fórmulas asumen la forma de la característica. La primera asume que un círculo es una descripción razonable de la característica. La segunda no solo asume que la característica es razonablemente descrita por una forma esférica, sino que además supone que esta es vista en proyección.

Parámetros como el diámetro circular o diámetro esférico son normalmente descritos como parámetros derivados, mientras que el área misma es una medida o parámetro fundamental. Esto refleja el hecho de que el área es directamente obtenida de la imagen misma, mientras que el diámetro equivalente es calculado después. Es posible realizar los cálculos de las variables derivadas en el momento que los parámetros son obtenidos de la imagen, pero también es posible hacerlo después mediante archivos o datos fundamentales almacenados. El primer método tiene la ventaja de que los datos son presentados al usuario casi inmediatamente (ya que el tiempo para calcular los datos es casi imperceptible). El segundo método cuenta con la ventaja de que parámetros derivados específicos pueden ser construidos y evaluados con calma después de que los datos medidos estén completos y el usuario puede reflexionar sobre la interpretación. Un sistema de análisis puede contener cualquiera de estos métodos o ambos (Russ, 2012).

2.8.3 Longitud y amplitud.

La longitud de una característica en una imagen binaria puede ser obtenida usando el método de acoplar un polígono convexo a la característica para obtener un área llamada área de cuerda tensa. Esto se logra mediante la medición de la longitud de la característica usando un conjunto de planos de coordenadas con rotación. Después de que el proceso es completado, el ángulo de rotación en el que

los valores máximos y mínimos produzcan la mayor diferencia. Esta diferencia es considerada como la longitud de la característica.

Si la característica tiene una dimensión máxima que no está alineada a ninguno de los ángulos de rotación que fueron utilizados, entonces el valor en la medición de longitud subestimara el valor real de longitud. Sin embargo, este error es pequeño, en el peor de los casos. Incluso utilizando solo 8 rotaciones (que corresponden a saltos de 22.5 grados) el error máximo posible es del 2% de la longitud medida. Si se utilizan 16 rotaciones el máximo error posible se reduce a 0.995%. Este rango de error en la medición típicamente es más pequeño que el tamaño de un pixel en la mayoría de los casos.

El valor de la longitud es referido algunas veces como máximo diámetro de Feret. Por analogía directa existe también un diámetro mínimo de Feret el cual corresponde a la amplitud. Esta amplitud de la característica puede ser encontrada usando el mismo método usado en la medición de la longitud, encontrando la menor diferencia entre las coordenadas máximas y mínimas en cualquier ángulo de rotación.

Sin embargo, mientras que la longitud es bastante insensible al número de ángulos de rotación utilizadas, este no es el caso para la amplitud. Si la característica es muy larga y delgada, cualquier desviación en la orientación del ángulo de rotación produce un error considerable. No hay un límite de que tan grande puede ser el error ya que es el producto del seno por la longitud de la característica, y no existe un número razonable de ángulos de rotación que aseguren una medida exacta. Más aún si la característica tiene una forma doblada o cóncava la medición de amplitud será equivocada.

Para obtener un estimado de la longitud de una característica curva como fibras o estructuras dobladas, está disponible el diámetro, para fibras muy largas y angostas la longitud sería una fracción del perímetro. Al aumentar el ancho de la característica es todavía posible estimar dimensiones de longitud útiles utilizando el perímetro, asumiendo que la característica tiene una forma de franja con un ancho constante y extremos circulares (Russ, 2012).

2.8.4 Posición.

La medición de la posición de una característica puede ser dentro de la imagen misma o con respecto a un sistema de coordenadas global, o con respecto a otros objetos. La posición es normalmente dada en las mismas unidades que la longitud, aunque no es obligatorio (ej. Latitud o longitud).

Una forma común para obtener la posición de una característica es mediante las coordenadas X y Y del centroide del objeto. Este es el punto en el cual el momento del objeto es minimizado, y representa el punto por el cual los ejes del objeto pasan. El centroide del objeto es calculado al integrar el momento de la característica sobre cualquier par de ejes X y Y, y dividiendo por el área.

$$C. G_x = \frac{\sum(X_i + X_{i-1})^2(Y_i - Y_{i-1})}{Area} \quad (4)$$

Las coordenadas resultantes son la locación del centroide. En formas convexas el centroide estará dentro de los límites de la característica, aunque en formas cóncavas o que contengan huecos internos ese puede estar fuera. La locación del centroide es un buen ejemplo de un parámetro que requiere trabajar con una imagen binaria. Tratar de calcular el centroide de un objeto en una imagen representada por bordes produce un resultado sesgado.

Uno de los usos para las mediciones de varias posiciones es obtener información de la posición relativa entre varios objetos dentro de una imagen, o de varias imágenes en un campo de visión. Esto puede llevar a descripciones de adyacencia o asociación.

Dentro de una imagen es posible determinar la distancia y si es necesario la dirección del vecino más cercano. En algunos casos el segundo, tercero, etc. Vecinos son considerados. Una de las formas de medir la distancia entre objetos es mediante la dilatación secuencial de la imagen mientras se cuenta el número de ciclos hasta que los dos objetos se tocan. Este método no es perfectamente isotrópico y es muy sensible a la forma original del objeto.

Un método alternativo consiste en la búsqueda exhaustiva de cada una de las posiciones de los centroides de todas las características. Esto únicamente se puede hacer después de que se ha encontrado cada una de las posiciones de las características y se han almacenado en archivos de datos. Hay varios trucos para acelerar el proceso de búsqueda como ordenar los datos por las dimensiones de las coordenadas X y Y. usualmente la búsqueda de datos entre varios miles de características es un proceso que lleva una cantidad de tiempo significativa (Russ, 2012).

2.9 Seguimiento.

Seguimiento o tracking basado en imágenes se refiere a la detección de objetos relevantes en series de imágenes en el tiempo, así como la determinación de sus posiciones espaciales en el tiempo. El seguimiento es aplicado para modelar y entender las dinámicas moleculares biológicas, también para entender la migración de las células. En un sentido amplio el seguimiento puede pensarse como un problema de segmentación dependiente del tiempo, y como en el caso de segmentación de imágenes estáticas, existe una variedad de métodos de aproximación.

El algoritmo de seguimiento más simple consiste en la identificación de potenciales características de interés en una imagen por medio de operaciones de discriminación por umbrales (posiblemente seguido por operaciones morfológicas para remover huecos e imperfecciones), y llevar a cabo una comparación para encontrar al vecino más cercano. Mientras esta técnica puede ofrecer muy buenos resultados para el seguimiento de objetos con formas uniformes, su efectividad disminuye cuando los objetos cambian su forma o tamaño con el progreso del tiempo.

La asociación de datos es un paso crucial cuando se hace el seguimiento de varios objetos desordenados. De hecho, debido a que en la etapa de detección se puede perder o detectar incorrectamente ciertos objetos, los métodos carentes de asociación crearan asociaciones inconsistentes llevando a trayectorias desconectadas en imágenes subsecuentes. Uno de los métodos desarrollados para evitar este tipo de problemas es el método del vecino más cercano (NN). El NN es un método muy sensible al ruido por lo que incluso niveles bajos de ruido provocaran errores en las trayectorias resultantes.

El seguimiento de múltiples objetos está directamente relacionado a la densidad de objetos, la calidad y precisión de su detección y a la inicialización del seguimiento. Vale mencionar que este campo metodológico permanece abierto y que muchos estudios aun necesitan ser emprendidos para resolver problemas persistentes en los métodos de seguimiento de partículas biológicas. Algunos estudios ya han integrado conceptos de super-resolución para mejorar la eficiencia de detección, en casos de partículas con una alta interacción, y facilitar el proceso de seguimiento. Otros estudios sugieren dividir el problema de seguimiento de múltiples partículas en varias subtarefas secuenciales (Rittcher, 2008).

Capítulo III

3 Metodología.

3.1 Características de la imagen/video a analizar.

Las cámaras digitales actuales ofrecen una calidad de imagen muy buena y existen muchos modelos comerciales específicamente diseñados para ser adaptados y grabar imágenes microscópicas a diferentes resoluciones, en el caso de imágenes de esperma usualmente se utilizan resoluciones cuadradas de 900 X 900 o mayores, aunque no es obligatorio si es recomendado. La frecuencia de captura debe ser suficientemente rápida como para no perder demasiada información debido al movimiento de las células entre cuadros, imágenes capturadas con una exposición menor a 0.02 segundos a una frecuencia mayor o igual a 60 cuadros por segundo es suficiente para mantener un seguimiento adecuado, sin una pérdida significativa. Es necesario remarcar que entre mayor sea la resolución de la imagen y la frecuencia de captura más precisos serán los datos obtenidos pero la exigencia de procesamiento será aumentada de forma exagerada por lo que se recomienda mantenerse en un rango cercano a los valores descritos anteriormente, cumpliendo con dicho mínimo, la resolución de las imágenes analizadas en este proyecto es de 1920x1080 y la frecuencia de captura es a 60 cuadros por segundo. Los anteriores son los parámetros más importantes que se deben cuidar en la imagen obtenida, otros menos importantes y normalmente presentes en la mayoría de las cámaras digitales son el formato de salida, que debe ser compatible con las librerías de programación utilizadas mp4, avi, etc. Un amplio rango dinámico, bajo nivel de ruido, entre otros.

3.2 Segmentación y detección de partículas.

El video a analizar está compuesto de una secuencia de fotografías que se presentan a una frecuencia determinada. A una frecuencia de 60Hz se obtienen 60 fotografías por cada segundo grabado. El primer paso para analizar estas imágenes es leer las imágenes independientes del video de forma ordenada, este orden representa la coordenada temporal. A cada una de estas imágenes se le aplicara un proceso de modificaciones para aislar los objetos de interés y asignarles una posición en el plano de observación que serán las coordenadas espaciales. Para dichas modificaciones es necesario tratar las imágenes como matrices en donde cada posición de la matriz representa un pixel de la imagen.

Filtro de suavizado. Se utiliza un filtro para suavizar la imagen eliminando el ruido generado por la cámara e incluso algunas partículas pequeñas que puedan estar presentes en la imagen, el suavizado se lleva a cabo mediante la convolucion de la imagen con un filtro kernel pasa-baja. En la librería de Opencv hay varios filtros de suavizado disponibles siendo el más apto el filtro gaussiano (kernel gaussiano). Para el filtro gaussiano se requieren dos parámetros. El tamaño del kernel de barrido que debe ser un número impar positivo y la desviación estándar que definirá los valores del kernel (Joshi, 2019), los parámetros óptimos utilizados en el programa son un kernel de 15 pixeles con una desviación estándar de 5.

Binarización. Con la imagen suavizada el siguiente paso es aislar los objetos de interés del fondo de la imagen para lo cual se hace uso de un filtro de binarización, el cual por medio de un umbral de grises asigna un valor a cada pixel de la imagen, si el valor del pixel original es menor al parámetro de binarización se le asigna un valor de cero, en cambio si es mayor al parámetro se le asigna el valor máximo, separando así las células blancas del fondo oscuro (Joshi, 2019).

Erosión y dilatación. Filtros de erosión y dilatación ayudan a eliminar ruido y rellenar pequeños huecos en la imagen que puedan haber persistido en los filtros anteriores y ayuda a obtener una forma más consistente en cada una de las células aisladas, se utiliza principalmente en imágenes binarias. La forma en que se aplica es mediante una convolución 2D de la imagen con un kernel del tamaño deseado que asignara a cada pixel un valor máximo o mínimo dependiendo del valor de los pixeles vecinos y el filtro utilizado, resultando en un desplazamiento de los límites de las formas binarias (Joshi, 2019).

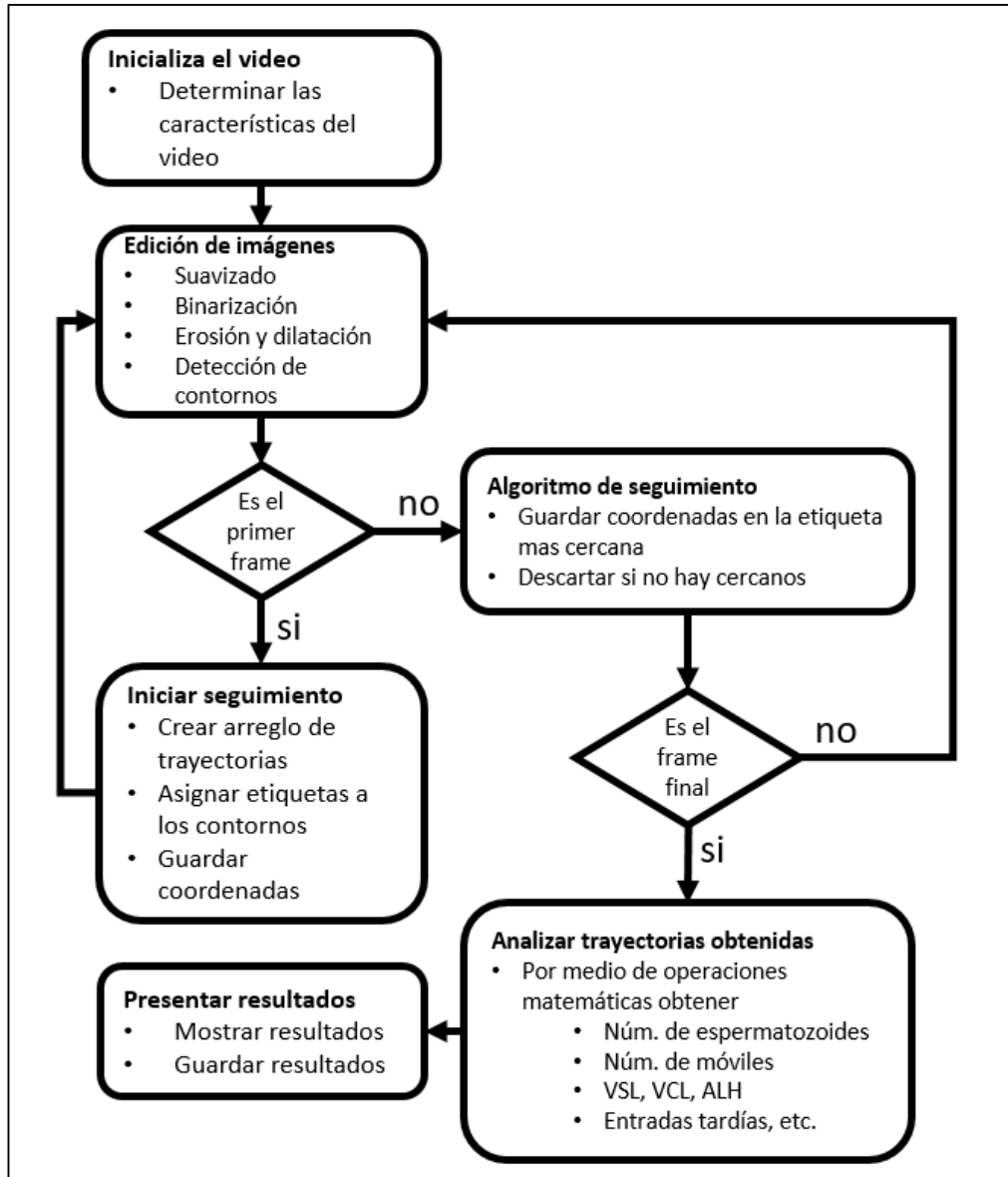
Detección de contornos. Con la función de detección de contornos se determina el número de células en cada cuadro y la posición de cada una de ellas en el plano de observación. Más aún se obtienen las dimensiones de cada contorno lo que ayuda en la discriminación de contornos que no representen un espermatozoide (Joshi, 2019).

Calibración de parámetros. Cada una de las funciones mencionadas anteriormente funciona basada en uno o varios parámetros que modifican la forma con la que cada filtro es aplicado a la imagen por lo que cada filtro se debe calibrar de forma independiente antes de aplicar el siguiente, una vez que todos los parámetros han sido ajustados al nivel deseado, permanecen igual durante el resto del proceso.

3.3 Estructura del software de análisis.

En la Figura 1 se muestran de forma general los procesos de edición por los que pasa cada uno de los cuadros del video.

Fig. 1. Diagrama de flujo del proceso de análisis.



Fuente: elaboración propia

Como partes principales de este diagrama de flujo está el módulo de detección de células que consta de la etapa de filtrado, segmentación y detección de contornos. El módulo de seguimiento que conecta las posiciones en cuadros consecutivos por medio de un algoritmo basado en la teoría del vecino más cercano (K-NN) que se explicará más adelante. Finalmente, el módulo de cálculo de variables que es donde se hacen los cálculos con los datos de las trayectorias obtenidos y se obtienen valores representativos de la muestra de esperma mostrada en el video.

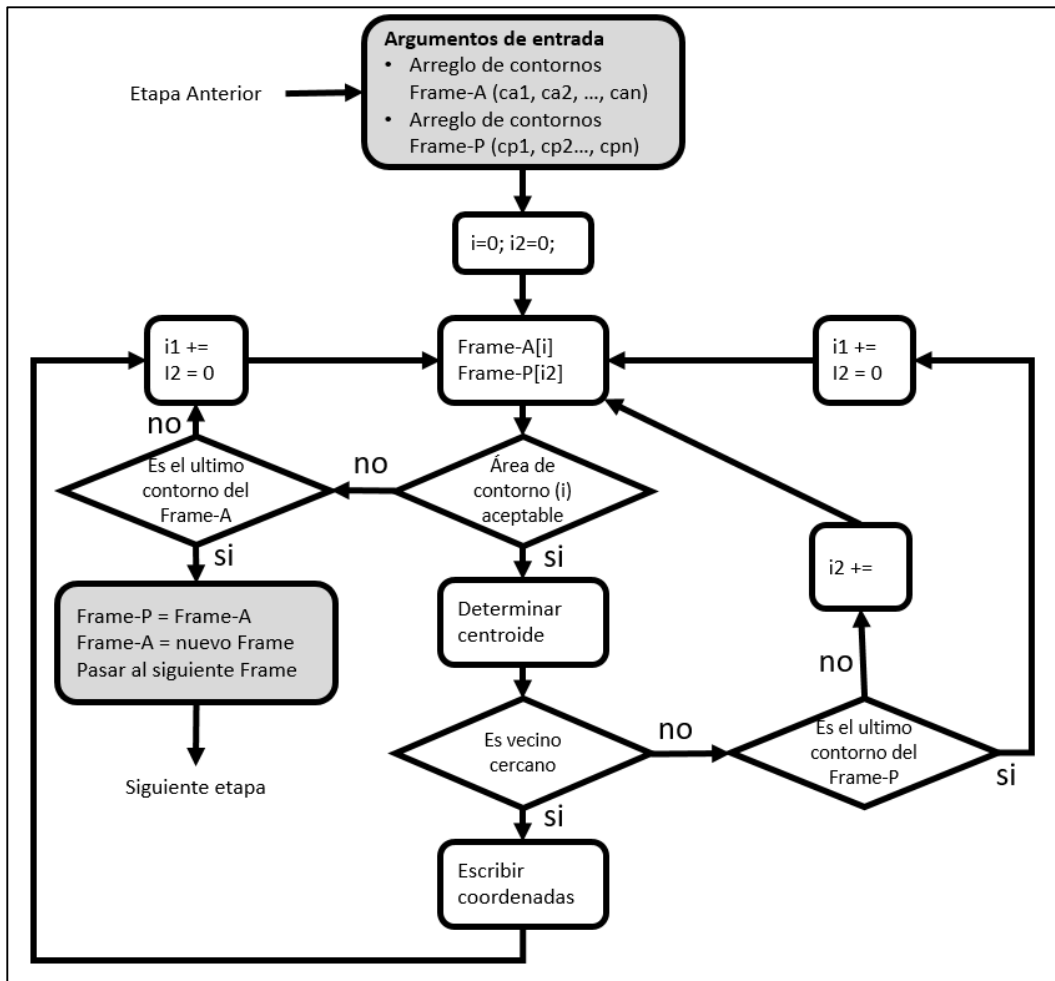
3.4 Algoritmo de seguimiento.

El algoritmo de seguimiento es el encargado de conectar la secuencia de coordenadas correspondientes a cada espermatozoide por separado, y determinar así las trayectorias que éstos siguen a través del tiempo. Como se ha mencionado antes existen algoritmos muy sofisticados que utilizan por ejemplo filtros de Kalman (Sørensen, 2008), el algoritmo multi-hipotesis (Urbano, 2016), o el algoritmo de asociación de datos probabilísticos conjuntos (Urbano, 2016). Este estudio desarrolla un algoritmo de seguimiento desde cero basado en el algoritmo del vecino más cercano (K-NN) que es el más utilizado en la gran mayoría de los instrumentos CASA comerciales el cual asume que las coordenadas más cercanas entre dos fotos consecutivas están conectadas.

Al analizar el primer frame se genera una matriz llamada matriz-T que contendrá las trayectorias de cada uno de los espermatozoides presentes en dicho frame, el tamaño de la matriz es determinado por la cantidad de células detectadas y el número de frames que contiene el video. Antes de pasar al siguiente frame se asignan las coordenadas de todos los contornos detectados al primer renglón de la matriz-T. Cada trayectoria entonces será representada por un par de columnas de la matriz-T en las que estará la secuencia de posiciones en el plano definida por coordenadas x , y .

Después de crear y asignar las posiciones iniciales de los espermatozoides se inicia el algoritmo de seguimiento mostrado en la figura 2 donde se comparan las coordenadas del conjunto de células detectadas en un frame con el conjunto de coordenadas del último renglón agregado a la matriz-T agregando renglones conforme se analiza la secuencia de video.

Fig. 2. Diagrama de flujo del algoritmo de seguimiento



Fuente: elaboración propia

Cuando se ha terminado de analizar el ultimo frame, la matriz contendrá los datos de cada una de las trayectorias seguidas por los espermatozoides presentes desde el comienzo del video. Este algoritmo cuenta con ciertos puntos negativos por ejemplo cuando dos o más células colisionan el algoritmo termina ahí dichas trayectorias excepto una y continua el seguimiento únicamente con esta. Otro punto es que las trayectorias guardadas serán únicamente las de las células presentes desde el principio del video excluyendo todos aquellos espermatozoides que ingresen al campo de visión durante el resto del video. Como añadido el algoritmo lleva un registro de la cantidad de células en cada frame y al final calcula el promedio de células totales en el video.

3.5 Análisis de datos

El análisis de los datos depende enormemente de los datos recolectados en la matriz-T además de otros datos obtenidos de las características del video y la cuenta de células en cada frame.

La primera variable a determinar es la cantidad de espermatozoides total en campo (TC) la cual podemos obtener mediante el registro de células iniciales en la matriz-T que es la cuenta de células iniciales en el video, aunque es más recomendable utilizar el promedio de células totales que toma en cuenta las células que entran y salen del campo de visión durante la secuencia de imágenes este dato es un parámetro de salida del algoritmo de seguimiento.

El porcentaje de móviles progresivas es obtenido dividiendo el número de trayectorias en la matriz-T que presenten un desplazamiento mínimo aceptable entre su posición inicial y final (MP) entre el número total de células presentes.

$$\%MP = (MP)/(TC) \quad (5)$$

La velocidad curvilínea (VCL) se calcula dividiendo la distancia total que la cabeza del espermatozoide cubre entre el periodo observado.

La velocidad de línea recta (VSL) es determinada por la distancia entre el punto inicial y final de la trayectoria entre el tiempo total.

La velocidad de trayecto promedio (VAP) se calcula suavizando la trayectoria original. Las coordenadas de cada punto en la trayectoria se promedian con las de los dos puntos adyacentes entonces se divide la distancia total de la trayectoria suavizada entre el tiempo.

La linealidad (LIN) rectitud (STR) y tambaleo (WOB) se calcula mediante los valores obtenidos de VSL, VCL y VAP con las relaciones siguientes.

$$LIN = \left(\frac{VSL}{VCL} \right) * 100 \quad (6)$$

$$STR = \left(\frac{VSL}{VAP} \right) * 100 \quad (7)$$

$$WOB = \left(\frac{VAP}{VSL} \right) * 100 \quad (8)$$

Finalmente, la amplitud de desplazamiento lateral de la cabeza (ALH) que representa el ancho total del movimiento lateral de la cabeza del espermatozoide es calculado mediante un valor llamado riser (Riser es la distancia que existe entre la posición de un punto de la trayectoria original hasta el punto de la trayectoria suavizada correspondiente al mismo frame), en algún momento el riser alcanza un valor máximo, entonces:

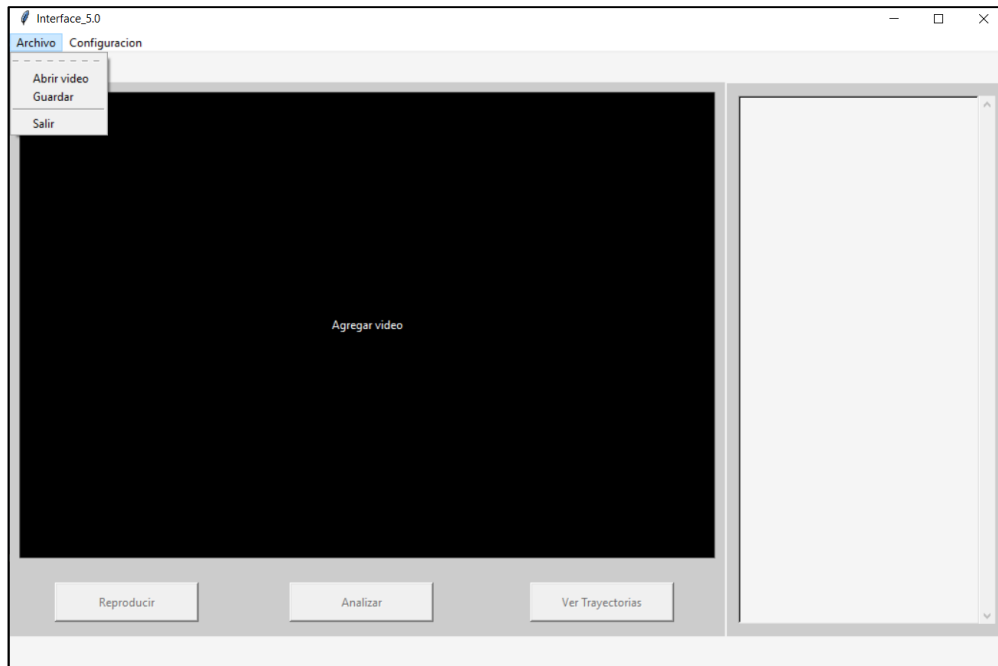
$$ALH = (riser \text{ máximo}) * 2 \quad (9)$$

Todas las características de distancia y velocidad son calculadas en pixeles/segundo y pixeles respectivamente por lo que se debe hacer la conversión de pixeles a micrómetros dependiendo de la amplificación de la imagen.

3.6 Presentación de los datos.

Diseño de interfaz de usuario. Los procedimientos descritos anteriormente se realizan de forma mayormente automática, sin embargo, se ha implementado con ayuda de la librería Tkinter una interfaz de usuario sencilla Figura 3 para facilitar el manejo del software incluyendo opciones como: selección de archivo, reproducción del video, opción para ver las trayectorias dibujadas sobre el video original, opción para configurar los parámetros utilizados, un área de visualización de datos, y opción de guardar los resultados de cada análisis en un archivo Excel o archivo de texto.

Fig. 3. Interfaz del software CASA



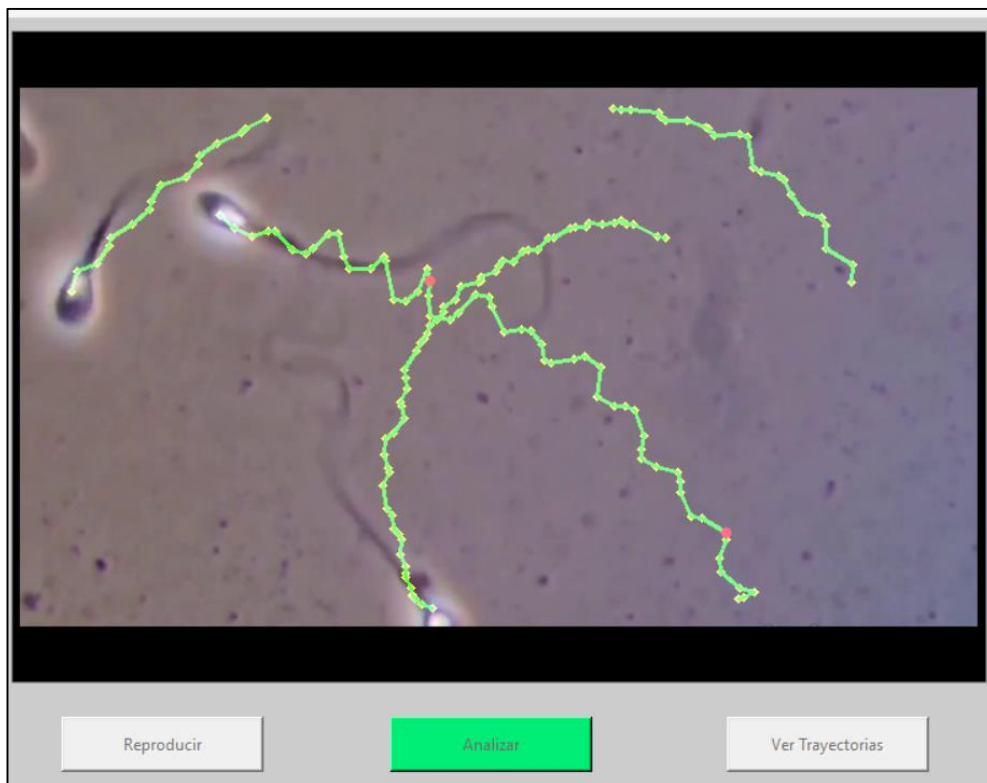
Fuente: elaboración propia

Presentación de los datos. Los resultados del análisis del video son mostrados en un contenedor de texto desplazable, se presentan las mediciones de número de espermatozoides en campo, número de móviles progresivos, los datos promedio de las características dinámicas de todas las trayectorias registradas en la matriz-T, además de otros datos como entradas tardías y duración del análisis.

Adicionalmente, se ofrece al usuario la posibilidad de ver en pantalla las trayectorias dibujadas sobre el video original como se muestra en la Figura 4, donde se puede apreciar de forma visual el comportamiento de cada célula.

Almacenamiento de los datos. No es solo de los datos mostrados en pantalla, se guardan todos los datos contenidos en la matriz-T, los resultados obtenidos del cálculo de variables y se identifica etiquetando a cada archivo guardado con la fecha y hora del análisis, al guardar los datos se ofrece al usuario la opción de elegir el formato del archivo que puede ser hoja de cálculo de Excel o archivo de texto.

Fig. 4. Trayectorias dibujadas sobre el video original.



Fuente: elaboración propia

Capítulo IV

4 Resultados.

Cada sistema de análisis CASA cuenta con particularidades asociadas a su desarrollador o compañía desarrolladora, pero todos están de alguna forma regulados bajo ciertos estándares de control establecidos y probados (Mortimer, 2000). Que aseguran resultados consistentes difiriendo uno de otro únicamente en sus métodos de análisis y en la precisión y exactitud de sus mediciones. Aplicar algoritmos de detección y seguimiento más robustos y sofisticados requiere a su vez de una mayor capacidad de procesamiento que gracias a los avances en electrónica y computación es cada vez más fácil de conseguir. El presente artículo enfocado principalmente en el análisis de semen bovino se crea con la finalidad de ofrecer a productores y veterinarios cuya producción depende de la inseminación artificial una alternativa a los costosos sistemas CASA comerciales usados comúnmente.

Capítulo V

5 Conclusiones.

La mayoría de los sistemas más sofisticados que implementan algoritmos de análisis basados en teorías probabilísticas, filtros de Kalman o multi-hipótesis. Son usados principalmente en clínicas especializadas en fertilidad humana, mientras que la mayoría de productores ganaderos y veterinarios interesados en inseminación artificial optan por sistemas comerciales con algoritmos de veinte o treinta años de antigüedad, pero como se ha mostrado en el presente documento, dichos algoritmos y su implementación en sistemas CASA basados en código abierto abre la posibilidad para que cualquiera pueda acceder a ellos en línea e implementar el que más convenga a sus necesidades sin costo alguno, comunidades como github.com ofrecen cada vez más alternativas de software en general y los sistemas CASA no son la excepción. Uno de los puntos más prometedores concernientes al análisis de imágenes es la aplicación de técnicas de machine learning y redes neuronales, dichas técnicas en el área de sistemas CASA está limitada por las escasas bases de datos disponibles con las que se pueda entrenar modelos de predicción, sin embargo, con el paso del tiempo más y más información está siendo almacenada en la nube y en algún momento dichas bases de datos serán una realidad lo que abrirá nuevas posibilidades.

Capítulo VI

6 Referencias.

1. Acott, T. S., Katz, D. F., & Hoskins, D. D. (1983). Movement characteristics of bovine epididymal spermatozoa: effects of forward motility protein and epididymal maturation. *Biology of reproduction*, 29(2), 389-399.
2. Mortimer, S. T. (1997). A critical review of the physiological importance and analysis of sperm movement in mammals. *Human reproduction update*, 3(5), 403-439.
3. Beresford-Smith, B., & Van Helden, D. F. (1994, November). Applications of radar tracking algorithms to motion analysis in biomedical images. In *Proceedings of 1st International Conference on Image Processing* (Vol. 1, pp. 411-415). IEEE.
4. Gray, J. (1958). The movement of the spermatozoa of the bull. *Journal of Experimental Biology*, 35(1), 96-108.
5. Amann, R. P., & Waberski, D. (2014). Computer-assisted sperm analysis (CASA): capabilities and potential developments. *Theriogenology*, 81(1), 5-17.
6. Daloglu, M. U., & Ozcan, A. (2017). Computational imaging of sperm locomotion. *Biology of reproduction*, 97(2), 182-188.
7. Jaqaman, K., Loerke, D., Mettlen, M., Kuwata, H., Grinstein, S., Schmid, S. L., & Danuser, G. (2008). Robust single-particle tracking in live-cell time-lapse sequences. *Nature methods*, 5(8), 695-702.
8. Aitken, R. J. (2006). Sperm function tests and fertility. *International journal of andrology*, 29(1), 69-75.
9. Mortimer, S. T. (2000). CASA—practical aspects. *Journal of andrology*, 21(4), 515-524.
10. Katz, D. F., Davis, R. O., Delandmeter, B. A., & Overstreet, J. W. (1985). Real-time analysis of sperm motion using automatic video image digitization. *Computer methods and programs in biomedicine*, 21(3), 173-182.
11. Sørensen, L., Østergaard, J., Johansen, P., & de Bruijne, M. (2008, March). Multi-object tracking of human spermatozoa. In *Medical Imaging 2008: Image Processing* (Vol. 6914, p. 69142C). International Society for Optics and Photonics.

12. Urbano, L. F., Masson, P., VerMilyea, M., & Kam, M. (2016). Automatic tracking and motility analysis of human sperm in time-lapse images. *IEEE transactions on medical imaging*, 36(3), 792-801.
13. Yang, H. F., Descombes, X., Prigent, S., Malandain, G., Druart, X., & Plouraboué, F. (2014, April). Head tracking and flagellum tracing for sperm motility analysis. In *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)* (pp. 310-313). IEEE.
14. García, M. J. L., Felices, A. U., & Povedano, M. C. (2012). *Manual de laboratorio para el análisis del semen*. OmniaScience.
15. Cardona Maya, W. (2010). Manual de procesamiento de semen humano de la Organización Mundial de la Salud-2010. *Actas urológicas españolas*, 34(7), 577-578.
16. Joshi, P. (2015). *OpenCV with Python by example*. Packt Publishing Ltd.
17. Foote, R.H. (2005). An Overview. In: *Highlights in Dairy Cattle Reproduction in the last 100 years*. The Internet-First University Press, Ithaca, New York. pp 5-15.
18. Foote, R.H. (2002). *The history of artificial insemination: Selected notes and notables*. American Society of Animal Science.
19. Milovanov, N.K. (1964). *Artificial Insemination of Livestock in the U.S.S.R.* Jerusalem Tech. Services, U.S. Dept. Commerce, Washington, DC.
20. Herman, H.A. (1981). History and development of artificial insemination. In: *Improving Cattle by Millions: NAAB and the Development and Worldwide Application of Artificial Insemination*. University of Missouri Press Columbia and London. pp 1-13.
21. Holt, W.V. (1997). Alternative strategies for the long-term preservation of spermatozoa. *Reproduction Fertility and Development*, 9: 309-319.
22. Watson, P.F. (2000). The causes of reduced fertility with cryopreserved semen. *Animal Reproduction Science*, 60: 481-492.
23. Holt, W.V. and North, R.D. (1984). Partially irreversible cold-induced lipid phase transitions in mammalian sperm plasma membrane domains: freeze-fracture study. *Journal of Experimental Zoology*, 230: 473-483.

24. Salisbury, G.W., VanDemark, N.L. and Lodge, J.R. (1978a). Extenders and extension of unfrozen semen. *Physiology of Reproduction and Artificial Insemination of Cattle*. 2nd ed. W.H. Freeman Co., San Francisco. pp. 442-493.
25. Amann RP, Katz DF (2004) Reflections on CASA after 25 years. *J Androl* 25:317–325.
26. Lu JC (2012) Standardization and quality control for sperm morphology analysis (Chinese). *Chin J Clin Lab Sci* 30: 834–836, 841.
27. Lu, J. C., Huang, Y. F., & Lü, N. Q. (2014). Computer-aided sperm analysis: past, present and future. *Andrologia*, 46(4), 329-338.
28. Inoue, S., and Spring, K. R. (1997). "Video Microscopy, The Fundamentals." Plenum Press, New York.
29. Castleman, K. R. (1993). Resolution and sampling requirements for digital image processing, analysis, and display. In "Electronic Light Microscopy" (D. Shotton, ed.), p. 71-94. Wiley-Liss, New York.
30. Ramm, P. (1994). Advanced image analysis systems in cell, molecular, and neurobiology applications. *J. Neurosci. Methods* 54, 131-149.
31. Shotton, D. (1993). An introduction to the electronic acquisition of light microscope images. In *Electronic light microscopy* (pp. 1-38). Wiley-Liss New York.
32. DURANGO, TERCER LUGAR NACIONAL EN EXPORTACIÓN DE GANADO BOVINO EN PIE A USA: SAGDR. (s. f.). Secretaría de Agricultura, Ganadería y Desarrollo Rural. <https://desarrollorural.durango.gob.mx/durango-tercer-lugar-nacional-en-exportacion-de-ganado-bovino-en-pie-a-usa-sagdr/>
33. Russ, J. C. (2012). *Computer-assisted microscopy: the measurement and analysis of images*. Springer Science & Business Media.
34. Wolf, D. E., & Sluder, G. (1998). *Video Microscopy*. Academic Press.
35. Rittscher, J., Machiraju, R., & Wong, S. T. (2008). *Microscopic image analysis for life science applications*. Artech House.
36. Murphy, D. B. (2001). *Fundamentals of light microscopy and electronic imaging*.

37. Zernike, F. (1955). How I discovered phase contrast. *Science*, 121(3141), 345-349.
38. Gonzales, R. C., & Wintz, P. (1987). *Digital image processing*. Addison-Wesley Longman Publishing Co., Inc.