

**DISEÑO E IMPLEMENTACIÓN DE UNA HERRAMIENTA COMPUTACIONAL
PARA EL ANÁLISIS ASISTIDO DE MOTILIDAD ESPERMÁTICA EN BOVINOS
MEDIANTE TÉCNICAS DE VISIÓN ARTIFICIAL**

DANIEL FERNANDO RODRÍGUEZ MONTAÑA



**UNIVERSIDAD DE CUNDINAMARCA
FACULTAD DE INGENIERÍA
INGENIERÍA ELECTRÓNICA
FUSAGASUGÁ
2016**

**DISEÑO E IMPLEMENTACIÓN DE UNA HERRAMIENTA COMPUTACIONAL
PARA EL ANÁLISIS ASISTIDO DE MOTILIDAD ESPERMÁTICA EN BOVINOS
MEDIANTE TÉCNICAS DE VISIÓN ARTIFICIAL**

DANIEL FERNANDO RODRÍGUEZ MONTAÑA

**Trabajo de grado presentado como requisito parcial
para optar por el título de Ingeniero Electrónico**

**Director:
MSc. EDGAR EDUARDO ROA GUERRERO**

**Línea de investigación:
Software, sistemas emergentes y nuevas tecnologías**

**UNIVERSIDAD DE CUNDINAMARCA
FACULTAD DE INGENIERÍA
INGENIERÍA ELECTRÓNICA
FUSAGASUGÁ
2016**

Nota de aceptación:

Firma del presidente del jurado

Firma del jurado

Firma del jurado

Fusagasugá ____ de _____ del _____

*A mi madre, por darme la vida y formarme como un ser humano íntegro
A mis hermanos, por brindarme su apoyo incondicional*

Agradecimientos

A mi director de tesis, Edgar Roa Guerrero, por su esfuerzo y dedicación, quien con su conocimiento y experiencia me guio durante todo el proceso investigativo, permitiendo la consecución de este proyecto.

A la doctora Claudia Jiménez, por abrirnos las puertas del laboratorio de andrología de la Universidad Nacional de Colombia y permitirnos adquirir el material necesario para el desarrollo del proyecto.

Al doctor Walter Cardona Maya, investigador del laboratorio de reproducción de la Universidad de Antioquia, por su amable colaboración en el análisis espermático de motilidad y su importante contribución en el proceso de validación de resultados.

Al zootecnista Diego Garzón, por atendernos en su laboratorio y aclararnos muchos conceptos de microscopía y motilidad espermática en bovinos.

A los docentes de la Facultad de Ingeniería, quienes contribuyeron en mi formación profesional. A mis compañeros de estudio, con quienes compartí momentos valiosos. A todos muchas gracias.

Contenido

Resumen	15
Abstract	17
Introducción	18
1. Capítulo I. El problema	19
1.1. Planteamiento del problema	19
1.2. Objetivos del estudio.....	20
1.2.1. Objetivo general	20
1.2.2. Objetivos específicos.....	20
1.3. Justificación	21
1.4. Alcances y limitaciones.....	22
1.4.1. Alcances.....	22
1.4.2. Limitaciones.....	22
2. Capítulo II. Marco teórico	23
2.1. Fundamentos teóricos	23
2.1.1. Ganadería bovina en Colombia	23
2.1.2. Fertilidad en toros.....	24
2.1.3. Motilidad espermática.....	25
2.1.4. Sistemas comerciales CASA	26
2.1.5. Visión artificial.....	28
2.2. Estado del arte.....	31
2.2.1. Algoritmos de análisis espermático	31
2.2.2. Evaluación de sistemas CASA	38
3. Capítulo III. Diseño metodológico	46
3.1. Adquisición y pre-procesamiento	46
3.2. Segmentación y detección de espermatozoides.....	46
3.2.1. Modelos gaussianos adaptativos.....	47
3.2.2. Operadores morfológicos	48
3.2.3. Análisis de componentes conectados	48

3.2.4.	Análisis discriminante lineal.....	48
3.3.	Asignación y construcción de trayectorias	50
3.3.1.	Algoritmo de asignación de Munkres.....	50
3.3.2.	Filtro de Kalman	51
3.4.	Análisis de los parámetros de motilidad espermática	54
3.4.1.	Velocidad curvilínea (VCL)	54
3.4.2.	Velocidad promedio (VAP)	54
3.4.3.	Velocidad rectilínea (VSL)	55
3.4.4.	Índice de linealidad (LIN)	55
3.4.5.	Índice de rectitud (STR).....	55
3.4.6.	Índice de oscilación (WOB)	55
3.4.7.	Motilidad total y progresiva (MT y MP)	56
4.	Capítulo IV. Análisis de resultados.....	58
4.1.	Procesamiento de video	58
4.1.1.	Segmentación y detección de espermatozoides	59
4.1.2.	Asignación y construcción de trayectorias.....	63
4.1.3.	Caracterización de motilidad espermática.....	65
4.2.	Validación de la herramienta computacional	67
4.3.	Análisis estadístico	69
	Conclusiones y trabajo futuro	75
	Bibliografía	78

Índice de tablas

Tabla 1. Ventajas y desventajas del análisis manual y el análisis asistido por computador (CASA). Traducida de Schleh & Leoni, 2013.	21
Tabla 2. Clasificación de motilidad masal. Tomado de Hafez, 1989.....	25
Tabla 3. Clasificación de motilidad individual. Tomado de Hafez, 1989.....	26
Tabla 4. Comparación entre espermatozoides existentes y detectados, y porcentajes de precisión. Tomado de Q. Li et al., 2012.	35
Tabla 5. Tabla de confusión de viabilidad en 14 individuos. Tomado de E. Roa et al., 2012.	36
Tabla 6. Tasas de error en la detección de trayectorias para los diferentes grados de clasificación tanto en semen normal como diluido. Tomado de M. Tomlinson et al., 2010.	39
Tabla 7. Variación de parámetros de motilidad espermática según velocidad de fotogramas. Tomado de C. Castellini et al., 2011.....	40
Tabla 8. Precisión de las metodologías evaluadas. Tomado de A. Nurhadiyatna et al., 2014.	43
Tabla 9. Ejemplo de coordenadas de una trayectoria en 8 fotogramas. Elaborada por el autor.	54
Tabla 10. Resultados del análisis manual por parte del técnico experto. Elaborada por el autor.	68
Tabla 11. Resultados del análisis semi-automático por parte de la herramienta computacional. Elaborada por el autor.	69

Índice de figuras

Figura 1. Árbol de problemas de la situación. Elaborada por el autor.....	20
Figura 2. Inventario bovino. Tomado de FEDEGÁN, 2013.	23
Figura 3. Representación visual de los parámetros de motilidad más comunes en sistemas CASA. Tomado de Amann & Waberski, 2014.....	27
Figura 4. Interfaz de resultados en sistemas CASA. (A) Análisis de motilidad (B) Análisis de morfología. Adaptado de Amann & Waberski, 2014.....	28
Figura 5. Representación matricial de una imagen digital. Elaborada por el autor.	29
Figura 6. (A) Imagen con ruido (B) Imagen filtrada. Tomado de <i>mathworks.com</i> . 29	
Figura 7. (A) Imagen original (B) Imagen segmentada. Tomado de La Serna & Román, 2009.....	30
Figura 8. Representación de una imagen binaria. Tomado de <i>en.wikiperdia.org</i> , 2016.....	31
Figura 9. Metodología propuesta por los autores. Tomado de J. Pascual et al., 2008.....	32
Figura 10. Metodología propuesta por los autores. Tomado de V. Abbiramy et al., 2010.....	33
Figura 11. Fotogramas 1-4. (a) Trasformada distancia euclidiana. (b) Algoritmo Watershed con centroide detectado. Tomado de M. Ravanfar et al., 2011.	34
Figura 12. Detección de flagelo en sucesión de 3 fotogramas. (a) En condiciones normales. (b) Bajo beta-mercaptoethanol (bME). Tomado de H. Yang et al., 2014.	37
Figura 13. (a) Todas las trayectorias detectadas. (b) Algunas seleccionadas. (c) identificación de parámetros de motilidad. Tomado de M. Elsayed et al., 2015.....	38
Figura 14. Velocidad curvilínea para cada especie y cada sistema a 25 y 50 Hz en 12 segundos de video. Tomado de S. Boryshpolets et al., 2013.	41
Figura 15. Comparación de los tres métodos para cada tipo de motilidad. (a) Progresiva rápida. (b) Progresiva lenta. (c) No progresiva. Tomado de I. Palacín et al., 2013.	42

Figura 16. “Clustering” de datos de motilidad como validación. Espermatozoides con un valor VAP alto tienden a presentar a su vez una linealidad alta. Tomado de R. Amann & Waberski, 2014.	44
Figura 17. Representación tridimensional de una combinación de modelos gaussianos. Tomado de <i>www.researchgate.net</i>	47
Figura 18. Agrupamiento de elementos en dos clases distintas. Tomado de <i>www.mathematica-journal.com</i> , 2016.....	49
Figura 19. Representación de una matriz de costos. Los elementos de n se relacionan a m a través de una asignación de mínimo costo total. Elaborado por el autor.	51
Figura 20. Diagrama del proceso de funcionamiento del filtro de Kalman discreto. Tomado de G. Prado, 2005.	52
Figura 21. Representación de la asignación de detecciones de espermatozoides a sus trayectorias mediante el algoritmo de Munkres y el filtro de Kalman. Elaborada por el autor.	53
Figura 22. Diagrama de flujo de la clasificación de los tipos de motilidad. Elaborada por el autor.	56
Figura 23. Diagrama de flujo del algoritmo. Elaborada por el autor.	58
Figura 24. (A) Fotograma original. (B) Ajuste de contraste. Elaborada por el autor.	59
Figura 25. Secuencia de fotogramas mostrando la segmentación de tres espermatozoides en movimiento por el método adaptativo de modelos gaussianos. Elaborada por el autor.	60
Figura 26. (A) Imagen binaria original (B) Filtrada por operadores morfológicos. Elaborada por el autor.	61
Figura 27. (A) Imagen filtrada. (B) Centroides calculados. Elaborada por el autor.	61
Figura 28. (A) Fotograma original. (B) Segmentación por análisis discriminante de Fisher. Elaborada por el autor.....	62
Figura 29. Separación de las clases de pixeles de espermatozoides y fondo mediante un plano discriminante. Elaborada por el autor.	63

Figura 30. Fotogramas en diferentes tiempos 1,2,3. (A) En puntos rojo se marcan las detecciones de centroides sin asignar. (B) En líneas azules se ilustran las trayectorias construidas. Elaborada por el autor.....	64
Figura 31. Diferentes comportamientos cinéticos observados en los videos procesados. En verde VCL, en azul VAP y en rojo VSL. Elaborada por el autor...	66
Figura 32. Relación de STR y VAP en la clasificación de motilidad de los espermatozoides analizados. (Rojo) Grado C. (Azul) Grado B. (Verde) Grado A.	67
Figura 33. Regresión lineal y coeficiente de determinación de los espermatozoides en clasificados en A+B. Elaborada por el autor.	70
Figura 34. Regresión lineal y coeficiente de determinación de los espermatozoides en clasificados en C. Elaborada por el autor.	70
Figura 35. Regresión lineal y coeficiente de determinación de los espermatozoides en clasificados en D. Elaborada por el autor.	71
Figura 36. Regresión lineal y coeficiente de determinación de la clasificación de espermatozoides en todas las categorías. Elaborada por el autor.....	71
Figura 37. Regresión lineal y coeficiente de determinación del cálculo del porcentaje de motilidad total. Elaborada por el autor.....	72
Figura 38. Regresión lineal y coeficiente de determinación del cálculo del porcentaje de motilidad progresiva. Elaborada por el autor.....	72
Figura 39. Comprobación Bland Altman para las clases A+B, C y D. Elaborada por el autor.	73
Figura 40. Comparación de los resultados de porcentaje de motilidad progresiva dados por el experto y por la herramienta computacional diseñada. Elaborada por el autor.	74

Glosario

Algoritmo: conjunto ordenado de operaciones sistemáticas que permite hacer un cálculo y hallar la solución de un tipo de problemas.

Andrología: parte de la medicina encargada del estudio de cualquier aspecto relacionado con la función sexual y reproducción masculina.

Automatización: aplicación de procedimientos automáticos en la realización de un proceso o en una industria.

CASA: análisis espermático asistido por computador

Centroide: coordenada que hace referencia al centro de masa de un objeto o figura bidimensional en una imagen.

Clustering: agrupación estructural de datos que permiten un análisis estadístico

Coefficiente de determinación: estadístico relacionado a la regresión lineal que hace referencia al nivel de correlación y capacidad de predicción de un conjunto de datos.

Concentración: cantidad de espermatozoides con respecto al volumen de la muestra.

Criopreservación: proceso en el cual células o tejidos son congelados a muy bajas temperaturas para disminuir sus funciones vitales y poderlos mantener en condiciones de vida suspendida por mucho tiempo.

Distancia euclidiana: función no negativa usada en diversos contextos para calcular la distancia entre dos puntos, ya sea en un plano o el espacio.

Distribución gaussiana: distribución de probabilidad formada por una función de densidad en forma de campana descrita por una media y una desviación estándar.

Escala de grises: es la representación de una imagen en la que cada pixel se dibuja usando un valor numérico individual que representa su luminancia, en una escala que se extiende entre blanco y negro.

Espermograma: examen que busca determinar las características de los espermatozoides y algunos parámetros físicos del semen.

Excentricidad: parámetro que determina el grado de desviación de una sección cónica con respecto a una circunferencia.

Fertilización in-vitro: técnica por la cual la fecundación de los ovocitos por los espermatozoides se realiza fuera del cuerpo de la madre.

Filtrado de imágenes: técnica de procesamiento digital de imágenes por medio de la cual se eliminan características no deseadas en una imagen.

Flagelo: apéndice movable en forma de látigo que permite a ciertos microorganismos, como los espermatozoides, desplazarse a través de un medio líquido.

Fotograma: imagen digital única que pertenece a una secuencia de imágenes que conforman un video.

Histograma: gráfico de la representación de distribuciones de frecuencias de los niveles de los valores de píxeles en una imagen a escala de grises.

Imagen binaria: imagen digital que tiene únicamente dos valores posibles para cada píxel, generalmente 1 y 0.

LIN: índice de linealidad de la trayectoria de un espermatozoide.

Matriz: arreglo bidimensional de números, donde cada uno cuenta con un doble subíndice dados por su posición en filas y columnas.

Microscopio de contraste de fase: instrumento que permite observar microorganismos mediante las pequeñas diferencias en los índices de refracción en las distintas partes de éstos.

Morfología: forma de la estructura de la cabeza y cuerpo del espermatozoide.

Motilidad: facultad de los espermatozoides de moverse o desplazarse.

OpenCV: biblioteca libre que permite el uso de algoritmos de visión artificial.

Operadores morfológicos: técnica para el análisis y tratamiento de estructuras geométricas presentes en imágenes binarias.

Pixel: menor unidad homogénea en color o escala de grises que forma parte de una imagen digital.

Progresividad: atributo de desplazamiento con tendencia lineal.

RGB: composición del color en términos de la intensidad de los colores primarios de la luz.

Segmentación: técnica de visión artificial, proceso de dividir una imagen digital en varias partes o grupos de píxeles.

Software: conjunto de programas y rutinas que permiten a la computadora realizar determinadas tareas. Conforman una herramienta computacional.

STR: índice de rectitud de la trayectoria de un espermatozoide.

VAP: velocidad promedio de un espermatozoide.

VCL: velocidad curvilínea de un espermatozoide.

Viabilidad: cantidad de espermatozoides vivos con respecto al total.

Visión artificial: subcampo de la inteligencia artificial que proporciona métodos para adquirir, procesar y analizar imágenes del mundo real.

VSL: velocidad rectilínea de un espermatozoide.

WOB: índice de ondulación de la trayectoria de un espermatozoide.

Resumen

El campo de la visión artificial ha tenido un gran impacto a nivel industrial en los últimos años y se ha convertido en una alternativa fundamental para una gran parte de procesos que requieran automatización. Las tareas repetitivas que normalmente realiza un operario tienden a ser ahora reemplazadas por las de una máquina o inteligencia artificial. Por lo tanto, las técnicas de visión artificial tales como el procesamiento digital de imágenes y video permiten brindar un soporte tecnológico confiable a procesos comúnmente manuales que presenten dificultades en su realización.

Uno de esos procesos es la observación y evaluación de muestras seminales bajo un microscopio con el fin de analizar motilidad espermática individual en bovinos para el área de zootecnia y reproducción animal. Este procedimiento normalmente llevado a cabo por un técnico experto de forma manual presenta los inconvenientes de ser muy subjetivo, consumir largos periodos de tiempo y ser exhaustivo. En consecuencia, los resultados en el diagnóstico de fertilidad presentan poca repetitividad y poca confiabilidad.

Los sistemas comerciales CASA, que son sistemas de visión por computadora semi-automáticos, tienen la capacidad de proveer una solución a este problema, sin embargo, presentan las desventajas de ser muy poco asequibles, altamente costosos y carecer de una estandarización adecuada. Por tanto, se convierte en una alternativa muy poco viable. Teniendo en cuenta esto, se propuso el diseño y desarrollo de una herramienta computacional que aporte al análisis de motilidad espermática con una precisión comparable a los sistemas CASA, pero con un costo sustancialmente bajo.

Las técnicas de visión artificial implementadas fueron la detección de movimiento por modelos gaussianos adaptativos, la segmentación de cabezas de espermatozoides a través de un análisis discriminante lineal de Fisher, el filtrado de imágenes binarias por operadores morfológicos, la detección de centroides mediante un análisis de componentes conectados y la asignación de detecciones a trayectorias empleando el algoritmo de Munkres junto con el filtro de Kalman. Los resultados obtenidos evidencian un coeficiente de determinación del 0.8143 en la correlación del análisis de motilidad con respecto al análisis manual por parte de un experto. Finalmente, mediante la prueba de comprobación Bland-Altman no se encontró variabilidad significativa entre la medida de motilidad espermática por

parte del experto y la herramienta computacional implementada, los resultados obtenidos muestran las diferencias de la medida manual con respecto al algoritmo, con una relación de 0,966 en un intervalo de confianza del 95% de (3,058 – 3,19), indicando que las medidas fueron muy similares. De esta manera, la metodología propuesta contribuirá al análisis de motilidad espermática en bovinos y en consecuencia beneficiará al sector veterinario y ganadero de la región.

Abstract

Artificial vision has had a major impact on an industrial field in recent years and has become a major alternative for a lot of processes that require automation. Repetitive tasks normally performed by an operator tend to be now replaced by a machine or artificial intelligence. Therefore, artificial vision techniques such as digital image and video processing allow us to give a reliable technological support to manual processes that commonly present difficulties in implementation.

One such process is the observation and evaluation of semen samples under a microscope in order to analyze single sperm motility in cattle for the area of animal husbandry and animal breeding. This procedure that is usually carried out manually by a skilled technician presents the drawbacks of being very subjective, consume long periods of time and be exhaustive. Consequently, the results in diagnosing fertility have low repetitiveness and low reliability.

CASA commercial systems, that are semi-automatic computer vision systems, have the ability to provide a solution to this problem, however, they have the disadvantages of not being easily affordable, being highly expensive and lack adequate standardization. Therefore, it becomes a low viable alternative. Given this, it was proposed the design and development of a computational tool that provides analysis of sperm motility with an accuracy comparable to CASA systems, but at a substantially lower cost.

The artificial vision techniques implemented were the motion detection by adaptive gaussian models, segmenting sperm heads through a linear discriminant analysis of Fisher, filtering binary images by morphological operators, detecting centroids by connected component analysis and assigning paths detections using Munkres algorithm along with the Kalman filter. The results show a coefficient of determination of 0.814 in the correlation of progressive motility analysis compared to manual analysis by an expert. Finally, through the verification test of Bland-Altman, no significant variability between the measurement of sperm motility by the expert and the designed computational tool was found, the results show the differences in manual measurement regarding to the algorithm with a ratio of 0,966 in a confidence interval of 95% (3.058 to 3.19), indicating that the measures were very similar. Thus, the proposed methodology will contribute to the analysis of sperm motility in cattle and thus benefit the veterinary and livestock sector in the region.

Introducción

En Colombia, la actividad agropecuaria y en su gran mayoría la ganadería bovina, tienen un gran impacto en la economía interna del territorio nacional (DANE, 2012). Nuestro país presenta un amplio comercio ganadero y, por ende, un mercado exigente relacionado a la ganadería bovina. Uno de los aspectos más importantes a la hora de comerciar con toros es su fertilidad, debido a que, siempre se requieren toros con una gran capacidad de preñar a grupos de hembras, así como también, su espermatozoide es utilizado para procesos de inseminación artificial (Contexto Ganadero, 2015). Por estas razones, se requieren exámenes andrológicos para determinar su capacidad de fertilizar, a través del análisis de la motilidad espermática, uno de los descriptores más importantes de la fertilidad en bovinos (SERIDA, 2005).

Hoy en día el análisis seminal generalmente se lleva a cabo de manera manual, es decir, mediante la observación física de los espermatozoides en el microscopio por parte de expertos en reproducción. No obstante, este procedimiento convencional está sujeto a errores debido a factores tales como la subjetividad en el diagnóstico, la falta de repetitividad en los resultados y el cansancio ocular del experto por largas jornadas de trabajo, afectando de esta manera el resultado del análisis (Roa, et al., 2012). Por esta razón, las técnicas de procesamiento de imágenes surgen como una propuesta prometedora para soportar los diagnósticos en fertilidad y caracterizar los criterios de análisis por parte de los técnicos expertos en reproducción animal.

Con motivo de aportar una solución, se propuso una metodología que disminuye los errores en los resultados del análisis de motilidad espermática mediante la implementación de una herramienta computacional que emplea técnicas de visión artificial a partir de videos, con el fin de proporcionar al especialista un soporte tecnológico, basado en resultados extraídos del análisis de parámetros objetivos. Estas técnicas de visión artificial se componen de diferentes métodos de procesamiento de imágenes y video, tales como: extracción de fondo, filtrado, operaciones morfológicas, análisis de regiones y detección de movimiento (Tekalp, 2015). A través de un algoritmo llamado filtro de Kalman se determinan trayectorias individuales de múltiples objetivos en movimiento. Finalmente, el método propuesto analiza las trayectorias de los espermatozoides y entrega resultados de motilidad espermática de acuerdo a los parámetros de motilidad convencionales (Fréour, et al., 2010) de los sistemas de análisis seminal asistido por computadora (CASA).

1. Capítulo I. El problema

1.1. Planteamiento del problema

El sector ganadero es uno de los sectores agropecuarios más preferidos por los campesinos debido a su alta productividad en los campos y fincas del territorio nacional (Mahecha & Gallego, 2002), particularmente en el departamento de Cundinamarca. Un ganadero siempre quiere conocer el estado de un toro objeto de comercialización antes de cerrar un negocio. La fertilidad, es entre otros, el criterio de calidad más importante que describe la preferencia de un toro sobre otro.

La fertilidad hace referencia a la probabilidad de un espécimen de toro de preñar a una o un conjunto de vacas. Este parámetro se determina a partir de un examen andrológico que involucra tanto la evaluación fisiológica de los órganos reproductivos del animal como el análisis del contenido seminal (Nossa & Suarez, 2011). El semen, a su vez, es evaluado extrayendo descriptores para la determinación de su calidad, tales como consistencia del fluido, concentración, viabilidad, morfología y motilidad espermática.

La motilidad espermática, que consiste en la capacidad de los espermatozoides en moverse rápidamente con un rumbo definido, es quizás el parámetro microscópico más relevante a la hora de diagnosticar la fertilidad de un espécimen (Sati & Huszar, 2015). Por lo tanto, un técnico experto en microbiología o reproducción debe observar en múltiples veces muestras seminales bajo microscopio con el fin de diagnosticar motilidad espermática. Sin embargo, este procedimiento tradicional está sujeto a la subjetividad, ya que los resultados dependen muchas veces de la experiencia personal del especialista, siendo ésta diferente entre uno y otro.

Además, las largas jornadas de trabajo pueden provocar en ellos fatiga visual, lo que por supuesto altera los resultados. Estos casos tienen un gran impacto en la baja repetitividad de los diagnósticos para una misma muestra seminal. Por consiguiente, es posible afirmar que el análisis de las características de los espermatozoides y la fertilidad de una muestra seminal llevado a cabo de manera convencional basado únicamente en la inspección visual del especialista podría llegar a ser poco confiable. En la figura 1 se resume brevemente la causa-efecto del problema.

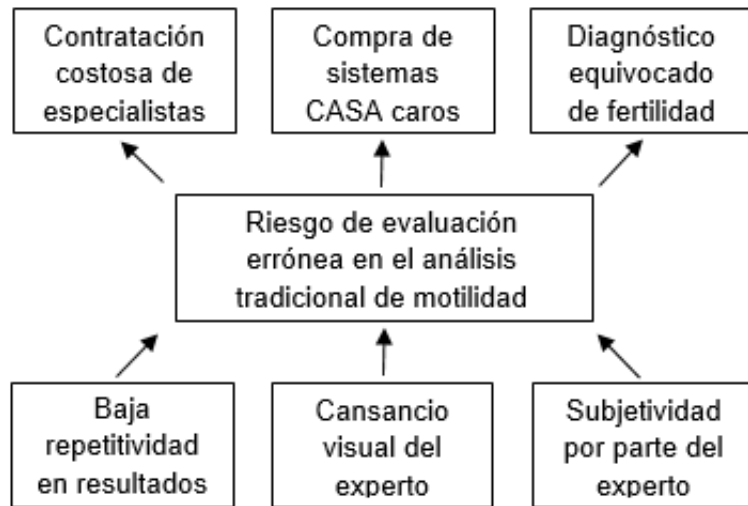


Figura 1. Árbol de problemas de la situación. Elaborada por el autor.

1.2. Objetivos del estudio

1.2.1. Objetivo general

Diseñar e implementar una herramienta computacional para determinar la motilidad individual espermática en bovinos mediante técnicas de visión artificial y procesamiento de vídeo.

1.2.2. Objetivos específicos

- Definir una técnica para la adquisición y pre-procesamiento de video de las muestras seminales.
- Segmentar y extraer los espermatozoides mediante técnicas de procesamiento de video.
- Establecer un método para detectar la trayectoria individual de cada espermatozoide sobre el video.
- Caracterizar y clasificar los espermatozoides mediante un estándar de motilidad individual espermática.
- Evaluar y comparar los resultados obtenidos en la caracterización de motilidad espermática con respecto al análisis manual por parte de un experto.

1.3. Justificación

Esta investigación tiene como trasfondo minimizar el riesgo de un diagnóstico equívoco al implementar una herramienta computacional que asista al experto en la realización del examen de semen en bovino y le permita determinar el estado de motilidad espermática de la muestra acogiéndose a protocolos preestablecidos (Amann & Waberski, 2014). Para la herramienta computacional la subjetividad es nula, puesto que ésta se basa en parámetros matemáticos medibles obtenidos de los vídeos de espermatozoides para la determinación de motilidad individual (Schleh & Leoni, 2013). También, se elimina el cansancio ocular del experto, ya que la herramienta será capaz de llevar a cabo un sinnúmero de análisis sin alteración alguna y finalmente, al tratarse de un software por computadora, superará enormemente la velocidad de evaluación que desempeñan los expertos y por lo tanto reducirá los periodos de examen de horas a solo algunos minutos. En la tabla 1 se evidencia el impacto de un sistema de análisis automático asistido por computador en el desempeño de la evaluación y en la confiabilidad de los resultados.

Tabla 1. Ventajas y desventajas del análisis manual y el análisis asistido por computador (CASA). Traducida de Schleh & Leoni, 2013.

Método	Ventajas	Desventajas
Análisis manual	- Bajo costo de adquisición	- Largos periodos de análisis - Evaluación subjetiva - Evaluación superficial
Análisis asistido por computador (CASA)	- Análisis rápido - Alta repetitividad con la misma configuración - Evaluación detallada - Alta significancia estadística debido al análisis objetivo	- Alto costo de adquisición - Necesidad de mantenimiento - Diferentes configuraciones alteran dramáticamente los resultados

Aunque hoy en día en el mercado existen diversos sistemas CASA que prestan un servicio de análisis muy completo y confiable, en Colombia no está disponible de manera abierta debido a que presenta altos requerimientos tecnológicos tanto de hardware como de software y su costo de adquisición es muy elevado. Por tanto, este proyecto se convierte en una alternativa más económica y accesible en el país, prestando un análisis confiable con una tasa de error relativamente baja. Así mismo, el método desarrollado será un aporte importante al campo de la investigación en la Universidad de Cundinamarca tanto en el área de la reproducción animal de

Zootecnia, como en el área de visión artificial de Ingeniería Electrónica. Y servirá de apoyo para futuros desarrollos de tecnologías de este tipo que se lleven a cabo en la institución.

1.4. Alcances y limitaciones

1.4.1. Alcances

- Procesamiento de videos de 640x480 pixeles y 30 fotogramas por segundo.
- Análisis de videos de espermatozoides de un espécimen de bovino.
- Evaluación de los parámetros de motilidad espermática más básicos encontrados en cualquier sistema CASA.

1.4.2. Limitaciones

- Procesamiento de videos de menos de 640x480 pixeles o con menos de 30 fotogramas por segundo.
- Análisis de videos de espermatozoides de otras especies alejadas taxonómicamente del bovino.
- Evaluación de otros aspectos del análisis seminal, tales como viabilidad, concentración y morfología.
- Análisis de videos de muestras seminales sin la correcta preparación o con demasiadas impurezas.

2. Capítulo II. Marco teórico

2.1. Fundamentos teóricos

A continuación, se exponen los conceptos teóricos principales con los cuales se sustentó el desarrollo del proyecto.

2.1.1. Ganadería bovina en Colombia

La ganadería en general es una actividad ampliamente presente en el territorio colombiano. Según (DANE, 2012), la ganadería incide en cerca del 20 % del PIB agropecuario y del 53 % del PIB pecuario. Los departamentos con mayor inventario en la región andina son Antioquia, Santander y Cundinamarca. En perspectiva nacional, la utilización del ganado se proyecta en un 16 % para la producción de leche, 48 % para la producción de carne y 36% para una doble utilidad.

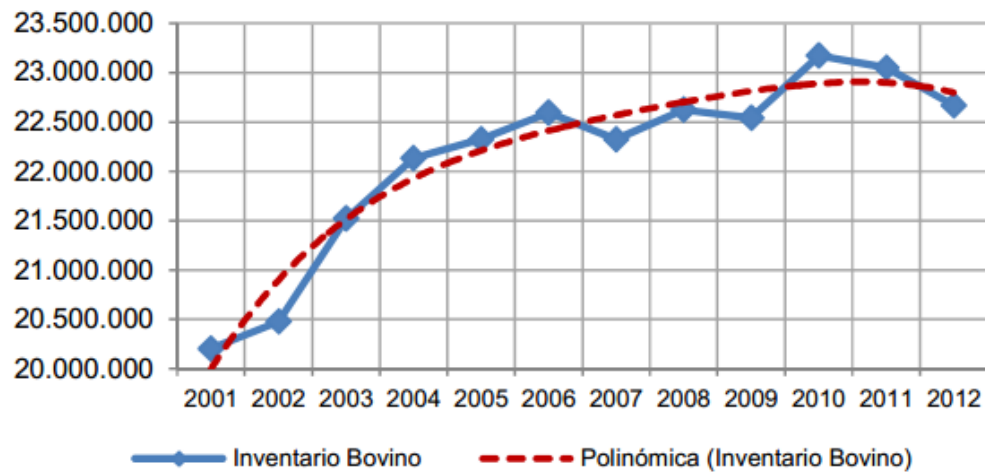


Figura 2. Inventario bovino. Tomado de FEDEGÁN, 2013.

En la figura 2 se muestra el inventario bovino desde 2001 hasta el 2012 según (FEDEGÁN, 2013) y se evidencia un conteo de hasta 22.9 millones de cabezas de ganado. A pesar de que en los últimos se ha registrado una disminución, este sector económico sigue siendo muy activo debido a la alta demanda de productos de carne y leche. En el hato ganadero se estima una población del 68 % de hembras y 32 % de machos.

La ganadería ha sido la actividad agropecuaria más importante históricamente en Colombia, y aún hoy en día es la más predominante en 27 de los 32 departamentos del país. No obstante, varios indicadores de productividad muestran un bajo nivel de implementación tecnológico en la ganadería colombiana (ABC-finkeros, 2013) quizás debido a la resistencia al cambio de los ganaderos con respecto a métodos de tecnificación y automatización de procesos.

2.1.2. Fertilidad en toros

Una baja capacidad reproductora significa una baja directa en la futura población de ganado y, por consiguiente, una pérdida de producción lechera y cárnica. Es por esto que la fertilidad tanto de macho como de hembra cobra importancia en el sector económico de la comercialización de cabezas de ganado. En el toro, la fertilidad se alcanza aproximadamente a los 16 meses de edad.

Para que un toro se considere fértil, debe contar con sus órganos reproductivos bien desarrollados y demostrar excelente libido. Además, debe producir semen de calidad, que presente un volumen de aproximadamente 5 mililitros y que contenga de entre 1 a 3 billones de células espermáticas, y que mínimo el 60 % éstas cuenten con una motilidad vigorosa (Gasque, 2008).

El examen andrológico es una evaluación que se realiza con el fin de conocer el estado o potencial reproductivo de los toros, esta evaluación se divide en varias partes: el examen físico, examen del pene, examen testicular, examen de órganos internos reproductivos y el examen de calidad seminal para la correcta determinación de la fertilidad del animal.

Las características del análisis seminal se dividen en dos grandes grupos. La primera corresponde a las características macroscópicas del semen y como su nombre lo indica abarca aquellas observaciones a simple vista de la calidad del semen extraído para evaluación (Angelino, 2009). El segundo grupo corresponde a características microscópicas, es decir, se necesitan instrumentos de observación como microscopios ópticos o electrónicos.

Parámetros macroscópicos

- Volumen: medido mediante graduación en mililitros.
- Color: tonalidades cremosas, lechosas o translúcidas.
- pH: nivel de acidez.

Parámetros microscópicos

- Viabilidad: porcentaje de espermatozoides vivos.
- Concentración: número de espermatozoides por centímetro cúbico.
- Morfología: forma y estado del acrosoma del espermatozoide
- Motilidad masal: estimación de la vigorosidad del movimiento de grandes grupos de espermatozoides.
- Motilidad individual: porcentaje de espermatozoides con desplazamiento rápido y progresivo.

2.1.3. Motilidad espermática

A su vez, el parámetro de motilidad se divide en dos grupos. La motilidad masal, que hace referencia al comportamiento colectivo y de fluidez del movimiento de la gran mayoría de los espermatozoides dentro de una muestra seminal y se analiza en un campo de acción grande en un microscopio. La motilidad individual, como su nombre lo indica, abarca únicamente a cada espermatozoide, aunque se analiza en grupos de algunas células espermáticas y verifica parámetros más detallados como su velocidad de desplazamiento y su tipo progresividad. Por lo tanto, el campo de acción requerido para el análisis de motilidad individual es pequeño.

El análisis de motilidad se realiza en muestras de semen puro (motilidad masal) y semen diluido (motilidad individual). La motilidad masal es la prueba más simple que se realiza en semen puro y cuando el semen se diluye (fresco o congelado), la motilidad individual es el mejor indicador de la vitalidad del espermatozoide (Angelino, 2009). Las categorías subjetivas de motilidad masal e individual se aprecian en las tablas 2 y 3 respectivamente.

Tabla 2. Clasificación de motilidad masal. Tomado de Hafez, 1989.

Muy buena	Movimiento en ondas vigorosas y en remolinos rápidos	80-100 %
Buena	Remolinos y ondas más lentas	60-79 %
Regular	Sin remolinos, pero con ondas generalizadas	40-59 %
Mala	Escasa o nula motilidad	< 40%

Tabla 3. Clasificación de motilidad individual. Tomado de Hafez, 1989.

Muy buena	80-100 % con motilidad progresiva
Buena	60-79 % con motilidad progresiva
Regular	40-59 % con motilidad progresiva
Mala	< 40 % con motilidad progresiva

2.1.4. Sistemas comerciales CASA

Los sistemas CASA (Computer Assisted Sperm Analysis o Análisis de Esperma Asistido por Computadora) son sistemas automáticos o semiautomáticos diseñados para proveer información significativa y precisa acerca de los valores descriptivos de una muestra de semen, tales como la viabilidad, concentración, morfología y dinámica de movimiento para luego llevar a cabo análisis estadísticos de valores objetivos (Lu, et al., 2013). Estos sistemas se basan en el procesamiento digital de una serie de imágenes consecutivas de espermatozoides capturadas gracias a una cámara digital especializada en la adquisición de imágenes en microscopios. Una computadora es la encargada de realizar el procesamiento tanto de imágenes como estadístico.

Aunque los sistemas CASA son normalmente costosos debido a sus implicaciones tecnológicas, proveen un soporte confiable para cualquier laboratorio especializado en exámenes andrológicos. Además, tienen la capacidad de llevar numerosos análisis de semen en periodos de tiempo muy cortos, con resultados de alta repetitividad y alta significancia, superando así el desempeño de un operador humano. A pesar de que estos sistemas son automáticos en el análisis de espermatozoides, generalmente es necesaria la presencia de un experto para realizar la preparación de la muestra y posterior configuración del sistema CASA.

Desde el punto de vista de la motilidad espermática individual, los sistemas CASA entregan al usuario variables de salida que representan el comportamiento cinético de los espermatozoides analizados. Estas variables de salida o parámetros espermáticos de motilidad son comunes en la gran mayoría de sistemas CASA (Boyers, et al., 1989). En la figura 3 se representan visualmente los parámetros de motilidad analizados por los sistemas CASA. La trayectoria real se muestra en verde, la trayectoria promedio se muestra en azul y el desplazamiento lineal se muestra en rojo. En la figura 4 se aprecian los resultados visuales que generalmente producen estos sistemas en el interfaz de usuario.

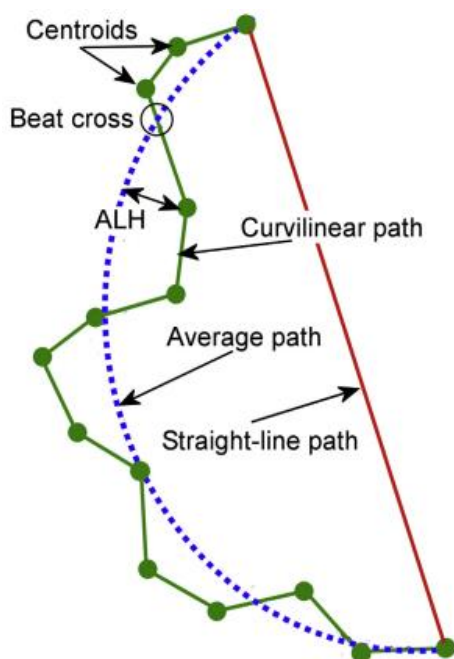


Figura 3. Representación visual de los parámetros de motilidad más comunes en sistemas CASA.
Tomado de Amann & Waberski, 2014.

- **Velocidad curvilínea (VCL - $\mu\text{m/s}$):** trayectoria real recorrida por el espermatozoide por unidad de tiempo.
- **Velocidad promedio (VAP - $\mu\text{m/s}$):** trayectoria media o suavizada recorrida por el espermatozoide por unidad de tiempo.
- **Velocidad rectilínea (VSL - $\mu\text{m/s}$):** distancia entre el punto inicial y el punto final del desplazamiento del espermatozoide por unidad de tiempo.
- **Índice de linealidad (LIN - %):** relación porcentual entre VSL y VCL.
- **Índice de rectitud (STR - %):** relación porcentual entre VSL y VAP.
- **Índice de oscilación (WOB - %):** relación porcentual entre VAP y VCL.
- **Porcentaje de motilidad total (MT - %):** relación entre los espermatozoides con motilidad respecto al número total.
- **Porcentaje de motilidad progresiva (MP - %):** relación entre aquellos espermatozoides con motilidad progresiva respecto al número total.
 - **Grado A:** espermatozoides con motilidad progresiva rápida.
 - **Grado B:** espermatozoides con motilidad progresiva lenta.
 - **Grado C:** espermatozoides con motilidad no progresiva.
 - **Grado D:** espermatozoides sin motilidad o estáticos.

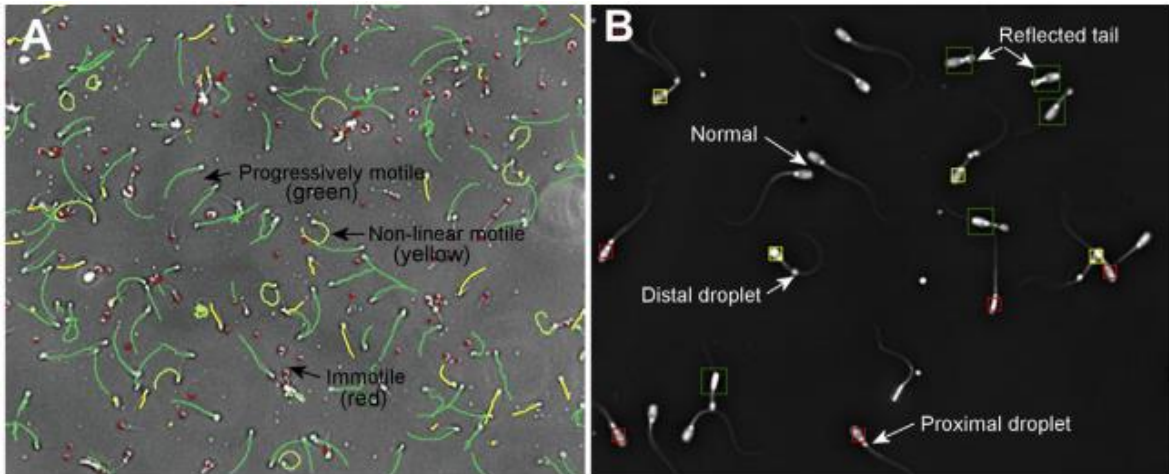


Figura 4. Interfaz de resultados en sistemas CASA. (A) Análisis de motilidad (B) Análisis de morfología. Adaptado de Amann & Waberski, 2014.

2.1.5. Visión artificial

La visión artificial o también conocida como visión por computadora es una rama de la ciencia de la inteligencia artificial que mediante la utilización de técnicas adecuadas permite la obtención, procesamiento y análisis de información a través de imágenes digitales. El impacto de la visión artificial sobre la labor del hombre se resume por ejemplo en la automatización de tareas de inspección repetitivas, realizar control de calidad sin contacto directo de gran cantidad de productos a gran velocidad y la reducción de periodos de ciclo de procesos industriales (Salgado, 2007). Las principales aplicaciones de la visión artificial son:

- Detección e identificación de objetos.
- Determinación de ubicación de objetos en el espacio.
- Determinación de relaciones espaciales entre objetos.
- Medición de coordenadas, distancias y ángulos.
- Clasificación y categorización.

Imagen y video digital

Una imagen es una agrupación de puntos en forma de matriz bidimensional que representan un espacio visual. Dichos puntos se denominan píxeles y comúnmente contienen tres componentes de color RGB (Rojo, Verde y Azul) que permiten representar cualquier color del espectro visible. Un vídeo, por su parte, es una secuencia de imágenes, cuadros o fotogramas que se desplazan a través del

tiempo. Los valores matemáticos de las imágenes no son más que valores binarios en cada campo de una matriz. Por lo tanto, están sujetos a la aplicación de operaciones matemáticas con el fin de procesarlas (Escalante, 2006).

88	82	84	88	85	83	80	93	102
88	80	78	80	80	78	73	94	100
85	79	80	78	77	74	65	91	99
38	35	40	35	39	74	77	70	65
20	25	23	28	37	69	64	60	57
22	26	22	28	40	65	64	59	34
24	28	24	30	37	60	58	56	66
21	22	23	27	38	60	67	65	67
23	22	22	25	38	59	64	67	66

Figura 5. Representación matricial de una imagen digital. Elaborada por el autor.

Técnicas de pre-procesamiento y filtrado

Las imágenes comúnmente presentan imperfecciones o efectos indeseados generalmente provocados por la falta de condiciones óptimas a la hora de tomarlas. Los filtros son algoritmos que permiten realizar un primer proceso a la imagen con el fin de eliminar o minimizar dichos efectos, tales como variaciones bruscas de nivel de pixel. Los filtros modifican el valor de cada pixel reduciendo su diferencia con respecto a sus pixeles vecinos. Cabe mencionar que es necesario convertir una imagen de tipo RGB a escala de grises para filtrarla.



Figura 6. (A) Imagen con ruido (B) Imagen filtrada. Tomado de *mathworks.com*.

Técnicas de segmentación

La segmentación subdivide una imagen en sus partes constituyentes u objetos, con el fin de separar las partes de interés del resto de la imagen (La Serna & Román, 2009). Existen diversas técnicas para la segmentación de imágenes, sin embargo, existen técnicas muy comunes.

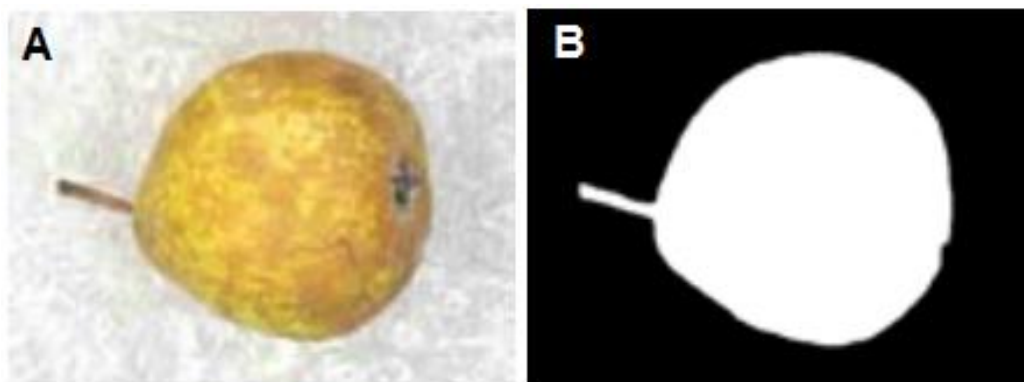


Figura 7. (A) Imagen original (B) Imagen segmentada. Tomado de La Serna & Román, 2009.

La detección de bordes, que consiste en la detección de límites o contornos de un objeto en primer plano contrastado con el fondo. La umbralización, es aquel método que discrimina el nivel de magnitud de los píxeles con respecto a un valor de referencia para así clasificarlo como objeto de interés o fondo de imagen. Con respecto a videos, el método de diferenciación de fotogramas se basa en la comparación del cambio de valor de los píxeles a través de una secuencia continua de imágenes, detectando así un objeto en movimiento. Como resultado, la segmentación entrega una imagen binaria del objeto detectado. Como complemento a estos procesos, frecuentemente se emplean operadores morfológicos, que tienen la finalidad de moldear o adecuar los objetos extraídos para un posterior análisis.

Análisis de componentes conectados

En una imagen binaria, los componentes conectados son aquellas agrupaciones de píxeles positivos que representan la región del objeto de interés. Mediante un análisis de estos píxeles conectados, es posible calcular valores geométricos que describan al objetos (Gonzalez & Woods, 2008). Estos valores generalmente corresponden a áreas, perímetros, centroides, ángulos, excentricidades.

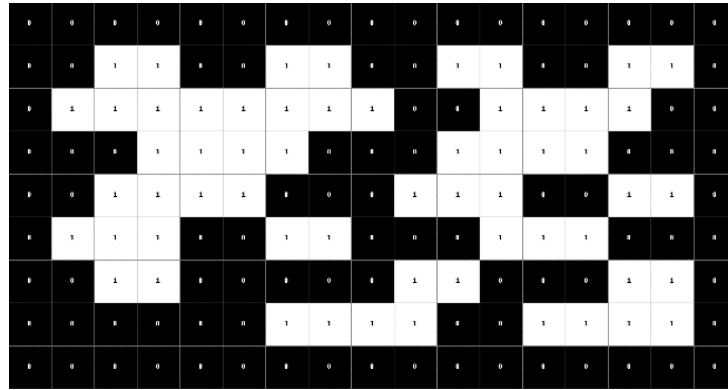


Figura 8. Representación de una imagen binaria. Tomado de *en.wikipedia.org*, 2016.

2.2. Estado del arte

A lo largo de la última década se han desarrollado múltiples sistemas computacionales para el análisis de parámetros espermáticos. Los avances tecnológicos en cuanto a microscopios, cámaras digitales, computadoras y algoritmos han impulsado el desarrollo de nuevas metodologías y ha incentivado su utilización como una alternativa efectiva frente a los métodos convencionales de evaluación seminal. A continuación, se presenta una breve recopilación de los avances significativos en el campo del desarrollo de sistemas para el análisis espermático asistido por computadora (CASA).

2.2.1. Algoritmos de análisis espermático

En (Pascual-Gaspar, et al., 2008) los autores desarrollan una metodología que consiste en el procesamiento digital de un video en formato .avi con una velocidad de 25 cuadros por segundo de una muestra microscópica con magnificación de 400X. Se procede a transformar la imagen de la secuencia a escala de grises y posteriormente se segmenta para obtener una imagen binaria. Las técnicas de segmentación se apoyan de diferencias de color, áreas, perímetro y excentricidad de los objetos para eliminar impurezas tales como células epiteliales, leucocitos, etc.

La última fase del proceso global mostrado en la figura 9 se compone por el módulo de seguimiento de espermatozoides que en este caso se lleva a cabo mediante una técnica llamada segmentación cinemática. Es necesario calcular el desplazamiento máximo de un espermatozoide entre un cuadro y otro para delimitar una 'ventana de rastreo' que se utiliza para evaluar cada objetivo. Los autores encontraron

concordancia de histogramas entre diferentes experimentos con videos distintos, presentado errores no mayores al 2%.

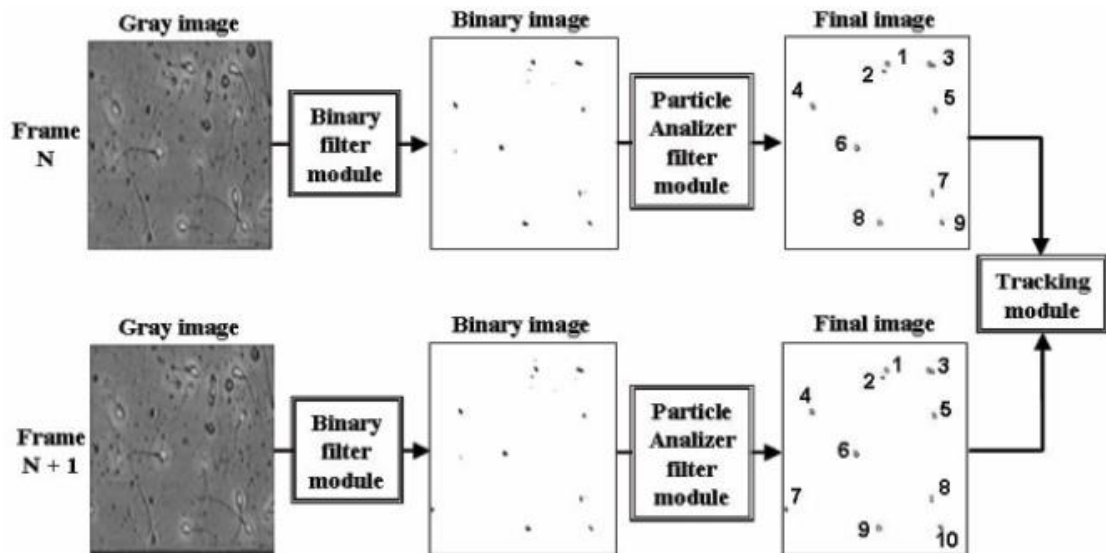


Figura 9. Metodología propuesta por los autores. Tomado de J. Pascual et al., 2008.

La técnica presentada en (Abbiramy, et al., 2010) consiste en un algoritmo que, a partir de tramas de imágenes consecutivas o videos de muestras microscópicas de semen, busca detectar, segmentar, procesar y hacer seguimiento del tipo de motilidad en los espermatozoides presentes con el fin de diagnosticar astenozoospermia, es decir, motilidad deficiente o nula de los espermatozoides. Esto permitiendo así una aproximación objetiva en el diagnóstico de motilidad espermática humana y por consiguiente en su fertilidad. El algoritmo, según se describe en la figura 10, comienza con una etapa de pre-procesamiento donde se convierte la imagen a grises, se filtra para eliminar ruido y, mediante la extracción de propiedades, se determinan centroides, áreas, y punto inicial, para luego ser almacenadas en una base de datos para cada cuerpo en cada cuadro de video.

Con estas propiedades es posible hacer una comparación entre el cuadro actual y el previo, para así identificar un mismo cuerpo móvil entre diferentes tramos del video. Por último, los autores describen una etapa de rastreo o seguimiento de espermatozoides, que consiste en dividir cada cuadro en secciones más pequeñas que son evaluadas cada una. En cada sección se comparan cuerpos de iguales propiedades y se determina su cambio de posición o desplazamiento. Este valor de distancia es almacenado conforme avanza cada cuadro del video. La velocidad de

cada espermatozoide se halla al multiplicar la media de desplazamiento por la velocidad de cuadros por segundo.

De esta manera, los autores clasifican la motilidad de cada espermatozoide en: motilidad progresiva rápida ($> 25 \mu\text{m/s}$), motilidad progresiva lenta, motilidad no progresiva ($< 5 \mu\text{m/s}$) y sin motilidad. Se reportan porcentajes de precisión de hasta 93% y por lo tanto es viable en el diagnóstico de fertilidad masculina.

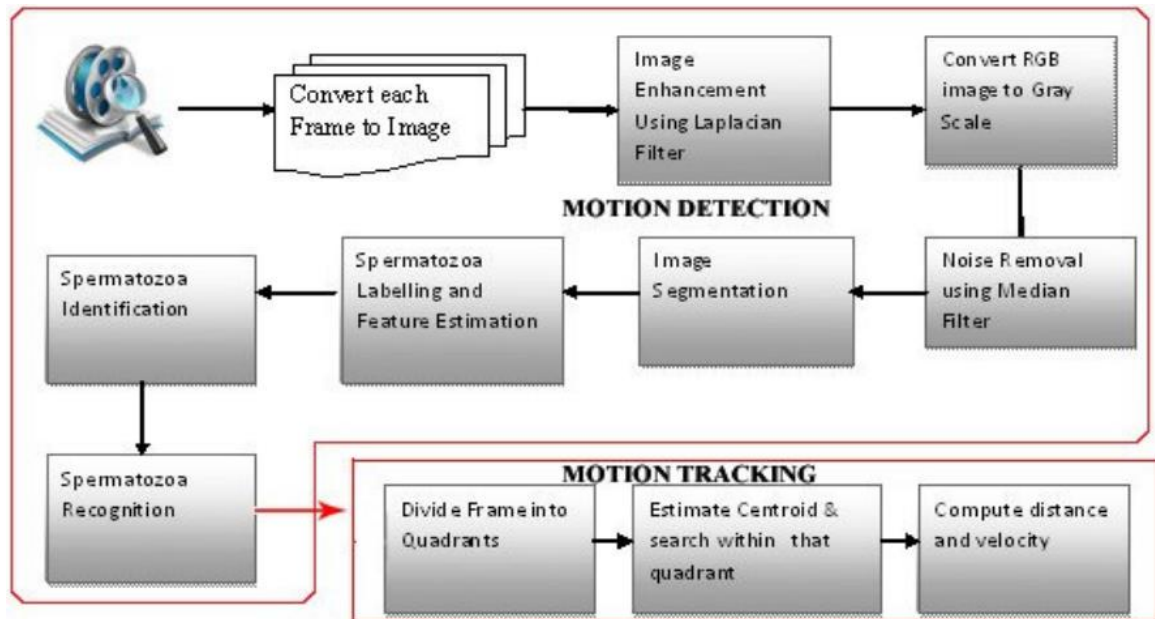


Figura 10. Metodología propuesta por los autores. Tomado de V. Abbiramy et al., 2010.

Los autores en (Ravanfar & Moradi, 2011) presentan en este proyecto una solución al inconveniente de la colisión entre espermatozoides en imágenes de bajo contraste que dificultan la correcta identificación de trayectorias de estos en los algoritmos de procesamiento de imágenes para la evaluación de motilidad espermática y el diagnóstico de fertilidad. La metodología propuesta funciona mediante la implementación de un algoritmo Watershed o de transformación divisoria y un filtro de partículas que mejora considerablemente la detección individual de los espermatozoides cuando se aglomeran.

Seguido a la segmentación de la imagen, se aplica un filtro de partículas, que consiste en un procedimiento estadístico en el cual se estima una distribución de probabilidades por medio de una red bayesiana para así crear un modelo de espacio de estados. Por medio de este espacio de estados se determina una distancia

euclidiana para cada partícula con el fin descartar elementos no pertenecientes a espermatozoides como por ejemplo células de sangre o restos citoplasmáticos. Luego de esto, la fase final de la segmentación se lleva a cabo con la aplicación del algoritmo Watershed que, mediante los conceptos de región profunda, cuenca hídrica y líneas divisorias (analogía de esta teoría aplicada al histograma de una imagen), detecta los bordes de dos o más cuerpos que colisionan para así separarlos de manera efectiva y dar paso a la etapa de rastreo, como se evidencia en la figura 11. Los autores encontraron que el error de cuadrado medio para la correcta detección fue de 2.72.



Figura 11. Fotografías 1-4. (a) Transformada distancia euclidiana. (b) Algoritmo Watershed con centroide detectado. Tomado de M. Ravanfar et al., 2011.

Otros autores en (Li, et al., 2012) se propone una alternativa a los algoritmos usados frecuentemente, que pretende eliminar las desventajas de la segmentación por niveles de grises mediante la adición de etapas de procesamiento correspondientes a la segmentación con criterios de modelamiento gaussiano e implementación de operadores morfológicos con el fin de aportar al mejoramiento del análisis tanto de concentración y motilidad espermática. La estructura completa del diseño se elabora bajo la arquitectura lógica del software OpenCV y se constituye en cuatro fases. La primera fase aplica técnicas de pre-procesamiento tales como un filtro gaussiano. La segunda fase realiza una segmentación del primer plano por medio de la aplicación de un modelo de fondo gaussiano.

De esta manera, se ignora el fondo y se extraen aquellos objetos pertenecientes a espermatozoides en movimiento. No obstante, la imagen segmentada resultante presenta imperfecciones como regiones de píxeles divididas o también dos regiones o segmentos independiente combinados erróneamente. Para corregir esto, en la tercera fase se aplican operadores morfológicos con el objetivo de moldear los

segmentos y eliminar imperfecciones para hacerlos coherentes con los espermatozoides de la imagen original. Las técnicas para esta fase son cierre, apertura y dilatación, respectivamente. La cuarta y última fase aplica un detector de contorno que etiqueta a cada segmento como espermatozoide individual en la imagen original. Según la tabla 4, los porcentajes de precisión del algoritmo propuesto oscilan alrededor del 96%.

Tabla 4. Comparación entre espermatozoides existentes y detectados, y porcentajes de precisión. Tomado de Q. Li et al., 2012.

Frame No	Actual No. of Objects	No. of Objects Detected	Accuracy %	Error Rate %
1	88	84	95.5	4.5
2	86	81	94.2	5.8
3	78	75	96.2	3.8
4	77	72	93.5	6.5
5	84	82	97.6	2.4
6	82	79	96.3	3.7
7	83	77	92.8	7.2
8	80	79	98.8	1.3
9	79	75	94.9	5.1
10	81	80	98.8	1.2

Por otra parte, en (Roa, et al., 2012) se presenta una herramienta computacional que detecta automáticamente viabilidad y concentración espermática a partir de una imagen en espacio RGB. La primera fase del algoritmo comienza por la adquisición de las imágenes microscópicas de las muestras seminales bajo la aplicación de un colorante vital Eosina-Y al 0.5%, lo que permite teñir de color rojizo las células espermáticas muertas, y así diferenciarlas de las vivas mediante color. La segunda fase consiste en el pre-proceso de las imágenes, en el cual se aplican un filtro adaptativo de Wiener, con el fin de reducir el ruido y suavizar las altas variaciones de color de los píxeles, de esta manera preparándola para el procesamiento.

La tercera fase consiste en la extracción de las cabezas de espermatozoides mediante la aplicación del análisis discriminante de Fisher y N-conectados, en donde se segmentan solo los espermatozoides, descartando los demás objetos incluido el fondo. Seguido en la cuarta fase se realiza un proceso de clasificación de dichos espermatozoides mediante “clustering” o agrupamiento en dos clases: vivos y muertos, específicamente mediante la técnica de K-medias. Y finalmente, se valida el algoritmo realizando una comparación entre éste mismo con análisis

llevados a cabo de forma manual por expertos de laboratorio en muestras seminales de 14 individuos, obteniendo así una correlación medible. En la tabla 5 los autores muestran que como resultado la metodología propuesta alcanza porcentajes de precisión de hasta 87.8% en la evaluación de viabilidad y concentración.

Tabla 5. Tabla de confusión de viabilidad en 14 individuos. Tomado de E. Roa et al., 2012.

	Vivos(algoritmo)	Muertos(algoritmo)
Vivos(real)	2918	355
Muertos(real)	207	1159
Exactitud		0,8788
Sensibilidad		0,934
Especificidad		0,76

Una novedosa propuesta (Yang, et al., 2014) fue desarrollada con el objetivo de detectar la cabeza del espermatozoide y su flagelo. La metodología se compone de una fase de detección de la cabeza, que consiste en la extracción de valores de esta tales como, coordenadas del centroide, eje mayor, eje menor y ángulo, y se consigue mediante la implementación de una versión mejorada del algoritmo Multiple Birth and Cut (MBC). Luego, una fase de rastreo de cabeza emplea una comparación entre pares de cuadros de video. Entre estos pares se genera una sub-imagen con el fin de servir de referencia para determinar desplazamiento y cambio de ángulo. Por último, La fase de seguimiento de flagelo parte del conocimiento de la posición de la cabeza del espermatozoide. Se realiza una observación de la zona oscura que representa la cola y se determina la convergencia o divergencia en valores de pixeles, es decir, en qué sentido existe un cambio de magnitud de pixel. Con esto, el algoritmo modela el flagelo desde la cabeza hasta que ya no encuentre cambios de pixeles toda la longitud del flagelo.

Este modelo se compone de una sucesión de círculos cuya doble sumatoria de radios determina la longitud, y cada círculo conserva un valor de ángulo que dicta la orientación. Mediante el gradiente del radio de los círculos se determina el borde del flagelo. Un algoritmo estadístico basado en el método de Monte Carlo por cadenas de Markov se emplea para realizar una estimación de la secuencia de círculos del flagelo. En la figura 12 se aprecia que la metodología propuesta se aproxima de manera efectiva a la forma real del flagelo y lo identifica en cada cuadro por segundo.

Para analizar los resultados se diseñó una comparación entre los datos generados por la computadora sobre la posición actual del flagelo y una anotación dada por una observación física sobre la imagen. Se evalúan dos métricas: el porcentaje de píxeles en la anotación alejados a más de x distancia de los datos generados (MDR), y el porcentaje de píxeles en los datos generados alejados a más de x distancia de la anotación (FDR). Para una distancia de 3 píxeles resultó que MDR y FDR equivalían a 6.12% y 1.78% respectivamente.

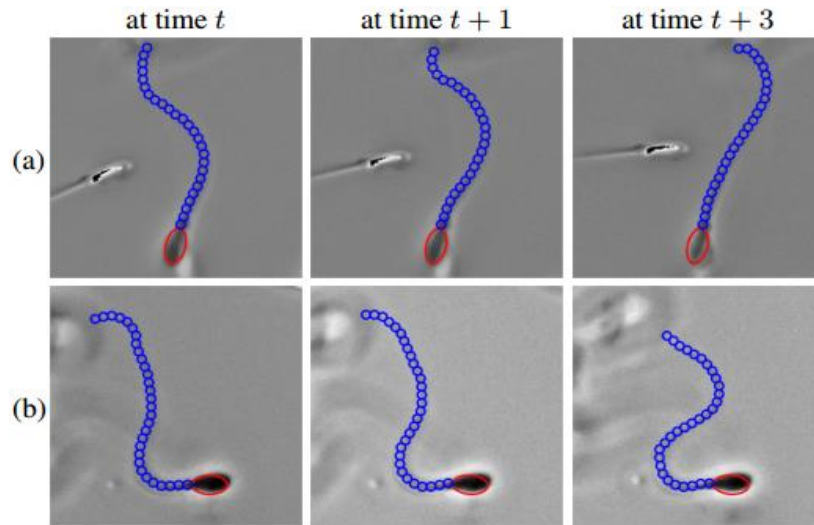


Figura 12. Detección de flagelo en sucesión de 3 fotografías. (a) En condiciones normales. (b) Bajo beta-mercaptoethanol (bME). Tomado de H. Yang et al., 2014.

El diseño presentado en (Elsayed, et al., 2015) tiene la capacidad de llevar a cabo un análisis de motilidad en muestras de semen dentro de dispositivos de microfluidos. Los dispositivos de microfluidos son aquellos que se utilizan mayormente en los procesos de fertilización *in-vitro* e inseminación artificial para seleccionar y clasificar determinados espermatozoides. Esto se facilita gracias a la estructura del dispositivo, que consta de canales controlados de microfluidos (en el orden de micrómetros). Sin embargo, estos microfluidos están en constante movimiento al igual que los espermatozoides. Por lo tanto, los autores presentan un algoritmo de licencia abierta para permitir el análisis de motilidad sobre microfluidos tal como lo haría un sistema CASA convencional.

El método propuesto consiste primero en realizar un mejoramiento en el contraste. Ya que los dispositivos microfluidos presentan en sus canales una altura relativamente mayor a su ancho, el movimiento vertical de los espermatozoides hace que su imagen se desenfoque y pierda contraste. Entonces, se implementó

una ecualización por histograma adaptativa para incrementar el contraste. Seguido a esto se aplica una conversión a escala de grises. Después, es extraído el fondo mediante la diferenciación consecutiva de entre 10 y 30 fotogramas, dejando solamente los espermatozoides para luego hacer binaria la imagen. Luego, operadores morfológicos se utilizan para eliminar el ruido en la imagen binaria y abrir paso al rastreo de partículas. Este rastreo se lleva a cabo al hallar la correspondencia de una partícula entre dos o más fotogramas. Para esto, los autores implementan el Algoritmo Greedy, que en términos sencillos realiza un análisis de todas las partículas existentes y determina cual es la más probable a ser la posición actual de una previamente determinada. Se reporta que la metodología propuesta identifica las trayectorias espermáticas en el microfluído, como se muestra en la figura 13 y alcanza un rango de precisión de entre 75-100%.

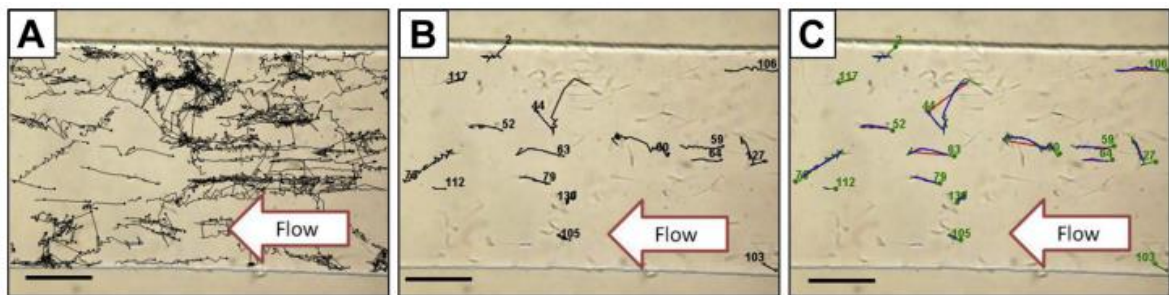


Figura 13. (a) Todas las trayectorias detectadas. (b) Algunas seleccionadas. (c) identificación de parámetros de motilidad. Tomado de M. Elsayed et al., 2015.

2.2.2. Evaluación de sistemas CASA

En este estudio (Tomlinson, et al., 2010) se pretende determinar la precisión de un nuevo sistema de análisis espermático asistido por computadora mediante la comparación de sus resultados con los de métodos manuales. La metodología inicia con una etapa de detección automática de cabezas. Análisis de píxeles y umbrales permiten determinar las cabezas espermáticas del fondo. No obstante, se deben aplicar operaciones morfológicas como erosión, dilatación y llenado de huecos o regiones para aproximar la imagen segmentada a la original. Regiones más grandes o más pequeñas del promedio de tamaño de cabezas son descartadas. Sin embargo, en este punto del proceso debe haber intervención de un usuario, quien debe descartar o incluir espermatozoides que fueron o no detectados erróneamente. Esto se debe a que los restos de material en el semen la mayoría de las veces se confunde fácilmente con esperma real, por lo que el software no puede identificar una diferencia y por consiguiente es necesaria la intervención humana.

El método de Monte Carlo por cadenas de Markov fue implementado para hacer seguimiento a cada espermatozoide a lo largo de 30 fotogramas en un segundo.

Este método permite que la detección de trayectorias se realice correctamente incluso cuando dos objetivos se interponen uno sobre el otro. Luego de obtener trayectorias, se categorizan cada una mediante la evaluación de sus velocidades. Las categorías son las convencionales. Grado A (>25 $\mu\text{m/s}$), grado B (5-25 $\mu\text{m/s}$), grado C (<5 $\mu\text{m/s}$) y grado D (inmóvil). Y los parámetros de motilidad son los tradicionales: VCL, VSL, VAP, etc. Para una muestra de semen preparada, es decir, diluida para una correcta visualización, la tasa de error en la detección de trayectos en grado A fue del 5.88% y en grado c aumenta hasta el 7.8% tal como se evidencia en la tabla 6.

Tabla 6. Tasas de error en la detección de trayectorias para los diferentes grados de clasificación tanto en semen normal como diluido. Tomado de M. Tomlinson et al., 2010.

	Grade a	Grade b	Grade c	Grade d	Overall
Sperm in semen					
Sperm evaluated	800	294	344	865	2,303
Tracking errors ^a	57	50	30	46	183
Error rate	7.1	17.0	8.7	5.3	7.9
Gradient prepared sperm					
Sperm evaluated	711	238	205	479	1,633
Tracking errors ^a	41	18	16	21	96
Error rate	5.8	7.6	7.8	4.4	5.88 ^a

Un estudio interesante presentado en (Castellini, et al., 2011) demuestra el impacto de la variación de la frecuencia de video en resultados del análisis de motilidad en sistemas CASA. La metodología propuesta consiste en la preparación de muestras de semen de humano, carnero, toro y conejo diluidos cuya concentración fue de 30 millones de espermatozoides por mililitro a 37°C. Cada muestra de esperma fue analizada a las frecuencias o velocidades de fotograma (FR) de 12, 25, 50, 100 y 200 Hz por un segundo. Mediante la implementación del sistema CASA los autores observaron parámetros de motilidad tales como VCL, VSL, VAP, LIN, ALH y BCF con los cuales se clasifica la motilidad espermática en grados A, B y C. En adición a esto, se realiza una comparación entre las cuatro especies al observar los resultados del sistema con 7 muestras para cada una.

Como resultado se halló que, para cualquier valor de FR, los espermatozoides de conejo presentaron la mayor velocidad de curvilínea (220.4 $\mu\text{m/s}$), seguido del carnero, el toro y por último el humano cuya velocidad fue de 97.5 $\mu\text{m/s}$. No

obstante, los espermatozoides humanos presentaron la mejor linealidad (VSL/VCL) con un valor de 64.8% en el mismo orden, pero en sentido contrario hallándose los espermatozoides de conejo con solo 26.5%. Por otra parte, los autores encontraron que independientemente de la especie en cuestión, los fotogramas por segundo no afectan significativamente el valor de porcentaje de motilidad, sin embargo, los demás parámetros variaron considerablemente, tal como lo evidencia la tabla 7. La velocidad curvilínea fue la más afectada, incrementando prácticamente 4 veces. Cabe destacar que esta significativa variación fue menos pronunciada para muestras seminales de humano y toro en comparación con carnero y conejo. Se encontró también que, el valor promedio de FR en adquisición de video de esperma humano cuya variación de parámetro de velocidad fuese mínima corresponde a 97 Hz. Los espermatozoides categorizados como grado C, es decir, que no presentan motilidad progresiva, son los más susceptibles a la variación de FR. Por tanto, se requieren frecuencias muy altas para hacer una parametrización más precisa de estos espermatozoides. El grado menos afectado es el B (progresivos lentos) dado que su velocidad y linealidad hacen que una frecuencia de video baja sea suficiente.

Tabla 7. Variación de parámetros de motilidad espermática según velocidad de fotogramas. Tomado de C. Castellini et al., 2011.

	FR (Hz)					Pooled SE
	12	25	50	100	200	
Incorrect sperm track reconstruction (%)	1.1	1.5	1.5	2.3	3.0	0.7
Motility (%)	76.5	75.2	76.6	74.3	73.1	8.2
VCL ($\mu\text{m/s}$)	58.5 ^a	88.1 ^b	156.7 ^c	206.8 ^d	217.2 ^d	56.5
VSL ($\mu\text{m/s}$)	47.3 ^a	54.4 ^b	53.6 ^b	56.1 ^b	57.6 ^b	20.9
VAP ($\mu\text{m/s}$)	42.3 ^a	69.1 ^b	92.3 ^c	126.8 ^d	137.7 ^e	50.1
Linearity (VSL/VCL %)	80.5 ^a	61.9 ^b	34.9 ^c	27.1 ^d	26.5 ^d	23.1
ALH (μm)	3.30 ^a	2.64 ^b	2.92 ^c	2.83 ^b	1.83 ^d	1.51
BCF (n)	2.6 ^a	6.1 ^b	15.2 ^c	23.5 ^d	29.6 ^e	10.2
WHO category a	65.0	51.2	24.8 ^b	9.9	9.0	3.4
WHO category b	25.0	14.5	4.6	1.0	0.6	1.2
WHO category c	10.0	34.3	70.6	89.1	90.4	5.6

Otros autores proponen en (Boryshpolets, et al., 2013) un análisis comparativo de algunas herramientas computacionales de análisis asistido de esperma con el fin de evidenciar que los resultados de motilidad de los diferentes análisis para una misma muestra pueden variar significativamente. Se utilizan muestras de semen de tres especies taxonómicamente separadas de peces: esturión esterlete, carpa común y trucha arcoíris, debido a que sus parámetros de motilidad son muy diferenciados, lo que permite un rango amplio de estudio. Y se obtuvieron videos tanto de 25 como de 50 cuadros por segundo, con duración de 12 segundos con el propósito de evaluar el impacto de la frecuencia en los resultados. Asimismo, se utilizan tres

herramientas computacionales para llevar a cabo el análisis comparativo: CRISMAS CASA, Hobson Sperm Tracker y Image J CASA plug-in.

Como resultado, se obtiene que las pruebas realizadas para videos de 50 Hz entregan sistemáticamente valores altos de VCL, VAP y VSL para las tres especies de peces comparados a un análisis similar de 25 Hz. Al pasar de 25 a 50 Hz, se evidenció un aumento de entre 9 y 15 $\mu\text{m/s}$ para las velocidades, entre ellas VCL. Según la figura 14. El sistema Hobson Sperm Tracker se diferenció ampliamente de sus similares en cuanto a la velocidad curvilínea de las especies carpa y trucha bajo muestras de 25 Hz. Sin embargo, para 50 Hz las pruebas se mostraron más consistentes para esta frecuencia, aunque cabe resaltar que se presenta un aumento del valor de VCL para cada sistema y especie. A su vez, los autores demuestran que los valores de velocidad son muy dependientes del periodo de tiempo del video usado en los experimentos.

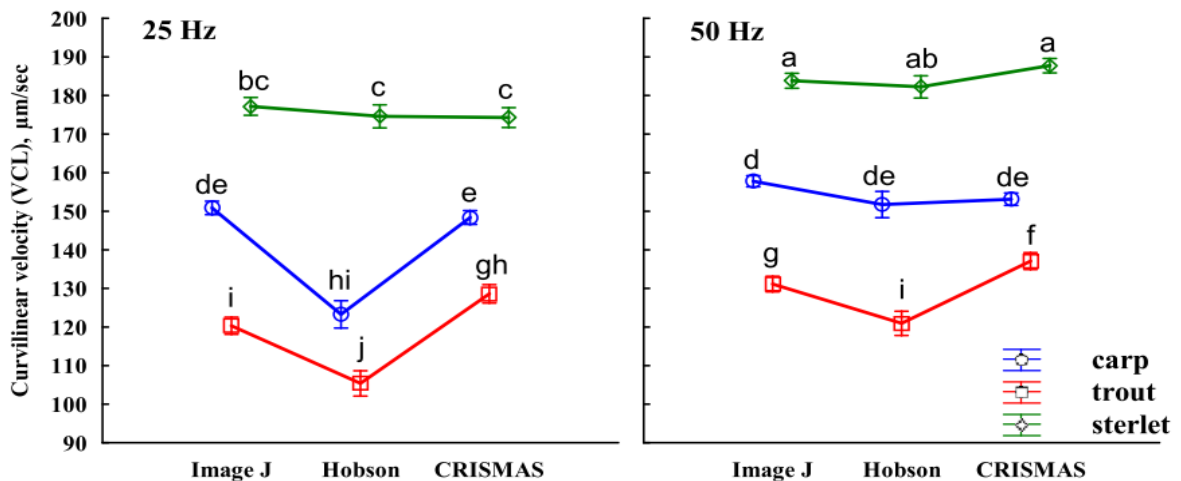


Figura 14. Velocidad curvilínea para cada especie y cada sistema a 25 y 50 Hz en 12 segundos de video. Tomado de S. Boryshpolets et al., 2013.

El objetivo del estudio en (Palacín, et al., 2013) es contribuir a una estandarización del análisis espermático CASA en caprinos evaluando el efecto de diferentes láminas, volumen de gota de la muestra y el campo de acción. Debido a que todos estos detalles, además de factores como técnica de recolección, concentración de muestra, ajustes de software, etc., pueden influenciar notablemente el resultado de los análisis de motilidad de células espermáticas. Los autores implementan un sistema CASA con un microscopio con 100x de magnificación y miden parámetros de motilidad convencionales.

El método propuesto por los autores consiste en tres experimentos. El primero compara dos gotas con volúmenes de 5 y 10 μl sobre una lámina con cubreobjetos y sobre una cámara de Makler. Para la lámina se observan 5 campos dentro de los límites de la muestra. El segundo método adiciona una cámara ISAS para la comparación, y se observan los efectos del número de espermatozoides por campo de acción en los resultados de motilidad del sistema CASA. El tercer método, a su vez, consiste en mezclar muestras frescas de semen caprino con muestras de espermatozoides muertos (forzadamente) para así analizarlos con el sistema CASA y determinar cómo afecta esta incidencia al número de espermatozoides no motiles.

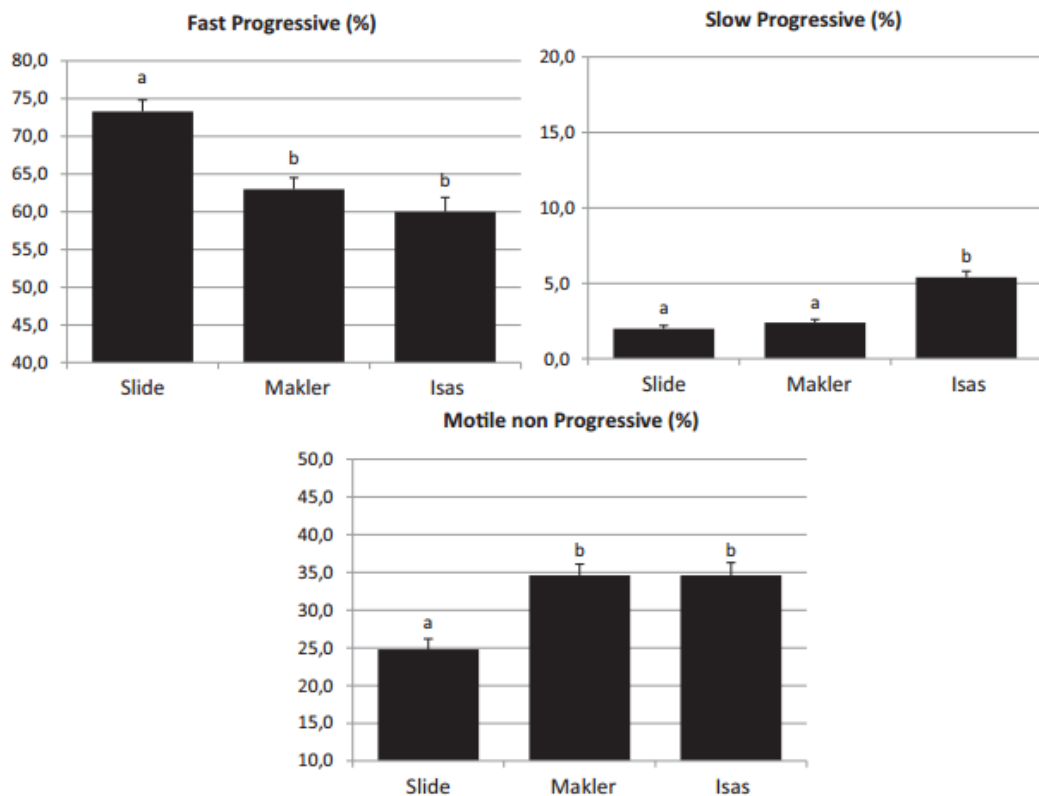


Figura 15. Comparación de los tres métodos para cada tipo de motilidad. (a) Progresiva rápida. (b) Progresiva lenta. (c) No progresiva. Tomado de I. Palacín et al., 2013.

Como resultado los autores evidencian que para los volúmenes de 5 y 10 μl y para los contenedores de muestra no hubo diferencia significativa en los parámetros de motilidad, a excepción de los valores de ALH y BCF para el caso de la laminilla con cubreobjetos con 5 μl , donde se obtuvieron resultados más altos. Se encuentra, además, que los parámetros de velocidad son distintos para cada método, y que presenta aún mayor diferencia en los parámetros de LIN, STR y WOB. Como lo

muestra la figura 15, para el conjunto de resultados la motilidad progresiva rápida se obtuvo un 73% para la lámina, un valor muy diferenciado con respecto al 63% de la cámara Makler y el 60% de la cámara ISAS.

En (Nurhadiyatna, et al., 2014) se presenta un estudio comparativo de algunas metodologías para la detección y rastreo de espermatozoides. Los autores desarrollan su propio algoritmo mediante la implementación de una combinación de modelos gaussianos con llenado de regiones (GMMHF) y un filtro de Kalman. Se evalúan otras metodologías tales como: diferencia de fotogramas, una versión más simple de Gaussian Mixture Model mostrado por Stauffer y Grimson, Pixel-Based Adaptative Segmenter (PBAS), Real-Time Tracking (Wren) y Adaptative Density Estimation (Zivkovic). Se calculan porcentajes de correlación para identificar aquellos métodos más eficientes para el análisis de motilidad y, por consiguiente, encontrar los más viables para medir la probabilidad de éxito de procedimientos de fertilización in-Vitro.

El porcentaje de precisión mayor obtenido corresponde al método GMMHF de los autores con 98.01% según se muestra en la tabla 8. No obstante los demás algoritmos no fueron significativamente inferiores, puesto que registraron porcentajes no menores a 97.44%. Además, se calculó el coeficiente de concordancia Kappa para estos algoritmos y se obtuvo el método GMM de Stauffer presentó el mayor valor con 0.48 lo que significa que para la mayoría de cuadros examinados presenta un buen rendimiento en la detección de espermatozoides ya que es sensible para pequeños cambios. Por otro lado, el menor valor Kappa corresponde para el método de diferencia de fotogramas, esto se debe a que este método de segmentación por movimiento necesita que los espermatozoides estén en constante movilización, es decir, si un espermatozoide no se mueve, es considerado como fondo y será ignorado. En conclusión, los métodos de distribución gaussiana son competitivos a la hora de detectar espermatozoides a través del video de una muestra microscópica.

Tabla 8. Precisión de las metodologías evaluadas. Tomado de A. Nurhadiyatna et al., 2014.

Methods	Frames										Avg.Accu(%)
	100	120	202	250	310	440	500	602	740	750	
Frame Difference	97.81	97.48	98.02	97.66	97.42	97.62	97.11	98.03	97.64	97.65	97.64
Stauffer and Grimson	97.65	97.69	98.11	97.72	96.77	97.42	96.89	97.74	97.20	97.23	97.44
PBAS	98.04	97.63	98.27	97.94	97.66	97.64	97.31	98.15	97.76	97.80	97.82
Wren	97.73	97.41	98.10	98.05	97.71	97.66	97.33	98.24	97.71	97.71	97.77
Zivkovic and Heijden	97.04	96.75	97.31	96.99	97.24	97.85	97.50	98.39	97.90	97.90	97.49
GMMHF	98.25	97.95	98.50	98.30	97.80	97.78	97.45	98.39	97.80	97.90	98.01

Y, por último, un importante estudio presentado en (Amann & Waberski, 2014) busca resaltar las capacidades y las limitaciones de los sistemas de análisis espermático asistido por computadora. Según los autores, es posible afirmar que un sistema CASA no predice precisamente la fertilidad de una muestra, sin embargo, mediante correctos preparativos y una validación suficiente, estos sistemas pueden ofrecer datos muy importantes para muestras seminales. Cámaras de espécimen, enfoque de lentes y cantidad de luz son condiciones cruciales que afectan de manera directa y profunda el resultado de cualquier análisis de un sistema CASA. Todos estos factores inciden en la calidad de las imágenes y/o videos. Usualmente los sistemas modernos requieren velocidades de captura iguales o mayores de 50 Hz, puesto que más centroides pueden reconstruir más precisamente una trayectoria, y por lo tanto tiene un impacto en los resultados de motilidad.

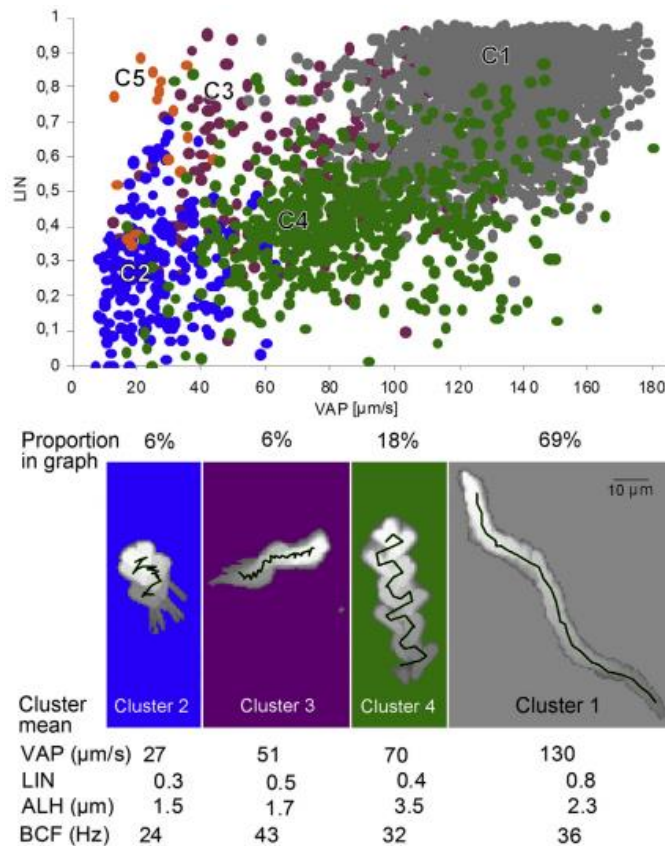


Figura 16. “Clustering” de datos de motilidad como validación. Espermatozoides con un valor VAP alto tienden a presentar a su vez una linealidad alta. Tomado de R. Amann & Waberski, 2014.

En ámbitos comerciales, se deben tener en cuenta que las muestras seminales a ser vendidas, deben no solo tener un análisis espermático, sino que también gozar de una garantía de calidad. Estas muestras deben ser validadas bajo múltiples análisis por medio de esta herramienta y bajo la supervisión de expertos que se encarguen de evaluar la correlación de los datos obtenidos con los datos observados. Desafortunadamente, muchos sistemas CASA no presentan altos niveles de control de calidad. Los procesos de validación no solo se dan en el análisis del resultado final, sino del proceso general, que comienza por supuesto con la preparación de las muestras seminales, e incluso con las técnicas de extracción en animales. En la figura 16 se observa un proceso de validación de resultados. Quizás el factor más importante del cual se deba tener mayor precisión es la velocidad curvilínea (VCL), ya que de ahí se extraen otros parámetros fundamentales.

3. Capítulo III. Diseño metodológico

Como primera actividad, se dispuso a preparar de manera rigurosa las muestras de semen de toro mediante los protocolos establecidos en los laboratorios de andrología y medicina veterinaria, teniendo en cuenta que este procedimiento puede influir notoriamente en los resultados finales del análisis de motilidad (Contri, et al., 2010). Para cada toro, se extrajeron las muestras del recipiente de nitrógeno líquido, se descongeló una pajilla mediante un baño de María durante 60 segundos y se colocaron 5 μ l de semen sobre un portaobjetos precalentado a 37°C.

3.1. Adquisición y pre-procesamiento

Una vez la muestra de semen de bovino se encuentre enfocada a x40 por medio de un microscopio de contraste de fase con cámara USB adaptada, se ejecuta el software controlador de la cámara denominado *touptek tuopview* desde el PC para capturar videos bajo las siguientes condiciones:

- Formato: WMV
- Resolución: 640x480 pixeles.
- Velocidad de fotogramas: 30 Hz.
- Velocidad de datos: 1572 kbps
- Duración: 5 segundos.

En la primera etapa del método diseñado, cada video capturado es guardado y posteriormente procesado matemáticamente. La herramienta computacional interpreta el video como una secuencia de imágenes continuas en el espacio de color RGB. Por lo tanto, cada fotograma se representa por medio de una matriz bidimensional donde cada elemento se compone de tres valores diferentes.

3.2. Segmentación y detección de espermatozoides

En la segunda etapa, previamente a la aplicación de modelos gaussianos adaptativos para la extracción del fondo y consecuente segmentación de las cabezas de células espermáticas, se ajusta el contraste del fotograma aumentando la diferencia entre pixeles claros y oscuros, con el fin de mejorar el desempeño de la segmentación por movimiento.

3.2.1. Modelos gaussianos adaptativos

Se emplea una técnica similar a la diferenciación de fotogramas. Esta técnica desarrollada en (Stauffer & W., 1999) se denomina por los autores como Modelos Adaptativos de Fondo y utiliza distribuciones gaussianas, como se visualizan en la figura 17, para analizar los pixeles de las imágenes consecutivas.

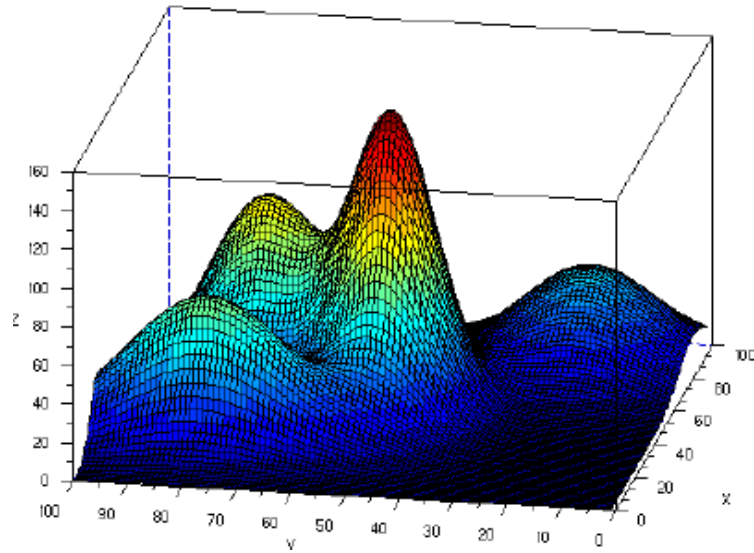


Figura 17. Representación tridimensional de una combinación de modelos gaussianos. Tomado de www.researchgate.net.

Cada pixel es analizado con respecto a múltiples modelos gaussianos creados de manera adaptativa, es decir, se busca definir la probabilidad de cada pixel a pertenecer a una región de fondo de imagen. Estas distribuciones gaussianas se modelan mediante el análisis de varianzas en los pixeles. Por lo tanto, cuando un pixel presenta un relativo valor de variabilidad en su magnitud a lo largo de una secuencia de fotogramas, se puede afirmar que pertenece al fondo de imagen, y no a un objeto en movimiento. La probabilidad de observar el valor del pixel se obtiene de la ecuación (1):

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} * \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (1)$$

Donde K es el número de distribuciones, ω es un peso estimado para modelo i por unidad de tiempo t , η es la densidad de probabilidad gaussiana, μ valor medio y Σ covarianza en la combinación de modelos. A su vez, la función de densidad de probabilidad gaussiana se muestra en (2):

$$(X_t, \mu_{i,t}, \Sigma_{i,t}) = \frac{1}{2\pi^2 |\Sigma|^{\frac{1}{2}}} * e^{-\frac{1}{2}(X_t - \mu_t)^T \Sigma^{-1} (X_t - \mu_t)} \quad (2)$$

De esta manera, a lo largo de los fotogramas, el algoritmo se entrena para la detección de objetos que se encuentran en el primer plano de la imagen, ya que éstos quedan fuera del modelo gaussiano de extracción de fondo. A medida que se procesan más fotogramas, el algoritmo se vuelve más robusto para la detección de movimiento y finalmente para la segmentación de los objetos de interés, en este caso, los espermatozoides.

3.2.2. Operadores morfológicos

Aunque la segmentación por los modelos gaussianos es efectiva para objetos en movimiento, normalmente se presentan pequeños objetos en la imagen binaria provocados por ruido o detecciones falsas. Por tanto, antes de llevar a cabo un análisis de componentes conectados es necesario optimizar la imagen segmentada mediante una serie de operaciones morfológicas. Las operaciones morfológicas son técnicas que permiten además de eliminar pequeñas partículas de ruido, moldear las regiones de interés en la imagen binaria mediante matemática morfológica (Maintz, 2005) con el fin de ajustarla previamente para el análisis de componentes conectados.

3.2.3. Análisis de componentes conectados

Una vez el fotograma se encuentre umbralizado, se lleva a cabo un análisis de componentes conectados. Los componentes conectados o n-conectados hacen referencia a los píxeles con valor 1 que conforman una región en la imagen binaria (Gonzalez & Woods, 2008). En este caso, cada conjunto o agrupación de píxeles conectados representan la cabeza de un espermatozoide en movimiento. El resultado de este análisis es el cálculo de los parámetros geométricos de los objetos de interés. Se obtiene así el centroide, que es el centro de masa del espermatozoide y corresponde a las coordenadas de éste sobre la imagen.

3.2.4. Análisis discriminante lineal

Teniendo en cuenta que el método de modelos gaussianos adaptativos solo puede segmentar espermatozoides en movimiento, el método de análisis discriminante diseñado en (Fisher, 1936) tiene como objetivo segmentar la totalidad de los

espermatozoides presentes en el último fotograma, sin importar si se encuentran en movimiento.

El análisis discriminante lineal de Fisher permite llevar a cabo una clasificación de datos entre dos o más clases. Su función consiste en maximizar la varianza inter-clase y minimizar la varianza intra-clase. De esta manera, se crea un vector discriminante, mostrado en la figura 18, que separe los pixeles pertenecientes al fondo de los pertenecientes a la cabeza de un espermatozoide. Este método requiere el uso de una base de datos de los valores de los pixeles en el espacio de color RGB para cada clase recolectada experimentalmente. Se calculan las matrices de dispersión inter-clase (3) e intra-clase (4).

$$S_B = \sum N_C (\mu_C - \mu)(\mu_C - \mu)^T \quad (4)$$

$$S_W = \sum N_{i \in C} (x_i - \mu_C)(x_i - \mu_C)^T \quad (5)$$

Donde μ es la media de la totalidad de los pixeles, μ_C es la media de cada clase, N_C es la cantidad de patrones en la clase y x_i es el valor de cada pixel. Posteriormente, se halla el vector discriminante (6) que separa las dos clases y permite llevar a cabo la segmentación. Siendo W la dirección del vector.

$$J(W) = \frac{W^T \cdot S_B \cdot W}{W^T \cdot S_W \cdot W} \quad (6)$$

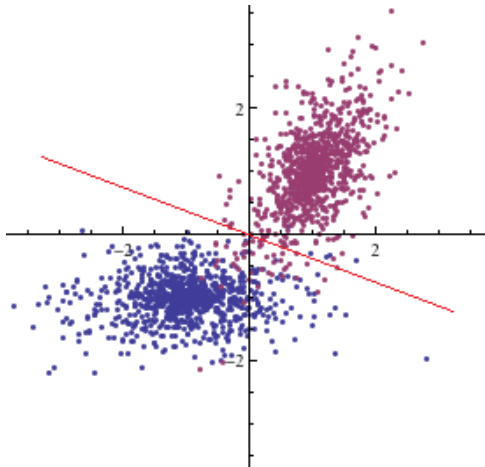


Figura 18. Agrupamiento de elementos en dos clases distintas. Tomado de *www.mathematica-journal.com*, 2016

3.3. Asignación y construcción de trayectorias

La asignación y construcción de las trayectorias descritas por lo espermatozoides se desarrolla en primera instancia por medio de la aplicación del algoritmo de asignación de Munkres, seguido de un filtro de Kalman. A continuación, se explica brevemente en qué consisten estas técnicas y cómo se aplicaron en la herramienta computacional.

3.3.1. Algoritmo de asignación de Munkres

La técnica implementada en esta sección corresponde a una variación del algoritmo de asignación de James Munkres o también conocido como Algoritmo Húngaro diseñado por (Munkres, 1957). Este modelo matemático tiene como fin realizar una asignación directa de m elementos a n destinos por medio de una matriz $m \times n$ denominada matriz de costes, como se muestra en la figura 19. Estos costes son valores relacionados a la probabilidad de asignar un elemento de la fila m a uno de la columna n . La finalidad de este proceso es realizar una asignación cuyo costo total sea mínimo. El tratamiento matemático de esta técnica modificada tiene los siguientes pasos (Burgeois & Lasalle, 1971).

1. Crear una matriz de costos C de tamaño $n \times m$.
2. Para cada fila, encontrar el menor elemento y restarlo de cada elemento en dicha fila.
3. Encontrar un cero (Z) en la matriz resultante. Si no hay ya un cero etiquetado con $*$ en su fila o columna, etiquetar Z , resultando Z^* .
4. Resaltar cada columna que contenga un Z^* . Si todas las columnas están resaltadas, los Z^* describen las asignaciones y el algoritmo termina en el paso 7, de lo contrario se continúa con el paso 5.
5. Hallar el valor menor no resaltado y sustraerlo al resto de los elementos no resaltados. Encontrar un cero y etiquetarlo como Z' . Si no hay ya un Z^* en la fila que contenga este Z' , continuar con el paso 6. De otra forma, resaltar esta fila y dejar de resaltar la columna que contenga el Z^* . Continuar de esta manera hasta que no haya ceros sin resaltar. Memorizar el valor sin resaltar más pequeño e ir al paso 7.
6. Construir una serie de ceros Z' y Z^* de tal forma que exista alternancia de estos elementos tanto en filas como en columnas. Continuar hasta que se halle un Z' que no tenga un Z^* en su columna. Convertir cada Z^* en Z y cada Z' en Z^* . Dejar de resaltar filas y columnas y regresar al paso 4.

7. El índice de cada Z^* describe la asignación. Para $C(i,j)$ el elemento i se asigna al elemento j .

	m_1	m_2	m_3	m_4		m_1	m_2	m_3	m_4	
n_1	5	3	2	8	→	n_1	5	3	2	8
n_2	8	4	6	7		n_2	8	4	6	7
n_3	3	7	6	9		n_3	3	7	6	9

total = 3 + 4 + 2 = 9

Figura 19. Representación de una matriz de costos. Los elementos de n se relacionan a m a través de una asignación de mínimo costo total. Elaborado por el autor.

Desde el punto de vista de la herramienta propuesta, en cada fotograma se presentan múltiples detecciones de espermatozoides que deben ser interpretadas y organizadas bajo un conjunto de trayectorias determinadas. Con el pasar de los fotogramas se tiene que garantizar que ciertas detecciones correspondan a una trayectoria en común para así tener registro de las coordenadas de movimiento. La matriz de costos $n \times m$ que se plantea en este caso, corresponde a n trayectorias y m detecciones. Cada elemento de la matriz representa la distancia entre el punto de referencia de una trayectoria establecida con el punto de detección para un fotograma determinado. El método de Munkres permite desarrollar un criterio para asignar una detección a una trayectoria según su posición de manera muy precisa.

3.3.2. Filtro de Kalman

El filtro de Kalman es un conjunto de ecuaciones matemáticas que proporcionan medios computacionales eficientes para estimar el estado de un proceso, en tal forma que minimiza la media del error al cuadrado. Este filtro permite realizar estimaciones de estados en tiempo pasado, presente e incluso futuro. Además, puede desempeñarse correctamente en los casos en donde se desconoce la naturaleza precisa de un sistema (Welch & Bishop, 2006). Este algoritmo, que fue diseñado por (Kalman, 1960), funciona mediante un mecanismo de predicción y corrección en la forma de un sistema de retroalimentación. Básicamente sirve como un estimador de un nuevo estado a partir de estimaciones previas agregando un valor de corrección, minimizando así errores de predicción.

La etapa de predicción funciona como una proyección de estados mediante las ecuaciones (6) y (7) del filtro de Kalman en tiempo discreto:

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1} \quad (6)$$

$$P_k^- = AP_{k-1}A^T + Q \quad (7)$$

Donde \hat{x}_k^- es el estado actual objeto de predicción, $k - 1$ representa el estado previo, A y B son matrices de transferencia de estados, P_k^- es la covarianza de error *a priori* y Q es la covarianza de ruido del proceso.

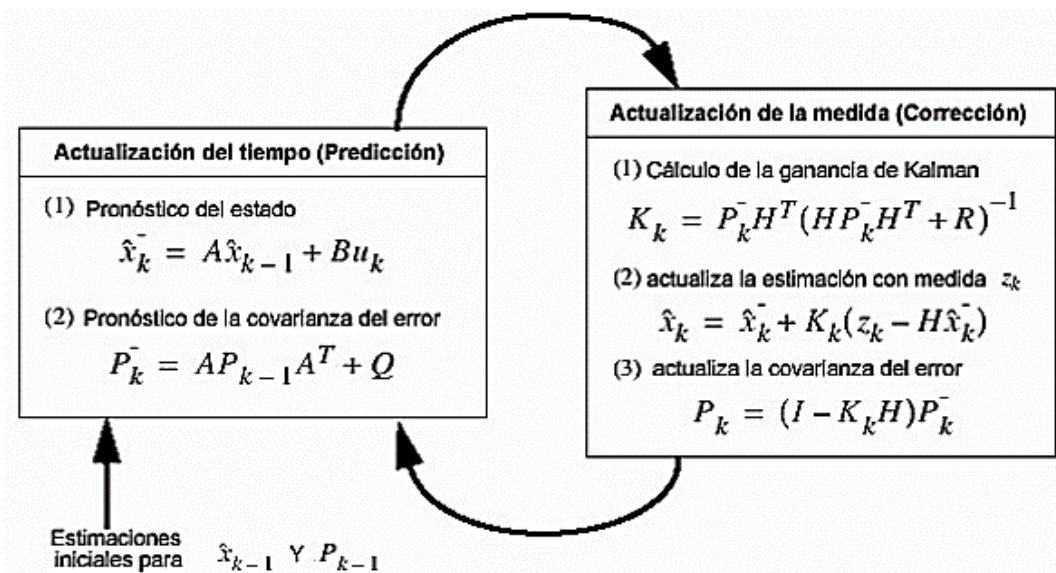


Figura 20. Diagrama del proceso de funcionamiento del filtro de Kalman discreto. Tomado de G. Prado, 2005.

La etapa de corrección es el resultado de la retroalimentación al actualizar las mediciones introduciéndolas a la estimación *a priori* mediante las operaciones (8), (9) y (10).

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \quad (8)$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \quad (9)$$

$$P_k = (I - K_k H)P_k^- \quad (10)$$

Donde K_k es la ganancia de Kalman que disminuye la covarianza de error *a posteriori*, H es una matriz relacionada al estado de la medición, R covarianza de

error de la medición y z_k es la medición actual del sistema. En la figura 20 se representa el modelo general del funcionamiento del filtro de Kalman.

De esta manera, el filtro de Kalman ajusta los datos de una medición mediante la acción de una predicción y corrección. Todas las mediciones previas constituyen un criterio para realizar la predicción de futuras mediciones. La corrección, a su vez, actualiza los datos y los incorpora al historial de información a través de un lazo de retroalimentación.

Por lo tanto, la aplicación del filtro de Kalman a la metodología propuesta, contribuye en la construcción de las trayectorias espermáticas llevando a cabo predicciones basadas en las posiciones previas de la cabeza del espermatozoide. El algoritmo de Munkres asigna la detección que más probabilidades tenga de pertenecer a determinada trayectoria basado en la configuración del filtro de Kalman (Figura 21). Esta asignación pasa a ser la posición actual del espermatozoide y así se actualiza mediante la corrección de su coordenada en el fotograma. Cuando un espermatozoide se desenfoca o de alguna manera no es detectado en la imagen, el filtro de Kalman realiza la predicción de su posición hasta que éste vuelva a aparecer y ser detectado cerca al punto de predicción actual del fotograma.

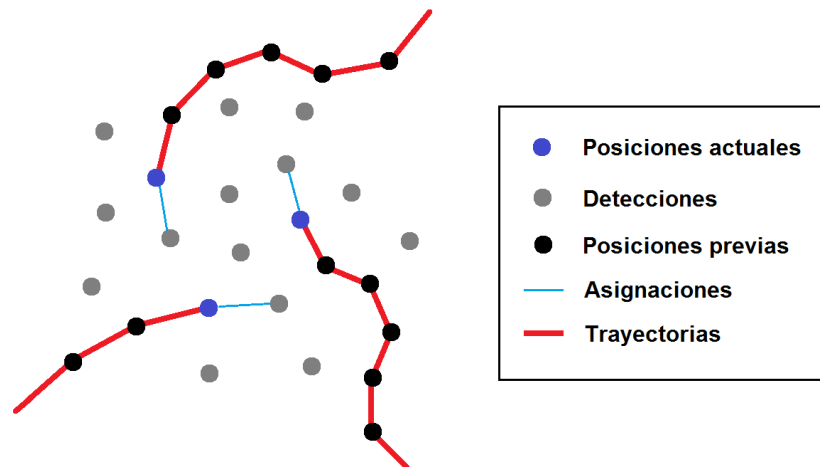


Figura 21. Representación de la asignación de detecciones de espermatozoides a sus trayectorias mediante el algoritmo de Munkres y el filtro de Kalman. Elaborada por el autor.

Aquellas detecciones que no son asignadas a ninguna trayectoria se toman como las posiciones iniciales de nuevas trayectorias. Las trayectorias inestables, es decir, aquellas que contienen pocas detecciones iniciales o presentan una carencia de

detecciones consecutivas muy grande son descartadas del proceso, ya que probablemente se traten de errores de detección.

3.4. Análisis de los parámetros de motilidad espermática

Se define entonces que la trayectoria de un espermatozoide está compuesta de una matriz bidimensional representando las coordenadas X y Y de las posiciones a través del tiempo dentro de la imagen.

Tabla 9. Ejemplo de coordenadas de una trayectoria en 8 fotogramas. Elaborada por el autor.

X	Y
658,42	473,97
659,21	474,54
660,15	483,07
660,52	484,97
663,64	485,33
657,85	487,77
656,72	489,56
655,59	491,34

A partir de las coordenadas obtenidas de cada trayectoria espermática, se calculan los parámetros de motilidad espermática más comunes entre los sistemas CASA (Boyers, et al., 1989).

3.4.1. Velocidad curvilínea (VCL)

Este es el parámetro fundamental de motilidad espermática individual. A partir de este valor se calculan los demás parámetros. La distancia total se obtuvo mediante la sumatoria de las distancias entre cada coordenada k en el espacio de la imagen y su multiplicación por un factor A que transforme coordenadas cartesianas (x_1, y_1) en micrómetros (μm). La velocidad se determinó en (11) al multiplicar esta distancia total por la frecuencia de captura del video F (s^{-1}) así:

$$VCL = A * F * \sum_{i=1}^k \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (11)$$

3.4.2. Velocidad promedio (VAP)

Debido a que esta velocidad corresponde a la trayectoria real suavizada, se aplicó un filtro de media para cada columna de la matriz de coordenadas y posteriormente

se computaron. Matemáticamente se calculó la media secuencial para cada elemento del vector columna (X_i, Y_i) obteniendo otro vector columna promedio (M_i, N_i) , mostrado en (12) y (13). Luego se aplica la misma ecuación utilizada anteriormente, pero adaptada a los nuevos valores (14):

$$M_i = \frac{\sum_i X}{i} \quad (12)$$

$$N_i = \frac{\sum_i Y}{i} \quad (13)$$

$$VAP = A * F * \sum_{i=1}^k \sqrt{(m_2 - m_1)^2 + (n_2 - n_1)^2} \quad (14)$$

3.4.3. Velocidad rectilínea (VSL)

Se computa la velocidad rectilínea desde el punto inicial (x_i, y_i) al punto final (x_f, y_f) de la trayectoria descrita por el espermatozoide, y de igual manera utilizando (15), se convierte a unidades de $\mu\text{m/s}$:

$$VSL = A * F * \sqrt{(x_f - x_i)^2 + (y_f - y_i)^2} \quad (15)$$

3.4.4. Índice de linealidad (LIN)

Se obtuvo al dividir la velocidad rectilínea sobre la velocidad curvilínea (16).

$$LIN = \frac{VSL}{VCL} \quad (16)$$

3.4.5. Índice de rectitud (STR)

Se obtuvo al dividir la velocidad rectilínea sobre la velocidad promedio (17).

$$STR = \frac{VSL}{VAP} \quad (17)$$

3.4.6. Índice de oscilación (WOB)

Se obtuvo al dividir la velocidad promedio sobre la velocidad curvilínea (18).

$$WOB = \frac{VAP}{VCL} \quad (18)$$

3.4.7. Motilidad total y progresiva (MT y MP)

Los porcentajes de motilidad total y motilidad progresiva son uno de los más importantes descriptores de la calidad de una muestra seminal (Vincent, et al., 2012). El criterio implementado para la clasificación de la motilidad espermática en las categorías de progresivo rápido, progresivo lento, no progresivo e inmóvil se basó en los criterios implementados en sistemas CASA para el análisis de muestras congeladas de semen bovino. Sin embargo, los umbrales utilizados para la categorización de adecuaron a las condiciones propias de nuestro diseño, debido a que para los sistemas CASA no existe un estándar único (Simonik, et al., 2015). En la figura 22 se observa el proceso de clasificación.

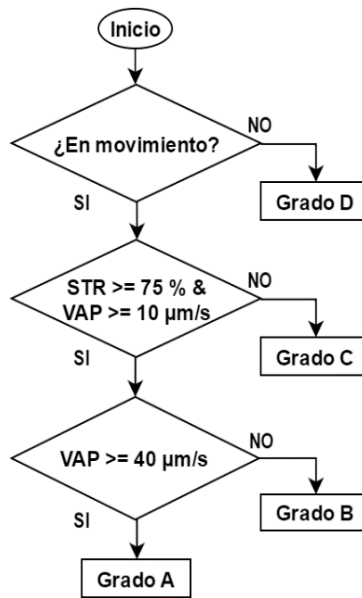


Figura 22. Diagrama de flujo de la clasificación de los tipos de motilidad. Elaborada por el autor.

- Existe motilidad progresiva rápida o lenta si el espermatozoide presenta un índice de rectitud (STR) mayor o igual al 75 % y una velocidad promedio (VAP) mínima de 10 µm/s.
- Hay motilidad progresiva rápida (Grado A) si el espermatozoide tiene una velocidad promedio (VAP) mayor o igual a 40 µm/s.
- Hay motilidad progresiva lenta (Grado B) si el espermatozoide tiene una velocidad promedio (VAP) entre 10 µm/s y 40 µm/s.

- Si el espermatozoide presenta una velocidad promedio menor a 10 $\mu\text{m/s}$ o un índice de linealidad (LIN) menor al 75 %, se considera motil, pero no progresivo (Grado C).
- Un espermatozoide que no presente movimiento propio en absoluto se clasifica en Grado D.

Mediante la sumatoria de espermatozoides analizados por video con respecto a las categorías A, B, C y D se calculan los porcentajes de motilidad total (19) y motilidad progresiva (20):

$$MT = \frac{A+B+C}{Total} \quad (19)$$

$$MP = \frac{A+B}{Total} \quad (20)$$

4. Capítulo IV. Análisis de resultados

4.1. Procesamiento de video

A continuación, se describen los resultados de la aplicación de las diferentes etapas de procesamiento de video del algoritmo resultante mostrado en la figura 23. Así mismo, se muestran los errores obtenidos y la explicación de su causa.

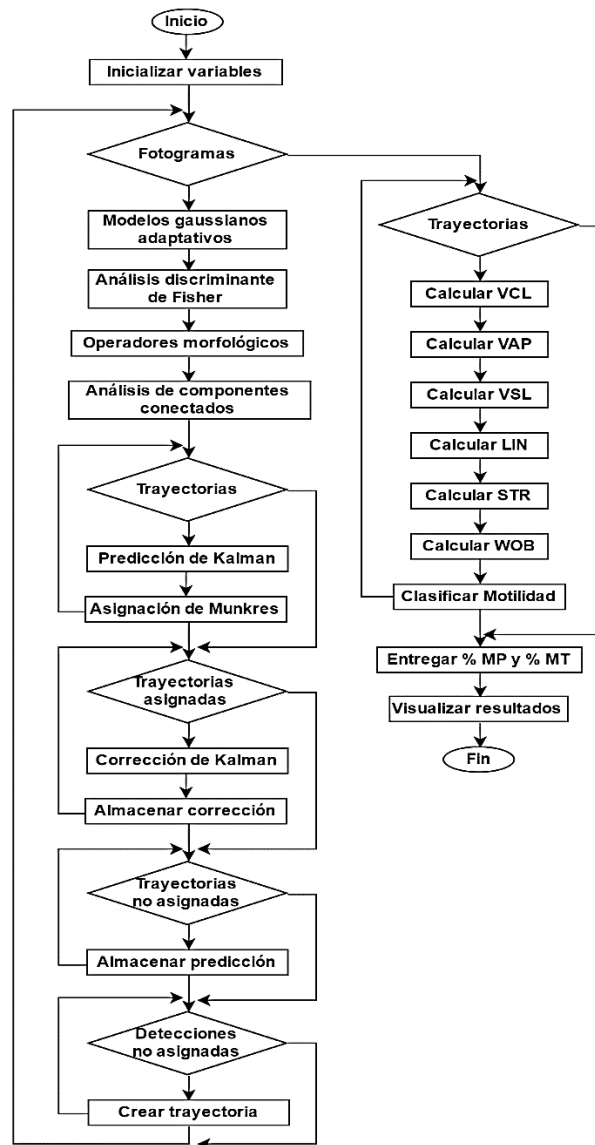


Figura 23. Diagrama de flujo del algoritmo. Elaborada por el autor.

4.1.1. Segmentación y detección de espermatozoides

El ajuste de contraste fue aplicado de manera experimental con el fin de evidenciar su impacto en la etapa de segmentación. Se encontró que un leve aumento de contraste del fondo con respecto a los espermatozoides favoreció el desempeño de la segmentación por modelos gaussianos. En la figura 24-B se muestra el cambio en los niveles de color de la imagen original.

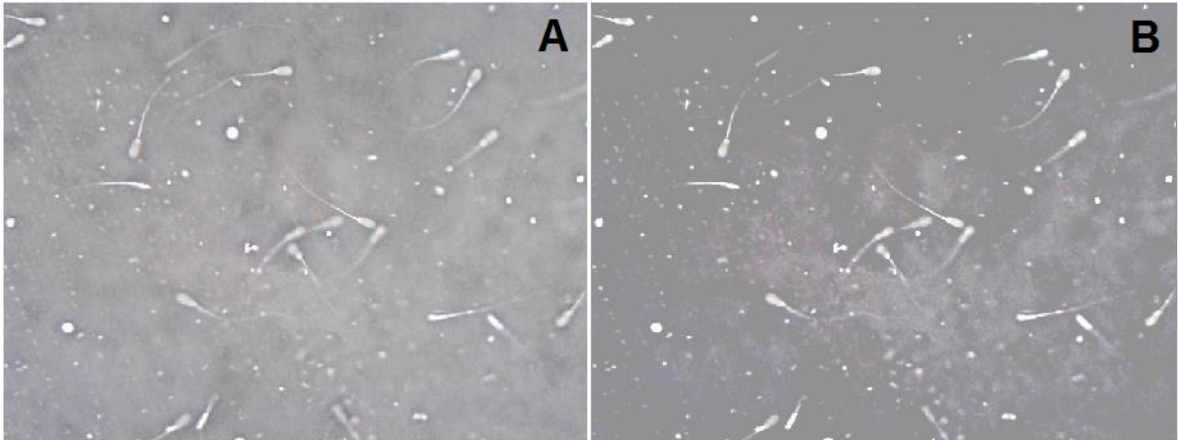


Figura 24. (A) Fotograma original. (B) Ajuste de contraste. Elaborada por el autor.

La segmentación mediante una combinación de modelos gaussianos adaptativos tuvo el efecto mostrado en las figuras 25-B1 a 25-B3. Se evidencia que aquellos espermatozoides en movimiento son extraídos del fondo junto con algunas partículas generadas por la vibración del semen provocadas por la turbulencia propia del movimiento flagelar de los espermatozoides.

También son extraídas las colas de los espermatozoides, ya sea completa o parcialmente. Así como aquellos espermatozoides con desplazamiento muy pobre o casi nulo son segmentados igualmente de forma parcial. Nótese además que los espermatozoides sin movimiento en absoluto no son segmentados por esta técnica, debido a que los pixeles que los conforman no presentan variabilidad y son prácticamente tomados como parte del fondo de imagen.

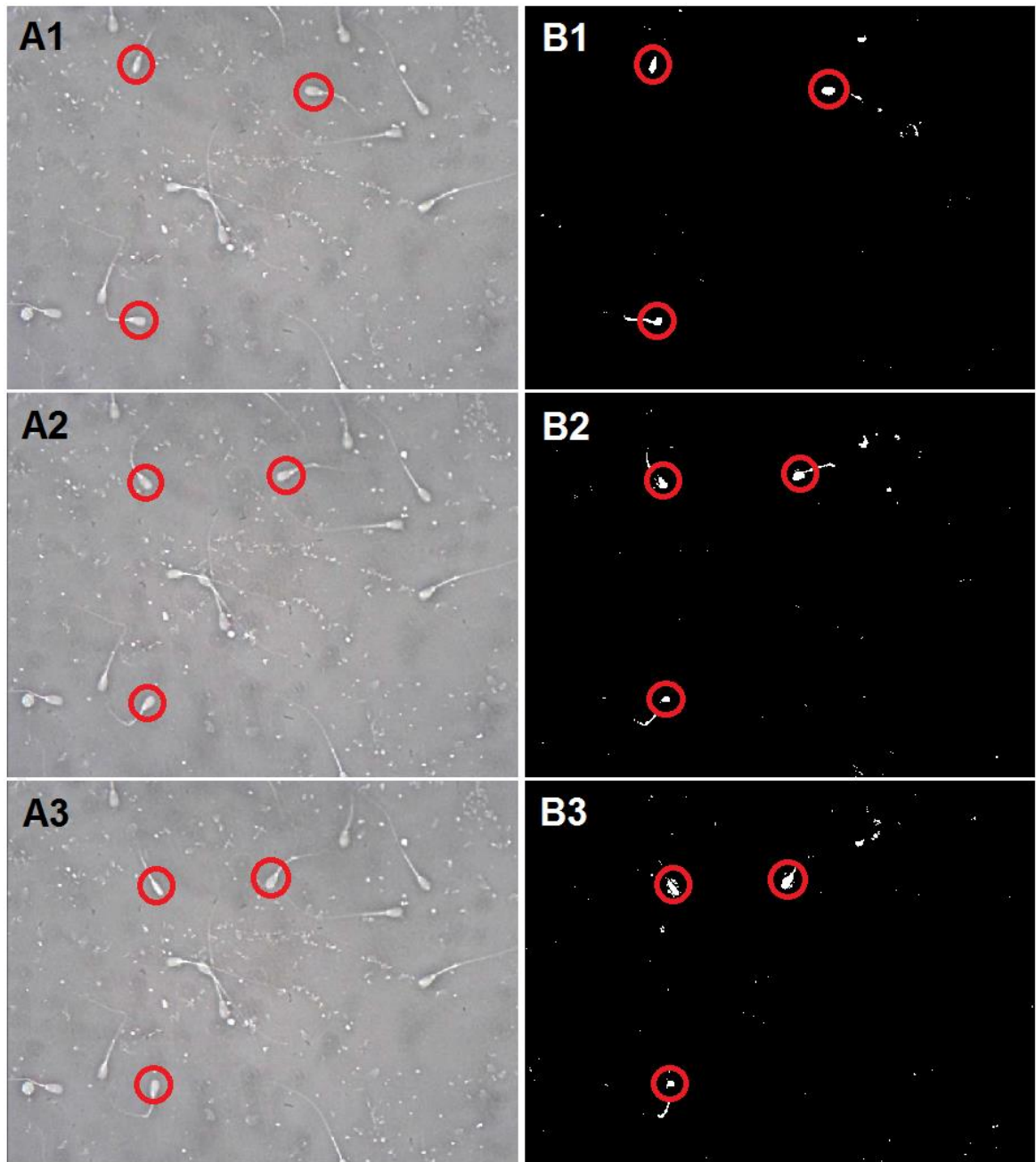


Figura 25. Secuencia de fotogramas mostrando la segmentación de tres espermatozoides en movimiento por el método adaptativo de modelos gaussianos. Elaborada por el autor.

La aplicación de operadores morfológicos a la imagen binaria resultante del método de segmentación permitió eliminar las pequeñas partículas generadas por el movimiento del fondo. También, como se evidencia en la figura 26-B, las colas de los espermatozoides son reducidas hasta ser desaparecidas, dejando así

únicamente las cabezas de los espermatozoides en movimiento. Sin embargo, el proceso de filtrado no es suficiente y en ocasiones permite la presencia de otros objetos de mayor tamaño en la imagen que no pertenecen a cabezas espermáticas. Se aplicó un operador de apertura seguido de uno de cierre, con un objeto estructurante circular, de radios de 2 y 4 píxeles respectivamente, obteniéndose la imagen filtrada mostrada en la figura 26-B.

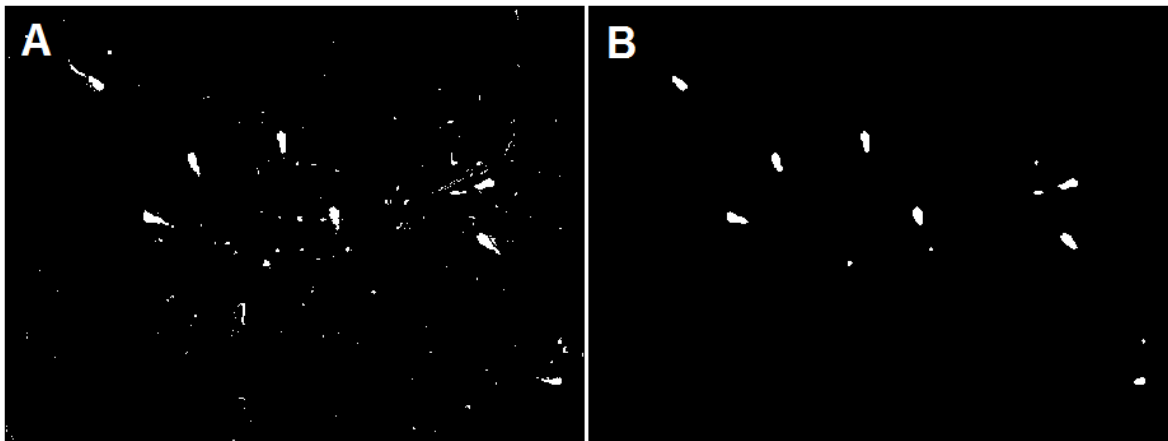


Figura 26. (A) Imagen binaria original (B) Filtrada por operadores morfológicos. Elaborada por el autor.

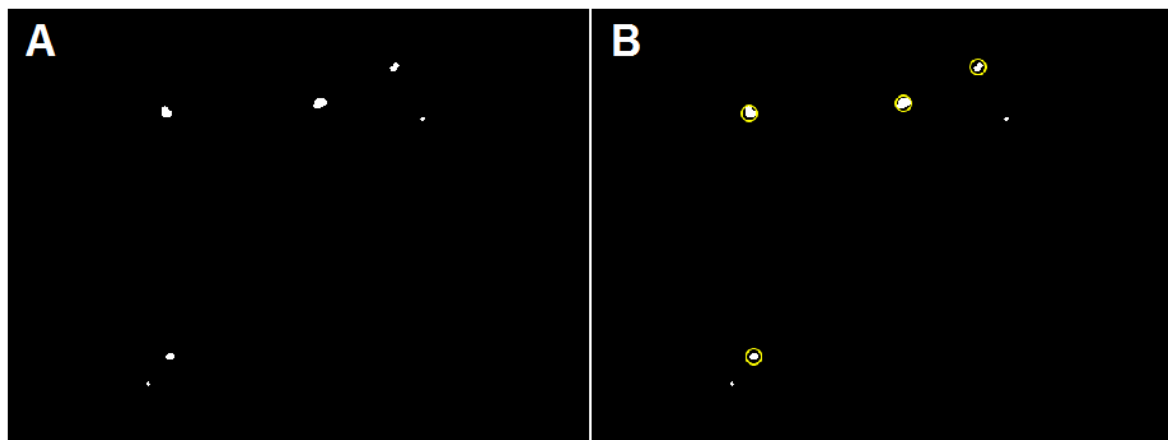


Figura 27. (A) Imagen filtrada. (B) Centroides calculados. Elaborada por el autor.

Seguidamente, mediante el análisis de componentes conectados se calculan el área de cada región de píxeles con valor 1 de la imagen binaria. Experimentalmente se observó que las cabezas espermáticas presentaban como mínimo 100 píxeles de área. Luego, a través de un barrido en la imagen se eliminaron aquellas regiones

que presentaron áreas menores a este umbral. Así, se descartan finalmente regiones conectadas no pertenecientes a los objetos de interés. Una vez garantizadas las regiones pertenecientes únicamente a cabezas de espermatozoides, se procedió a calcular sus centroides. En la figura 27-B se observa el resultado del análisis de componentes conectados, donde las áreas señaladas corresponden a espermatozoides en movimiento. La obtención de centroides permitió entonces la utilización del filtro de Kalman y el algoritmo de Munkres.

Puesto que los modelos gaussianos adaptativos no segmentaron los espermatozoides sin motilidad, fue necesario otro método en conjunto. El análisis discriminante de Fisher permitió segmentar las cabezas de las células espermáticas independientemente de su movilidad o vitalidad. En la figura 28-A se observa el fotograma final de un video de prueba al cual se le aplicó este análisis. Su resultado se visualiza en la figura 28-B, donde las cabezas son extraídas del fondo de manera efectiva presentando tan solo un par de errores.

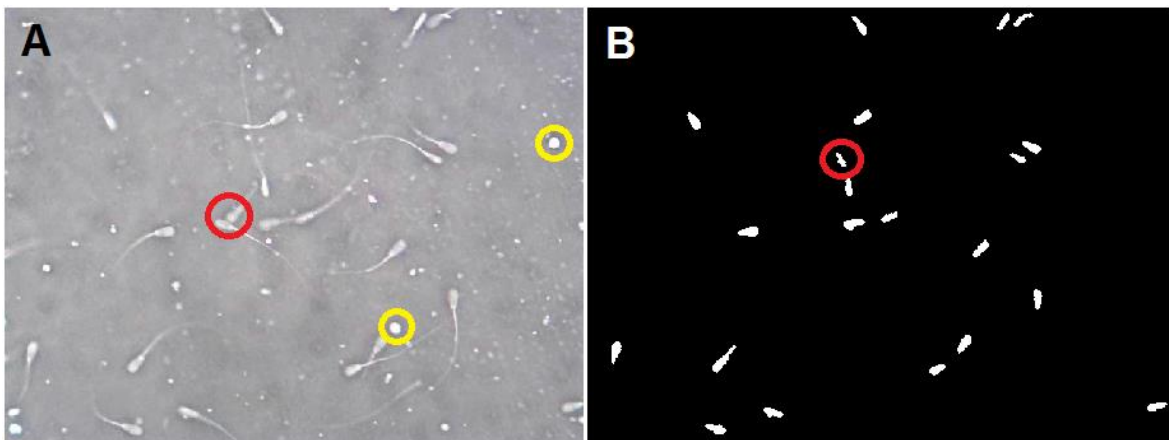


Figura 28. (A) Fotograma original. (B) Segmentación por análisis discriminante de Fisher.
Elaborada por el autor.

En el fotograma original se observan dos espermatozoides muy próximos (en rojo) que en el proceso de análisis de áreas son descartados por su alto valor y por eso no se segmentan en el resultado. También, en la figura 28-B se muestra la segmentación errónea de la cola de un espermatozoide (en rojo). Sin embargo, es positivo resaltar que los objetos de color y área similar a una cabeza espermática pero que no corresponde a esta (en amarillo), son descartados de la segmentación gracias a una discriminación de excentricidades, puesto que las cabezas no son perfectamente circulares sino ovaladas. De esta forma, se mejoran los resultados

de la segmentación por Fisher aportando al análisis de motilidad total de la muestra. El desempeño en el análisis discriminante de Fisher también puede observarse en la figura 29, donde se separan de manera efectiva las dos clases por medio de un vector discriminante cruza en el punto de mayor varianza inter-clase. En azul se muestran los valores en el espacio de color RGB de pixeles pertenecientes al fondo de los fotogramas. En rojo se muestran los pixeles pertenecientes a las cabezas de las células espermáticas. La inclinación del vector se debe a la desviación de algunos pixeles rojos en la clase azul, como se evidencia en la figura 29. Así, se demuestra la capacidad del análisis discriminante de Fisher para la segmentación de una imagen a color.

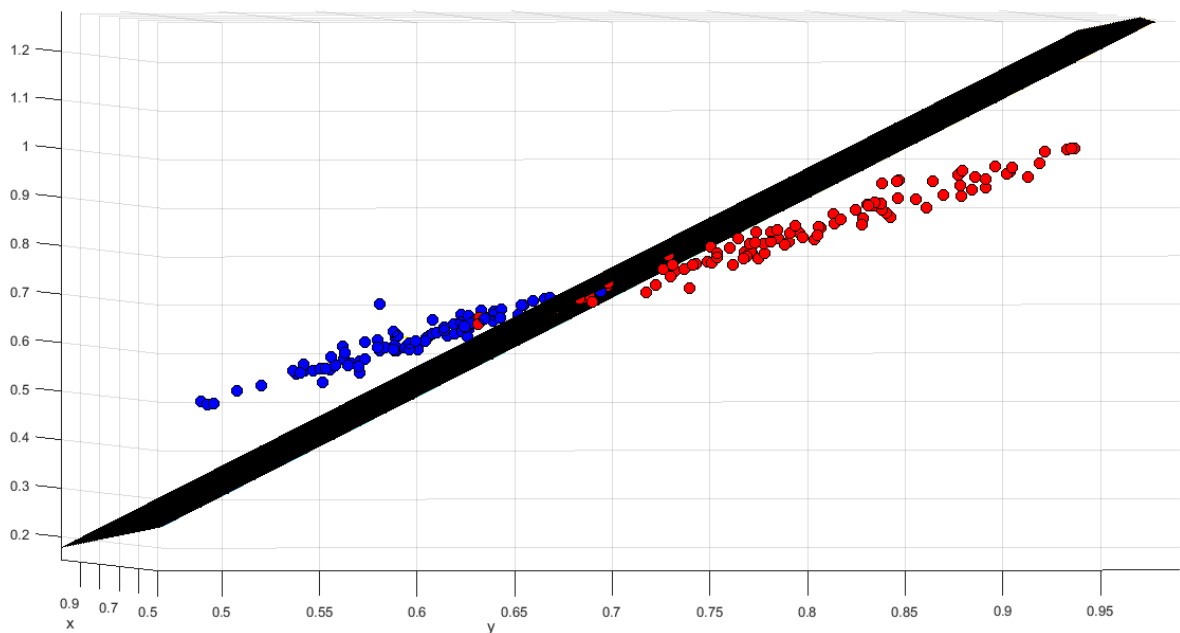


Figura 29. Separación de las clases de pixeles de espermatozoides y fondo mediante un plano discriminante. Elaborada por el autor.

4.1.2. Asignación y construcción de trayectorias

En los primeros fotogramas del video, los primeros centroides detectados fueron considerados como el inicio de nuevas trayectorias. Esto también sucede en posteriores fotogramas, donde nuevas detecciones fuera del rango de trayectorias existentes son consideradas nuevas trayectorias. Por lo tanto, la posición inicial junto con las condiciones iniciales de covarianza de posición, velocidad y ruido de medición hicieron posible el funcionamiento del filtro de Kalman para cada trayectoria presente en el video. Así mismo, el algoritmo de asignación de Munkres

utiliza los datos del filtro de Kalman para relacionar las detecciones a sus correspondientes trayectorias de manera correcta.

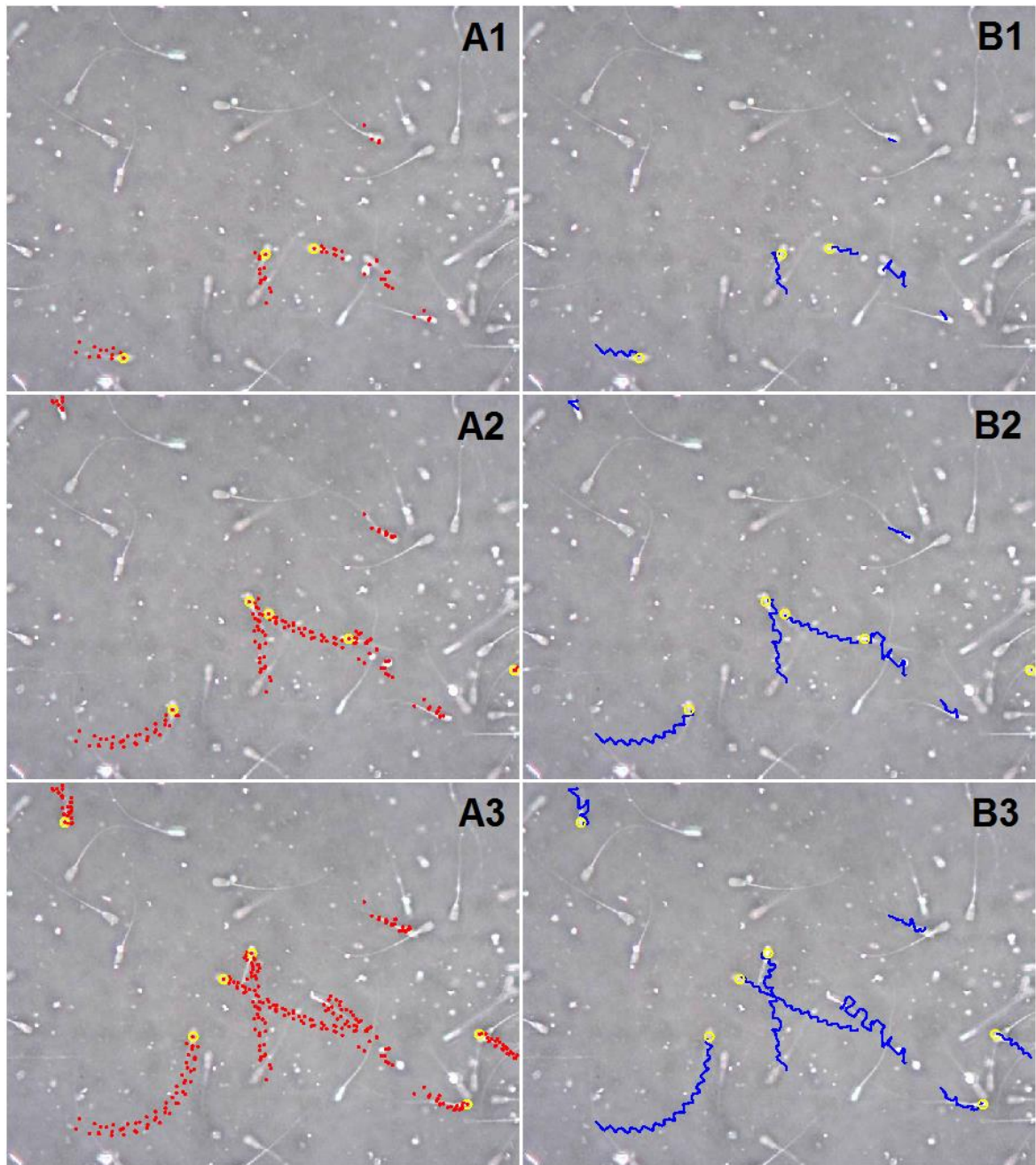


Figura 30. Fotogramas en diferentes tiempos 1,2,3. (A) En puntos rojo se marcan las detecciones de centroides sin asignar. (B) En líneas azules se ilustran las trayectorias construidas. Elaborada por el autor.

Básicamente, el algoritmo de Munkres utilizó la predicción del filtro de Kalman para asignar correctamente las nuevas posiciones de una trayectoria. La predicción del filtro de Kalman a su vez funciona mediante la corrección de la posición en píxeles de nuevos centroides detectados en los fotogramas y la retroalimentación de estos datos para la construcción del modelo de trayectoria de cada espermatozoide en el video. El modelo de trayectorias que se implementó fue el de velocidad constante, puesto que el comportamiento cinético observado de los espermatozoides deja en evidencia su modo de desplazamiento prácticamente sin cambio de velocidad con respecto al tiempo.

En la figura 30, se evidencia el efecto de la aplicación del filtro de Kalman y el algoritmo de Munkres. En las figuras 30-A1 a 30-A3 se observa la transición de tres fotogramas en diferentes tiempos, mostrando las posiciones de los centroides detectados. En las figuras 30-B1 a 30-B3 se muestra la asignación y corrección de estas posiciones a las trayectorias reales de los espermatozoides. Nótese que en las figuras 30-A3 y 30-B3 se presenta la superposición de dos trayectorias espermáticas. Sin embargo, la asignación de Munkres permite diferenciar las dos trayectorias. Se observó además que cuando la posición actual del espermatozoide no es detectada, ya sea debido a errores de segmentación, el filtro de Kalman entra a hacer la predicción de dicha posición en cada momento, para que cuando la detección se haga de nuevo presente, se lleve a cabo una corrección de posición al computar el modelo de predicción con la medición actual.

4.1.3. Caracterización de motilidad espermática

Con el filtro de Kalman, se logró definir la trayectoria de cada espermatozoide en movimiento presente en el video. Como las trayectorias son descritas como una serie de coordenadas cartesianas en la imagen, se procedió a aplicar las ecuaciones (11) a (18) a estos puntos con el fin de caracterizar y clasificar cada espermatozoide dentro de un grado de motilidad de manera similar a cómo lo hacen los sistemas CASA comerciales.

En la figura 31-A se observa un espermatozoide con una buena velocidad curvilínea e índice de rectitud ($VCL = 75.3 \mu\text{m/s}$ y $STR = 90.2 \%$ respectivamente), sin embargo, presenta un índice de ondulación bajo ($WOB = 48.9 \%$), lo que significa que tiene una oscilación muy pronunciada. En la figura 31-B se muestra un espermatozoide con una velocidad lineal buena ($VSL = 34.1 \mu\text{m/s}$), índices de rectitud y ondulación altos ($STR = 99.4 \%$ y $WOB = 89.3 \%$), lo que lo hace un espermatozoide con motilidad progresiva rápida, entrando en la categoría A.

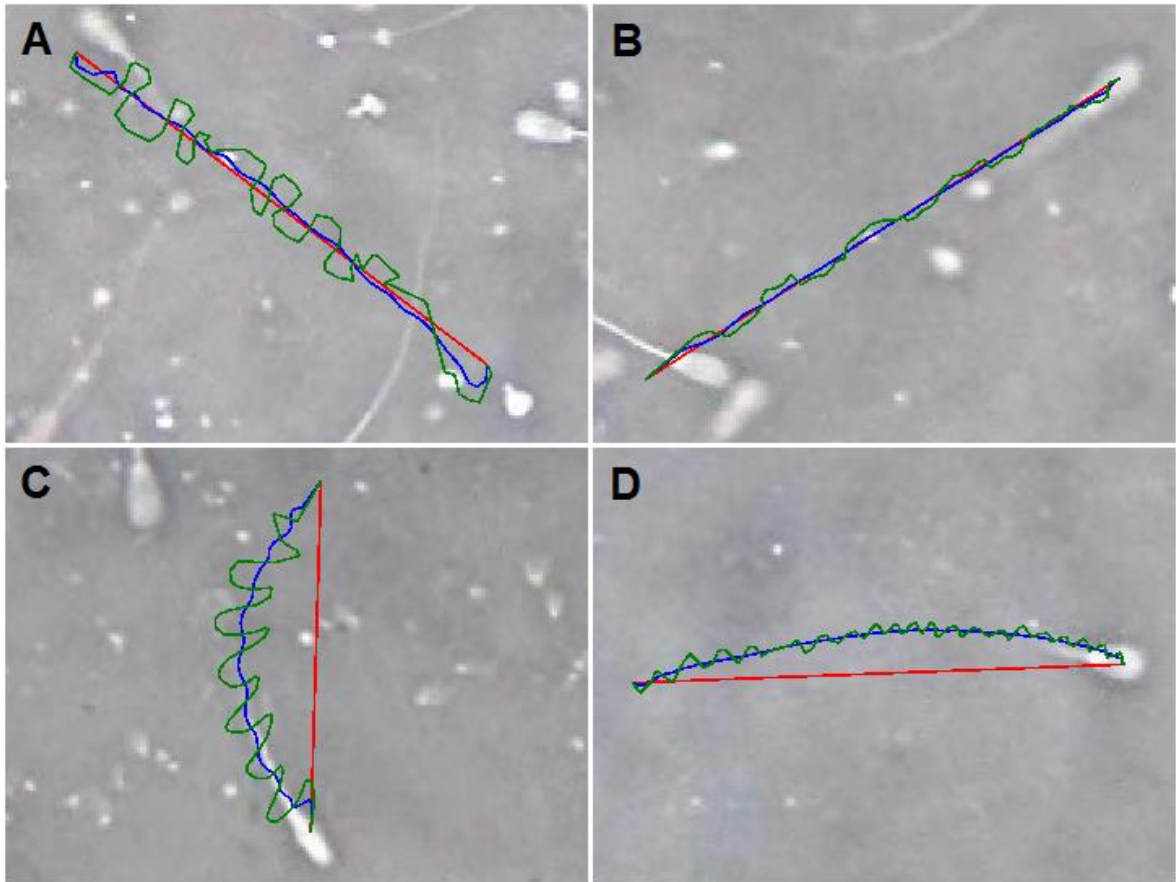


Figura 31. Diferentes comportamientos cinéticos observados en los videos procesados. En verde VCL, en azul VAP y en rojo VSL. Elaborada por el autor.

También, en la figura 31-C se observa la trayectoria muy ondulante de un espermatozoide ($WOB = 47.3 \%$) combinada con una linealidad muy pobre ($LIN = 39.1 \%$), lo que lo hace un espermatozoide sin progresividad perteneciendo a la categoría C. Finalmente, en la figura 31-D se evidencia un espermatozoide regular con una velocidad promedio ($VAP = 23.9 \%$) y una rectitud ($STR = 97.1 \%$). Esta última es la subpoblación más predominante en las muestras observadas haciendo parte de la categoría B de clasificación de motilidad espermática. En la figura 32 se evidencia la clasificación de los espermatozoides a través de su velocidad promedio y el índice de rectitud, los círculos rojos indican los pertenecientes al grado C, los círculos azules indican los pertenecientes al grado B Y los círculos verdes muestra los espermatozoides en grado A.

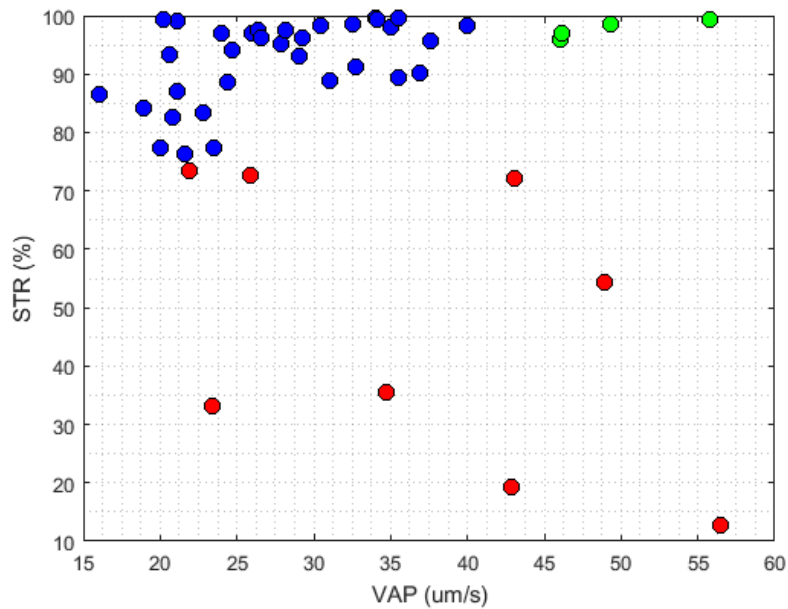


Figura 32. Relación de STR y VAP en la clasificación de motilidad de los espermatozoides analizados. (Rojo) Grado C. (Azul) Grado B. (Verde) Grado A.

Los hallazgos evidenciaron que en general las muestras presentan una motilidad total y progresiva baja. No obstante, se debe tener en cuenta que se utilizaron muestras que previamente habían sido sometidas a ciclo de crio-preservación, lo que tiene un impacto notable en la capacidad motil de los espermatozoides (Sundararaman, et al., 2012).

4.2. Validación de la herramienta computacional

El objetivo de la herramienta computacional es llevar a cabo un análisis de motilidad espermática a partir de una secuencia de imágenes obtenidas de un microscopio. Este análisis involucra el seguimiento de cada uno de los espermatozoides, el cálculo de sus trayectorias y velocidades, para que posteriormente se clasifiquen en las categorías de motilidad anteriormente mencionadas. A partir del conteo de la totalidad de espermatozoides y de la cantidad de ellos presente en cada categoría, se obtiene el análisis de motilidad progresiva y total. Por consiguiente, el resultado principal de la herramienta propuesta son el porcentaje de motilidad progresiva y el porcentaje de motilidad total.

Teniendo en cuenta esto, el análisis manual de la motilidad espermática por parte de un experto también tiene como finalidad hallar un porcentaje de motilidad progresiva y total, siendo la progresiva la más significativa, dada su estrecha relación

con la calidad de la muestra seminal. Por tanto, la validación correspondiente en este caso es la comparación de los resultados en cuanto a motilidad progresiva y total entregadas tanto por el algoritmo como por el especialista. Finalmente encontrando el nivel de correlación entre ambos resultados y, por ende, aprobando la herramienta computacional.

La evaluación llevada por parte del experto se basó en el mismo criterio de clasificación descrito en el estado del arte, a excepción de las categorías A y B. Según el experto, en los nuevos manuales de análisis de motilidad espermática los espermatozoides progresivos se categorizan en una sola clase sin necesidad de diferenciar aquellos muy rápidos o lentos, por lo tanto, las categorías A y B son unidas en una sola clase. Teniendo en cuenta esto, la clasificación parte de la siguiente manera.

- Clase 1: espermatozoides con motilidad progresiva (Categoría A+B).
- Clase 2: espermatozoides con motilidad no progresiva (Categoría C).
- Clase 3: espermatozoides sin motilidad (Categoría D).

De estas clases se derivan el cálculo de la totalidad de espermatozoides detectados (Total), el porcentaje de motilidad progresiva (% MP) y el porcentaje de motilidad total (% MT). De esta forma, la observación y evaluación obtenidas a partir de los 10 videos por parte del experto en andrología se evidencian en la tabla 10, donde se muestran el número de espermatozoides clasificados en cada categoría y los porcentajes de motilidad progresiva y total.

Tabla 10. Resultados del análisis manual por parte del técnico experto. Elaborada por el autor.

Video	Cat. A+B	Cat. C	Cat. D	Total	% MP	% MT
1	7	2	19	28	25,00	32,14
2	6	1	13	20	30,00	35,00
3	3	1	10	14	21,43	28,57
4	2	2	3	7	28,57	57,14
5	5	0	4	9	55,56	55,56
6	7	1	9	17	41,18	47,06
7	2	3	7	12	16,67	41,67
8	5	2	5	12	41,67	58,33
9	5	0	9	14	35,71	35,71
10	6	1	11	18	33,33	38,89

Mientras tanto, se dispuso el análisis de estos 10 videos de muestras seminales de diferentes toros y se procedió a ejecutar el algoritmo con el fin de evidenciar sus resultados en cuanto a los parámetros de motilidad observados por el experto. En la tabla 11, se muestran los resultados obtenidos por la herramienta computacional.

Tabla 11. Resultados del análisis semi-automático por parte de la herramienta computacional. Elaborada por el autor.

Video	Cat. A+B	Cat. C	Cat. D	Total	% MP	% MT
1	7	1	19	27	25,93	29,63
2	6	0	13	19	31,58	31,58
3	2	2	9	13	15,38	30,77
4	3	0	4	7	42,86	42,86
5	5	1	2	8	62,50	75,00
6	8	2	7	17	47,06	58,82
7	3	0	12	15	20,00	20,00
8	4	0	7	11	36,36	36,36
9	5	2	7	14	35,71	50,00
10	6	1	11	18	33,33	38,89

4.3. Análisis estadístico

Con el fin de interpretar los datos obtenidos, se llevó a cabo un análisis de estadístico donde se compararon los datos del algoritmo y del experto para cada clase. Se halló el coeficiente de determinación (R^2), el cual se deriva del coeficiente de Pearson para cada conjunto de datos. Además, se observa la línea de tendencia en cada gráfico de dispersión. La interpretación del coeficiente de determinación consiste en verificar la similitud entre dos conjuntos de datos (Arriaza, 2006). Este valor oscila entre 0 y 1. Un R^2 muy próximo a 0 indica una baja correlación y por tanto una diferencia significativa en los datos pareados. Un R^2 muy próximo a 1 indica una alta correlación, lo cual da evidencia de una alta similitud y también, una alta capacidad predictiva.

En la figura 33, se observa la correlación de los espermatozoides en la categoría A+B para el experto y el algoritmo. Se demuestra que la herramienta computacional fue capaz de clasificar en su mayoría los espermatozoides progresivos ($R^2 = 0.8542$). Se puede afirmar entonces que la aplicación de una combinación de modelos gaussianos adaptativos para la detección de movimiento presentó buenos resultados en la clasificación de motilidad progresiva.

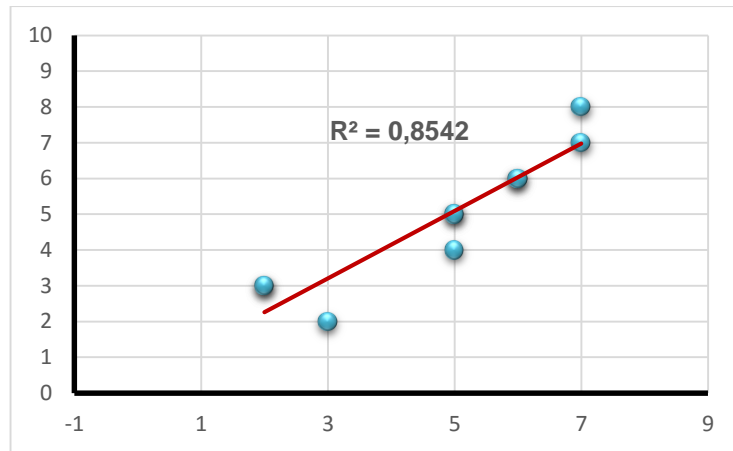


Figura 33. Regresión lineal y coeficiente de determinación de los espermatozoides en clasificados en A+B. Elaborada por el autor.

Sin embargo, se presentaron diferencias pronunciadas en la detección y clasificación de los espermatozoides pertenecientes a la categoría C, puesto que en la figura 34 se observa que la correlación fue baja ($R^2 = 0.3952$). Esto se debió principalmente a la dificultad de ajustar una sensibilidad mínima para considerar un espermatozoide en movimiento y no progresivo, ya que la inspección visual del experto es muy subjetiva y no existe un regimiento estricto, por ejemplo, cuando considerar un espermatozoide con desplazamiento casi nulo o con desplazamiento extremadamente corto.

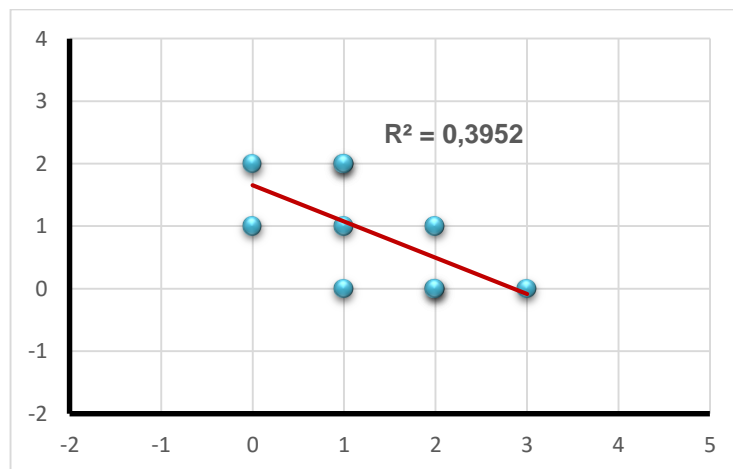


Figura 34. Regresión lineal y coeficiente de determinación de los espermatozoides en clasificados en C. Elaborada por el autor.

En la figura 35 se muestra la correlación para la detección de espermatozoides inmóviles. Se presentó una buena correlación ($R^2 = 0.9214$), por lo tanto, el análisis discriminante de Fisher, que fue el encargado de calcular los espermatozoides inmóviles tuvo un buen desempeño, como se puede evidenciar en la siguiente figura.

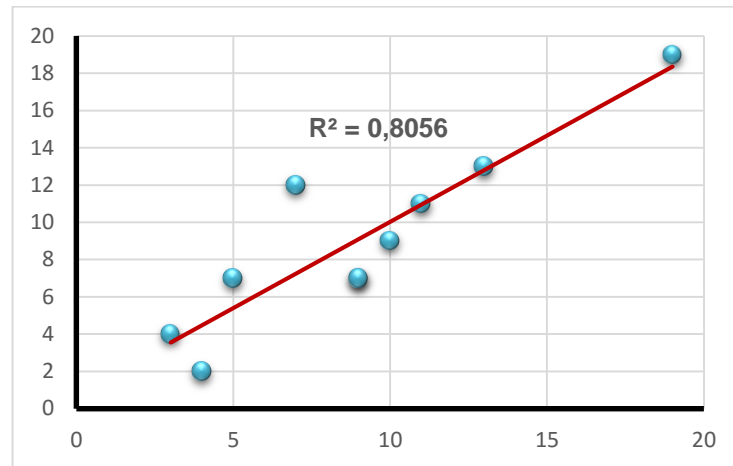


Figura 35. Regresión lineal y coeficiente de determinación de los espermatozoides en clasificados en D. Elaborada por el autor.

En términos generales, se puede afirmar que la etapa de clasificación de motilidad del algoritmo presentó buena concordancia con el análisis manual del experto, puesto que al juntar los datos de las tres clases podemos ver una correlación alta ($R^2 = 0,8754$) como se muestra en la figura 36.

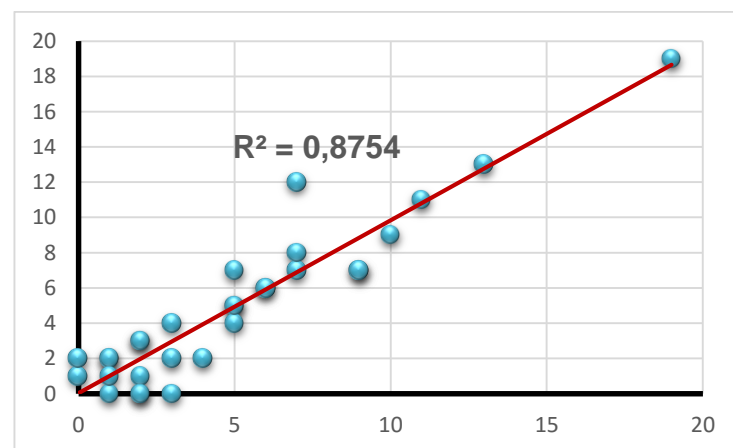


Figura 36. Regresión lineal y coeficiente de determinación de la clasificación de espermatozoides en todas las categorías. Elaborada por el autor.

Al igual que en el caso de la clasificación de la motilidad no progresiva (Categoría C), el cálculo del porcentaje de motilidad total (% MT) fue bastante variable entre el análisis manual y el análisis semi-automático, mostrándose en la figura 37.

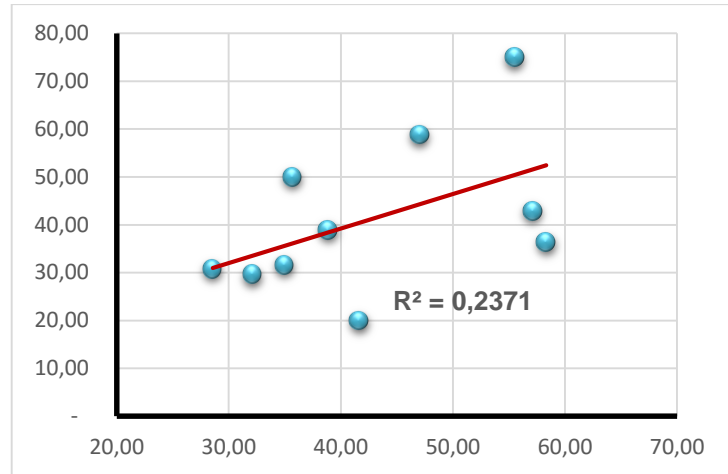


Figura 37. Regresión lineal y coeficiente de determinación del cálculo del porcentaje de motilidad total. Elaborada por el autor.

Esto es causado por la baja correlación presentada en la figura 34, donde al ser computados el número de espermatozoides en motilidad no progresiva con el número total es éstos, provoca también una baja concordancia en la validación de este parámetro. No obstante, como en los sistemas CASA convencionales, el parámetro de motilidad total no es el más significativo y nuestro algoritmo tiene como objetivo calcular el porcentaje de motilidad progresiva.

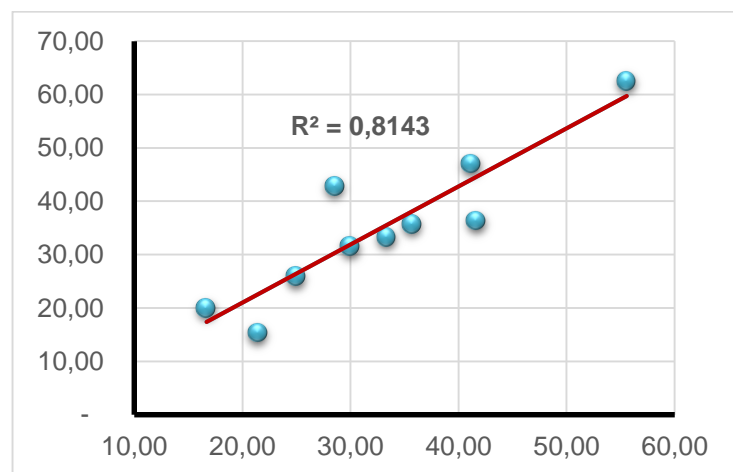


Figura 38. Regresión lineal y coeficiente de determinación del cálculo del porcentaje de motilidad progresiva. Elaborada por el autor.

Así pues, la validación de los resultados de la herramienta computacional diseñada con respecto al análisis por parte de un experto se evidencia en la figura 38. Se obtuvo un coeficiente de determinación alto ($R^2 = 0.8143$) debido principalmente a la alta concordancia en los datos de la categoría A+B, es decir, espermatozoides progresivos.

La mayor variación para esta correlación se presentó en video 4, donde se produce un error del 33.33 % con respecto al dato proporcionado por el experto. Esta desviación fue pronunciada debido a que esa muestra en particular contenía la menor concentración de espermatozoides de todas, por lo que cualquier dato equivocado implicaba un error significativo. No obstante, la correlación para las muestras con mayor concentración fue muy buena, con lo cual el algoritmo diseñado tiene mejores resultados con más espermatozoides analizados. Por tanto, es posible decir que la implementación de los algoritmos de Munkres y Kalman tuvieron un desempeño notable en el seguimiento de múltiples trayectorias presentes en los videos.

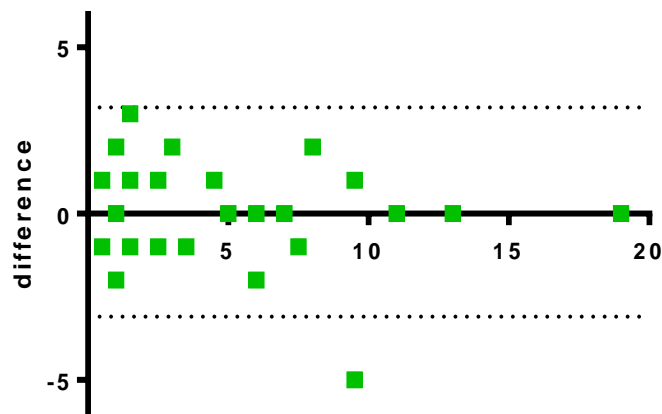


Figura 39. Comprobación Bland Altman para las clases A+B, C y D. Elaborada por el autor.

Finalmente, mediante la prueba de comprobación Bland-Altman mostrada en la figura 39, que representa la diferencia entre las metodologías implementadas, no se encontró variabilidad significativa entre la medida de motilidad espermática por parte del experto y la herramienta computacional implementada. Los resultados obtenidos muestran las diferencias de la medida manual con respecto al algoritmo con una relación de 0,966 en un intervalo de confianza del 95% de (3,058 – 3,19), indicando que las medidas fueron muy similares. Si las dos metodologías fueran iguales la razón esperada sería 1. En definitiva, los resultados obtenidos a partir de la herramienta mostraron una buena correlación y siguieron la tendencia de los resultados del experto, como se ilustra en la figura 40.

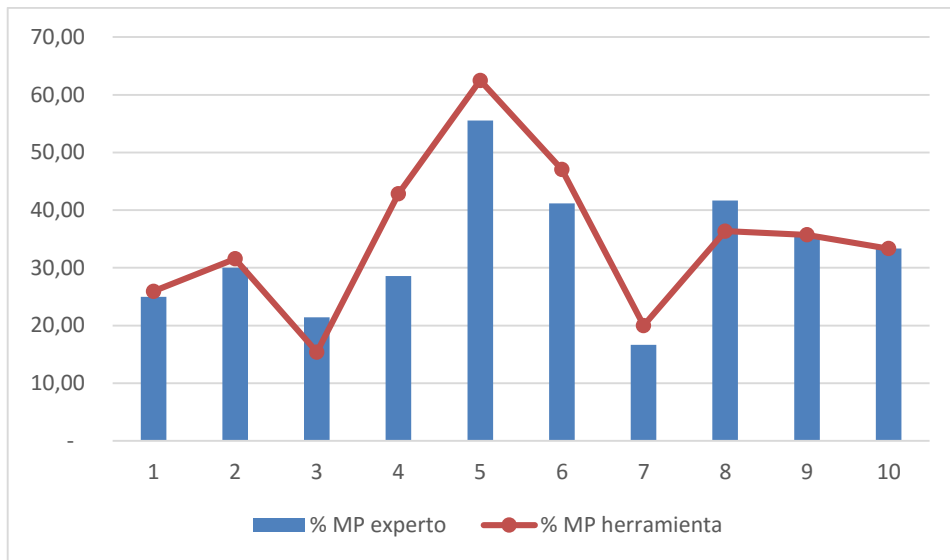


Figura 40. Comparación de los resultados de porcentaje de motilidad progresiva dados por el experto y por la herramienta computacional diseñada. Elaborada por el autor.

Conclusiones y trabajo futuro

A través del trabajo realizado queda demostrada la versatilidad de la visión artificial en procesos que requieran automatización. La aplicación en conjunto de diferentes técnicas de procesamiento digital de video permite desarrollar tecnologías muy efectivas que tengan la capacidad de reemplazar en gran medida la labor humana y por tanto eliminar ciertos inconvenientes intrínsecos de su intervención. La observación y evaluación de muestras seminales en bovinos llevadas a cabo por especialistas puede ser reemplazada por una máquina o herramienta computacional que además de proveer un análisis objetivo pueda brindar resultados de alta precisión y en cortos periodos de tiempo.

La colecta y preparación de las muestras seminales congeladas se llevaron a cabo de acuerdo a los protocolos establecidos en los laboratorios de medicina veterinaria. A su vez, la captura de imágenes se obtuvo mediante una configuración adecuada de iluminación y enfoque de un microscopio de contraste de fase, junto con una cámara digital adjunta al ocular del mismo. Esto permitió la correcta adquisición de videos y por lo tanto la obtención de un material con la capacidad de ser analizado por un computador. Aunque se presentaron algunos videos con objetos no deseados en sus fotogramas, se minimizaron significativamente mediante el procesamiento de imágenes posterior.

La aplicación del análisis de modelos gaussianos adaptativos permitió la detección de movimiento presente en la secuencia de fotogramas de los videos. Inicialmente se presentaron inconvenientes en esta etapa debido a los casos en los que ciertos espermatozoides se desenfocaban por el cambio de altura en la muestra, lo que provocaba en algunas instancias una carencia de segmentación. El ajuste de contraste permitió mejorar esta situación al facilitarle al método el reconocimiento de variación de valores de pixeles a lo largo de los fotogramas. Debido a la vibración del semen provocada por el movimiento oscilatorio de las colas de los espermatozoides se produjo una detección exagerada de partículas no correspondientes a los mismos, por lo tanto, se implementaron operadores morfológicos con el fin de eliminarlas en el campo de la imagen binaria. De esta forma, se obtuvieron únicamente las regiones pertenecientes a las cabezas espermáticas, las cuales a través de un análisis de componentes conectados se extrajeron de los valores de las coordenadas de posición en cada fotograma.

La reconstrucción de cada trayectoria espermática en el video fue posible gracias a la implementación del algoritmo de asignación de Munkres que tuvo la capacidad de asignar los nuevos centroides a las trayectorias existentes mediante el criterio de predicción y corrección del filtro de Kalman. Este filtro, cuyos valores iniciales fueron ajustados de manera experimental, funcionó correctamente en el rastreo individual de cada espermatozoide durante todo el periodo de procesamiento del video. Por su parte, el análisis discriminante de Fisher dio lugar a la segmentación de los espermatozoides inmóviles mediante una base de datos extraída experimentalmente. De esta forma, con el número de espermatozoides total y el número y atributos de espermatozoides en desplazamiento se procedió a desarrollar la etapa de análisis de motilidad individual.

La caracterización de los espermatozoides en movimiento se llevó a cabo teniendo en cuenta el estado del arte de los sistemas CASA en la evaluación de motilidad individual de espermatozoides de bovino previamente congelados y reactivados. A pesar de no contar con un protocolo único para los sistemas CASA, se adecuó uno para nuestro diseño. A través de la serie de coordenadas que describen la trayectoria de cada espermatozoide móvil en el video se consiguió calcular los parámetros de motilidad individual de acuerdo a los sistemas comerciales CASA. De la velocidad curvilínea se derivó un conjunto de atributos cinéticos que permitieron la clasificación de cada espermatozoide dentro de las categorías observadas en el estado del arte. Partiendo de la obtención de estos parámetros y la clasificación de motilidad, para cada video fue obtenido un porcentaje de motilidad progresiva, el cual fue sometido a validación teniendo como referencia la evaluación por parte de un técnico experto en andrología.

La validación con respecto a la observación y el análisis manual por parte de un técnico experto evidenció que existió una buena concordancia en la clasificación de motilidad progresiva. No obstante, la clasificación de la motilidad no progresiva se vio afectada fundamentalmente por la complejidad de la determinación objetiva de los límites de velocidad de los espermatozoides no progresivos o con desplazamiento casi nulo, puesto que la sensibilidad del algoritmo es muy diferente a la del ojo humano. Esto también afectó al análisis de motilidad total, pero los resultados mostraron seguir la línea de tendencia de los datos proporcionados por el experto. Por otra parte, el análisis de motilidad progresiva, el más significativo de un análisis seminal por parte de un sistema CASA, tuvo un coeficiente de determinación alto ($R^2 = 0.8143$) y, por lo tanto, una muy buena correlación.

De esta manera, se desarrolló una herramienta computacional semi-automática modular y configurable para el análisis asistido de motilidad espermática individual de muestras de semen bovino, que tiene la capacidad de eliminar los errores humanos en la interpretación del parámetro de motilidad y así proveer resultados confiables. En Colombia, esta herramienta brindaría una alternativa económica y accesible para el análisis, estudio e investigación en el campo de la medicina veterinaria. Futuros desarrollos involucran en primer lugar la minimización de los errores presentados en el presente trabajo, posteriormente el diseño de un sistema CASA completo con la capacidad de llevar análisis tanto de motilidad como de viabilidad, concentración y morfología espermática. Además, se propone a ajustar el diseño para brindar la opción al usuario de evaluar tanto muestras de semen bovino como de otras especies, como por ejemplo muestras humanas. Mientras tanto, este proyecto supone un aporte importante en el proceso de investigación en campos interdisciplinarios de la Universidad de Cundinamarca.

Bibliografía

Abbiramy, V. S., Shanthi, V. & Allidurai, C., 2010. Spermatozoa detection, counting and tracking in video streams to detect asthenozoospermia. *International Conference on Signal and Image Processing*, pp. 265-270.

ABC-finkeros, 2013. *El problema de la ganadería en Colombia*. [En línea]

Disponible en: <http://abc.finkeros.com/el-problema-de-la-ganaderia-en-colombia/> [Último acceso: 2015].

Amann, R. & Waberski, D., 2014. Computer-assisted sperm analysis (CASA): Capabilities and potential developments. *Theriogenology*, pp. 5-17.

Angelino, J., 2009. Manual de evaluación de semen en bovinos. En: Veracruz: s.n., pp. 21-49.

Arriaza, M., 2006. En: *Guía práctica de análisis de datos*. Córdoba: s.n., pp. 112-128.

Boryshpolets, S. y otros, 2013. Different computer-assisted sperm analysis (CASA) systems highly influence sperm motility parameters. *Theriogenology*, pp. 758-765.

Boyers, S., Davis, R. & Katz, D., 1989. Automated semen analysis. *Curr. Probl. Obstet. Gynecol. Fertil.*, Issue 12, pp. 172-200.

Burgeois, F. & Lasalle, J. C., 1971. An extension of the Munkres algorithm for the assignment problem to rectangular matrices. *Communications of the ACM*, 14(12), pp. 802-804.

Castellini, C., Bosco, A. D., Ruggeri, S. & Collodel, G., 2011. What is the best frame rate for evaluation of sperm motility in different species by computer-assisted sperm analysis?. *Fertility and Sterility*, Vol. 96, No. 1, pp. 24-27.

Contexto Ganadero, 2015. *Semen de toro, clave en producción de leche con más proteína y grasa*. [En línea]

Disponible en: <http://contextoganadero.com/internacional/semen-de-toro-clave-en-produccion-de-leche-con-mas-proteina-y-grasa> [Último acceso: 2015].

Contri, A. y otros, 2010. Effect of semen preparation on casa motility results in cryopreserved bull spermatozoa. *Theriogenology*, Volumen 74, pp. 424-435.

DANE, 2012. Insumos y factores asociados a la producción agropecuaria. pp. 1-4.

Elsayed, M., El-Sherry, T. & Abdelgawad, M., 2015. Development of computer-assisted sperm analysis plugin for analyzing sperm motion in microfluidic environments using Image-J. *Theriogenology* 84, pp. 1367-1377.

Escalante, B., 2006. Procesamiento digital de imágenes. pp. 1-6.

FEDEGÁN, 2013. *Análisis del inventario ganadero colombiano. Comportamiento y variables explicativas*, s.l.: s.n.

Fisher, R. A., 1936. The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, Volumen 7, pp. 179-188.

Fréour, T. y otros, 2010. Computer-Assisted Sperm Analysis (CASA) parameters and their evolution during preparation as predictors of pregnancy in intrauterine insemination with frozen-thawed donor semen cycles. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, pp. 186-189.

Gasque, R., 2008. Pág. 10: Reproducción bovina. En: *Enciclopedia bovina*. México DF: s.n., pp. 391-397.

Gonzalez, R. C. & Woods, R. E., 2008. Some basic relationships between pixels. En: *Digital Image Processing, 3rd Edition*. New Jersey: Prentice Hall, pp. 68-84.

Kalman, R. E., 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, Issue 82, pp. 35-45.

La Serna, N. & Román, U., 2009. Técnicas de segmentación en Procesamiento Digital de Imágenes. *Revista de Ingeniería de Sistemas e Informática*, 6(2), pp. 9-16.

Li, Q. y otros, 2012. Automatic human spermatozoa detection in microscopic video streams based on OpenCV. *5th International Conference on BioMedical Engineering and Informatics*, pp. 224-227.

Lu, J. C., Huang, Y. F. & Lu, N. Q., 2013. Computer-aided sperm analysis: past, present and future. *First International Journal of Andrology*, Issue 46, pp. 329-338.

Mahecha, L. & Gallego, L., 2002. Situación actual de la ganadería de carne en Colombia y alternativas para impulsar su competitividad y sostenibilidad. *Universidad de Antioquia*, pp. 213-225.

Maintz, T., 2005. Mathematical morphology. En: *Digital and Medical Image Processing*. Universiteit Utrecht: s.n., pp. 121-160.

Munkres, J., 1957. Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics*, 5(1), pp. 32-38.

Nossa, N. & Suarez, E., 2011. *Evaluación de fertilidad en toros mediante variables fisiológicas y reproductivas en tres fincas del municipio de Fusagasugá*, Fusagasugá: s.n.

Nurhadiyatna, A. y otros, 2014. Comparison and implementation of motion detection methods for sperm detection and tracking. -, pp. 1-5.

Palacín, I., Vicente-Fiel, S., Santolaria, P. & Yániz, J., 2013. Standardization of CASA sperm motility assessment in the ram. *Small Ruminant Research*, pp. 128-135.

Pascual-Gaspar, J. y otros, 2008. A simple and effective system for computer-assisted semen analysis. *Advances in Medical, Signal and Information Processing Conference*, pp. 1-4.

Ravanfar, M. R. & Moradi, M. H., 2011. Low contrast sperm detection and tracking by Watershed Algorithm and Particle Filter. *18th Iranian Conference on BioMedical Engineering*, pp. 260-263.

Roa, E. y otros, 2012. Evaluación asistida por computador de la viabilidad espermática en humanos. *Revista Ingeniería Biomédica*, pp. 17-28.

Salgado, L., 2007. *Visión artificial: fundamentos y aplicaciones*, Madrid: s.n.

Sati, L. & Huszar, G., 2015. Sperm motility and viability: Overview of the cellular and physiological aspects that support these functions. *European Medical Journal*, pp. 74-80.

Schleh, C. & Leoni, A.-L., 2013. How to optimize the benefits of computer assisted sperm analysis in experimental toxicology. *Journal of Occupational Medicine and Toxicology*, pp. 1-4.

- SERIDA, 2005. Análisis del semen bovino. *Tecnología agroalimentaria*, pp. 39-43.
- Simonik, O., Šichtař, J., Stádník, L. & Biniova, Z., 2015. Computer assisted sperm analysis - The relationship to bull field fertility, possible errors and their impact on outputs: A review. *Indian Journal of Animal Sciences*, 85(1), pp. 3-11.
- Stauffer, C. & W., G., 1999. Adaptive background mixture models for real-time tracking. *The Artificial Intelligence Laboratory, MIT*, pp. 1-7.
- Sundararaman, M., Kalatharan, J., Thilak, K. & Jawahar, P., 2012. Computer assisted semen analysis for quantification of motion characteristics of bull sperm during cryopreservation cycle. *Veterinary World*, 5(12), pp. 723-726.
- Tekalp, A. M., 2015. En: *Digital video processing. 2nd Edition*. s.l.:Prentice Hall, pp. 273-389.
- Tomlinson, M. y otros, 2010. Validation of a novel CASA system using multitarget-tracking algorithms. *Fertility and Sterility* 93, pp. 1911-1920.
- Vincent, P. y otros, 2012. Bovine semen quality control in artificial insemination centers. *Anim Reprod*, 9(3), pp. 153-165.
- Welch, G. & Bishop, G., 2006. An introduction to the Kalman Filter. *University of North Carolina*, pp. 1-16.
- Yang, H. F. y otros, 2014. Head tracking and flagellum tracing for sperm motility analysis. pp. 310-313.