Universidad Politécnica de Madrid
Escuela Técnica Superior de Ingeniería Agronómica,
Alimentaria y de Biosistemas.
Departamento de Ingeniería Agroforestal

**Extracción de Características en Viñedos Mediante Algoritmos de
Clasificación Difusa y Técnicas de Inteligencia Artificial**

Tesis Doctoral

Christian Correa Farías

Magister, Ingeniero Civil Agrícola

Madrid, 2016

Universidad Politécnica de Madrid
Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de
Biosistemas.
Departamento de Ingeniería Agroforestal

Extracción de Características en Viñedos Mediante Algoritmos de
Clasificación Difusa y Técnicas de Inteligencia Artificial

Tesis Doctoral

Doctorando:

Christian Correa Farías (Magister en Ingeniería Agrícola,
Ingeniero Civil Agrícola)

Director:

Constantino Valero Ubierna (Doctor Ingeniero Agrónomo)

Madrid, 2016

A mis padres, por su cariño y apoyo incondicional.

AGRADECIMIENTOS

Deseo agradecer la financiación del presente trabajo a la COMISIÓN EUROPEA a través del proyecto RHEA. “Robot Fleets for Highly Effective Agriculture and Forestry Management” del 7 Programa Marco (proyecto nº 245986). Se agradece la contribución de todos los participantes en el proyecto: Agencia Estatal Consejo Superior de Investigaciones Científicas – CSIC-CAR (Centro de Automática y Robótica, Instituto de Ciencias Agrarias, Instituto de Agricultura Sostenible), CogVis GmbH, Forschungszentrum Telekommunikation Wien Ltd., Cyberbotics Ltd, Università di Pisa, Universidad Complutense de Madrid, Tropical, Soluciones Agrícolas de Precisión S.L., Universidad Politécnica de Madrid - UPM (ETS Ingenieros Agrónomos, ETS Ingenieros Industriales), AirRobot GmbH & Co. KG, Università degli Studi di Firenze, Institut national de recherche en sciences et technologies pour l'environnement et l'agriculture -IRSTEA, CNH Belgium NV, CNH France SA, Bluebotics S.A. and CM Srl.

Este trabajo ha sido llevado a cabo en el seno del grupo de investigación LPF_TAGRALIA de la Universidad Politécnica de Madrid, y también ha recibido financiación parcial del proyecto de investigación “VITISENSOR: integración de tecnologías avanzadas de detección en una plataforma móvil multisensor para el estudio de la variabilidad espacio-temporal del viñedo” AGL2011-23673 perteneciente al Plan Nacional de I+D y concedido por el MICCIN a la Universidad de La Rioja.

INDICE DE CONTENIDOS

I.	RESUMEN.....	19
II.	ABSTRACT	21
III.	ESTADO DEL ARTE Y DEFINICIÓN DEL PROBLEMA.....	23
	• 1.Caracterización del viñedo	23
	• 2.Sensores y tecnologías de imagen.....	26
	• 3.Metodologías de análisis y clasificación.....	30
IV.	HIPÓTESIS	33
V.	OBJETIVOS.....	34
1.	Objetivo Generales	34
2.	Objetivos Específicos.....	34
VI.	METODOLOGIA.	35
	Síntesis metodológica y relación con trabajos publicados.....	35
	Trabajos publicados:	37
1.	PROCEDIMIENTOS DE ADQUISICIÓN DE IMÁGENES... ..	38
1.1.	Primera campaña de adquisición	38
1.1.1.	Adquiridas de modo estático.	38
1.2.	Segunda campaña de adquisición	38
1.2.1.	Adquiridas de modo dinámico.	38
1.2.2.	Tercera campaña de adquisición	41
1.2.3.	Adquiridas de modo dinámico.	41
2.	TÉCNICAS DE PROCESAMIENTO EMPLEADAS.....	45
A.	Algoritmos de Clasificación Supervisados.....	45
B.	Algoritmos de Clasificación No Supervisados.	45
C.	Métodos de identificación o etiquetado de grupos.....	46
2.1.	Primera campaña	47

2.1.1. Procedimiento aplicado.	47
2.1.2. Método de Validación.	48
2.1.3. Selección del espacio de color.	48
2.1.4. Métodos de Agrupamiento no supervisados	49
2.1.4.1. Algoritmo de agrupamiento difuso	49
2.1.4.2. Agrupamiento mediante el algoritmo K-means:	51
2.1.4.3. Segmentación basada en Delaunay.	52
2.1.5. Métodos de Agrupamiento supervisados.	52
2.1.5.1. Agrupamiento mediante el algoritmo Vecinos más Cercanos:	52
2.1.5.2. Agrupamiento mediante la distancia de Mahalanobis.	53
2.1.6. Detección a través de la forma.	53
2.1.6.1. Transformada circular de Hough.	53
2.1.6.2. Transformada de Simetría Radial.	54
2.1.7. Detección a través de Análisis de textura.	54
2.1.7.1. Wavelet de Haar.	55
2.1.7.2. Segmentación de texturas.	55
2.2. Segunda Campaña de adquisición	57
2.2.1. Procedimiento Aplicado.	57
2.2.2. Imágenes obtenidas de forma estática.	57
2.2.3. Imágenes obtenidas de forma Dinámica.	57
2.2.3.1. Análisis Imágenes Infrarrojas.	57
2.2.3.2. Análisis Imágenes visibles.	58
2.2.3.3. Selección de algoritmos de clasificación difusos	58
2.2.3.4. Algoritmos de agrupamiento difuso evaluados	59

2.2.3.4.1. Fuzzy C-Means (FCM).....	59
2.2.3.4.2. Algoritmo Possibilistic C-Means (PCM).....	60
2.2.3.4.3. Fuzzy Possibilistic C-Means.....	61
2.2.3.4.4. Robust Fuzzy-Possibilistic C-Means.....	63
2.2.3.4.5. Gustafson-Kessel Clustering	65
2.2.3.4.6. Fuzzy c-Mean con Restricciones Espaciales y Contornos de Regiones.	66
2.2.3.4.7. Kernelled Robust Fuzzy C-means	68
2.3. Tercera campaña de adquisición	69
2.3.1. Estrategia de Aceleración de Convergencia e Identificación de Grupos.	69
2.3.1.1. Clasificador basado en Máquinas de Vectores de Soporte:	71
2.3.1.2. Creación de Mosaicos Mediante Transformada de Características de Escala Invariante (SIFT).	72
2.3.1.3. Descripción del marco metodológico final.	72
2.4. Cámara de Profundidad.....	76
2.4.1. Síntesis de la solución final planteada.....	76
2.4.1.1. Validación del Algoritmo.	79
3. PREDICCIÓN DE RENDIMIENTO Y DE ÁREA FOLIAR. ..	80
3.1. Sección experimental	80
3.2. Sitio experimental	80
3.3. Defoliación, raleo de racimos y la evaluación del área foliar y fruta eliminados	80
3.4. Procesamiento de Imágenes (Algoritmo de clústeres)	82
3.5. Implementación del software.....	85

3.6. Etapas de la metodología de extracción de características.....	87
3.7. Validación del algoritmo	88
3.8. Correlaciones, estimación del área de hojas y rendimiento	88
VII. RESULTADOS	90
1. Primera campaña de adquisición	90
1.1. Métodos de clasificación por color	90
1.1.1. Selección del espacio de color.	90
1.1.2. Métodos de agrupamiento por color.	93
1.1.2.1. K-Means	93
1.1.2.2. FCM	97
1.1.2.3. Desempeño de los clasificadores.....	100
1.2. Métodos de clasificación por forma.	102
1.2.1. Transformada circular de Hough	102
1.2.2. Transformada de Simetría Radial (TSR)	102
1.3. Análisis de textura.	104
1.3.1. Estadísticos de segundo orden para caracterizar textura.	104
1.3.2. Wavelet de Haar.....	107
1.3.3. Segmentación de texturas.....	110
1.4. Métodos de agrupamiento supervisados.....	112
1.4.1. Vecinos más cercanos.	112
1.5. Métodos de agrupamiento no supervisados.....	117
1.5.1. Segmentación basada en Delaunay.	117
2. Segunda campaña de adquisición.....	119
2.1. Índices multiespectrales.	119

2.2. Métodos de clasificación de agrupamiento difusos...	120
2.2.1. Fuzzy c-mean con Restricciones Espaciales y Contornos de Regiones.	120
2.2.2. Kernelled Robust Fuzzy C-means.....	122
2.3. Evaluación cuantitativa de la Segmentación.....	123
2.3.1. Errores de Clasificación del método seleccionado, FCM-GK.	127
3. Tercera campaña de adquisición.	129
3.1. Resultados de la estrategia de aceleración de convergencia.....	129
3.1.1. Desempeño de la clasificación.....	129
3.1.2. Tiempos de ejecución.....	136
3.1.3. Resolución.....	136
3.1.4. SIFT.....	137
3.1.5. Identificación de grupos mediante redes neuronales.	139
3.1.6. Identificación de grupos mediante SVM.....	139
3.2. Cámara de profundidad.....	139
3.2.1. Discusión comparada de resultados.....	144
3.3. Resultados predicción de rendimiento y área foliar .	151
3.3.1. Validación del algoritmo.....	151
3.3.2. Desempeño del clasificador.....	153
3.3.3. Estimación de rendimiento a cosecha y área foliar ...	158
VIII. CONCLUSIONES.....	162
1. Conclusiones Predicción de rendimiento.....	162
2. Conclusiones Métodos De Agrupamiento.....	163
3. Conclusiones Estrategia de Aceleración.....	164

4.	Conclusiones Finales	165
IX.	TRABAJOS FUTUROS.....	166
X.	REFERENCIAS	167

ÍNDICE DE FIGURAS

Figura 1. Remolque empleado para la adquisición de las imágenes, dispone de una cama de video, una cámara IR, un DGPS y un computador portátil para el almacenamiento de la información.	39
Figura 2. Esquema de la plataforma de adquisición empleada en los ensayos de terreno. Dispone, además de las cubiertas para evitar los efectos de la luz, de un sistema de estabilización mecánico y electrónico diseñado para mantener la cámara de forma paralela a la hilera.....	41
Figura 3. Sistema de posicionamiento y estabilización de la cámara, accionado por dos servo motores y controlados desde un microcontrolador.....	43
Figura 4. Plataforma empleada en los ensayos de terreno, en la que se aprecia parcialmente el sistema de cobertura para evitar los efectos de las variaciones en iluminación. a) Vista frontal b) vista en perspectiva.	44
Figura 5. Esquema del marco metodológico propuesto para el análisis de imágenes RGB. ...	72
Figura 6. Diagrama de flujo final de la metodología de análisis de imágenes RGB y de Profundidad propuesta.	78
Figura 7. a) Selección aleatoria de 2000 píxeles de una imagen típica de una vid (imagen I0 en el espacio de color RGB). b) plano RG. Elipses representan puntos equidistantes. c) plano RB. d) Plano GB.	85
Figura 8. Captura de pantalla del proceso de selección manual de los píxeles de referencia.	86
Figura 9. Imagen de una vid en el espacio de color RGB y sus canales.	91
Figura 10. Imagen de una vid en el espacio de color HSV y sus canales.....	91
Figura 11. Imagen de una vid en el espacio de color Otha y sus canales de color.	92
Figura 12. Imagen de una vid en el espacio de color $L^*a^*b^*$ y sus canales.	92
Figura 13. Imagen de una vid descompuesta en los cuatro canales del espacio de color CMYK.	93
Figura 14. Clases generadas por K-Means usando la Distancia del Coseno sobre canales a^*b^*	95
Figura 15. Resultado de aplicar K-Means sobre la Clase 2 de la Figura 14. Nótese que las imágenes b), c) y d) contienen hojas, mientras que la imagen e) contiene segmentos de sarmientos.....	96
Figura 16. a) Imagen original. b) Grupo racimos generado por K-Means empleando la distancia euclidiana y ocho grupos.	96

Figura 17. Resultados obtenidos con los diversos métodos de clasificación. Encerrado en los círculos se observan zonas mal clasificadas por Fuzzy C-Means.	99
Figura 18. Transformada circular de Hough sobre el canal b^* . Busca círculos de radio entre 3 y 9 píxeles, sin tolerancia a incluir círculos concéntricos.	102
Figura 19. a) Imagen original en escala de grises. b) Mapa de simetría para búsqueda de simetría en dos radios (4 y 5 píxeles). c) Zonas con mayor índice de simetría. d) mapa de simetría para búsqueda de un sólo radio (5 píxeles).....	103
Figura 20. Imágenes de textura generadas. a) Imagen original. b) Espacio de color $L^*a^*b^*$, filtro entropía local. c) Espacio de color $L^*a^*b^*$, filtro de desviación estándar local. d) Espacio de color $L^*a^*b^*$, filtro del rango. e) Espacio de color RGB, filtro entropía local. f) Espacio de color RGB, filtro de desviación estándar local. g) Espacio de color RGB, filtro del rango. h) Espacio de color HSV, filtro de entropía local. i) Espacio de color HSV, filtro de desviación estándar. j) Espacio de color HSV, filtro del rango. k) Espacio de color XYZ, filtro de entropía local.	106
Figura 21. a) Imagen original, b) Imagen comprimida mediante Wavelet de Haar de tres niveles aplicada sobre la imagen original a). c) Píxeles identificables como uva.....	107
Figura 22. Tres niveles de descomposición e imágenes de detalles en tres orientaciones. .	108
Figura 23. A la izquierda, imagen de un racimo en envero espacio de color RGB. A la derecha imagen comprimida a nivel de descomposición 5, mediante Wavelet de Haar.	108
Figura 24. Imagen de un racimo en envero, el espacio de color RGB, con distintos niveles de compresión mediante Wavelet de Haar.	109
Figura 25. A la izquierda, Imagen de un racimo en envero, el espacio de color HSV. A la derecha imagen después de la compresión mediante Wavelet de Haar.....	109
Figura 26. A la izquierda, Imagen de un racimo en envero, el espacio de color $L^*a^*b^*$. A la derecha imagen después de la compresión mediante Wavelet de Haar.	110
Figura 27. a) Imagen original en falso color de un segmento de la Figura 20 a). b) Vector de características para escala de la textura. (Imagen segmentada). c) Vector de características para la orientación de la textura en el eje x d) Vector de características para la orientación de la textura en el eje y. e) Vector de características para la magnitud de la textura. f) Vector de características para la imagen de intensidad.....	111
Figura 28. Ejemplo de la interface de usuario creada para capturar áreas con píxeles representativos para cada clase. Encerrado en líneas rojas, las áreas seleccionadas por el usuario como píxeles de referencia para cada clase.	113
Figura 29. Resultados de la clasificación mediante el método los vecinos más cercanos. a) Píxeles pertenecientes a la clase Hojas Viejas. b) Píxeles pertenecientes a la clase Sarmientos. c) Píxeles pertenecientes a la clase Fondo 1. d) Píxeles pertenecientes a la clase Fondo 2. e) Píxeles pertenecientes a la clase Racimo. f) Píxeles pertenecientes a la clase	

Racimo 2. g) Píxeles pertenecientes a la clase Tronco. h) Píxeles pertenecientes a la clase Hojas Nuevas.	116
Figura 30. Impresión de pantalla del software donde se muestra el resumen de los porcentajes de cada clase generada.	117
Figura 31. a) Imagen Original. b) Imagen en escala de grises agrupada mediante el método de Delaunay. c) Imagen en color como la original con los 64 grupos generados por el método de Delaunay. d) Imagen falso color con grupos fusionados manualmente.	118
Figura 32. Índices espectrales umbralizados generados partir de la imagen infrarroja.....	119
Figura 33. En a) se observa la imagen original, de b) a h) corresponden a las diversas clases generadas.....	120
Figura 34. a) Algoritmo ejecutado con 200 iteraciones. b) Algoritmo ejecutado con 400 iteraciones. Encerrado en amarillo una zona de mal clasificada.....	121
Figura 35. a) Imagen original de 320x240. b) Resultado de la clasificación mediante Kernelled Robust Fuzzy C-Means.....	122
Figura 36. Imágenes generadas por la segmentación por agrupamiento, empleando ocho clases sobre los canales a^* , b^* del espacio de color $L^*a^*b^*$, con resoluciones de 640x480 y 320x240.	125
Figura 37. Ejemplos de gráficos de los grupos y centros de grupos generados a partir del agrupamiento de píxeles sobre el plano $a^* b^*$ usando ocho grupos.	126
Figura 38. a) Imagen original, con baja presencia de uva. c) Grupo Uvas creado al aplicar el algoritmo de clasificación una vez. c) Grupo uvas al aplicar el clasificador sobre la imagen b).....	128
Figura 39. Partición por agrupamiento, realizada sobre el plano a^*b^* del espacio de color L^*b^* Las marcas grandes en negro marcan los centros de cada grupos o prototipos de los grupos.	129
Figura 40. a) Imagen original. b) Grupos generados por FCM-GK. c) Imagen de grupos generado por K-means con inicialización aleatoria. d) Imagen de grupos generado por K-means mediante la estrategia de inicialización.	131
Figura 41. a) Imágenes originales en RGB. b) Grupo Hojas viejas. c) Grupo Hojas Nuevas. d) Grupo Hojas Secas. e) Grupo Fondo. f) Grupo Uva.	134
Figura 42. a) Grupo Uvas proveniente de K-means con inicialización aleatoria. b) Grupo Uvas proveniente de K-means con inicialización al azar cuando se ejecuta el algoritmo por segunda vez. c) Grupo Uva proveniente de FCM-GK. d) Grupo Uva proveniente de K-means inicializado con los centroides FCM-GK.....	135

Figura 43. a) Mosaico creado por el algoritmo basado en SIFT. La imagen muestra la unión mediante la superposición de tres imágenes que producen un mosaico 3 m de longitud. b) Regiones de interés usadas para acelerar la velocidad de convergencia de SIFT.	138
Figura 44. Captura de pantalla del software de adquisición de imágenes desarrollado para la cámara Kinect. En rojo objetos más cercanos.....	140
Figura 45. Imagen usando como mascara los objetos más cercanos, marcados en rojo en la Figura 44 derecha.	141
Figura 46. Grupos resultantes del proceso de agrupamiento con K-means inicializado con FCM-GK.....	142
Figura 47. Grupo Hojas resultantes del proceso de fusión de grupos realizado por SVM....	143
Figura 48. a) Grupo identificado como uva. b) imagen binaria del grupo Uva antes de las operaciones morfológicas y de filtrado. c) Imagen binaria después de las operaciones morfológicas y del proceso de filtrado (excentricidad entre 0.70-0.98). d) Grupo Uva enmascarado con la imagen binaria filtrada.	143
Figura 49. Resultados de clasificación para diversos métodos basados en color publicados por Liu et al. (2013). a) Imagen original. b) Resultados histograma 3D. c) resultados umbralización RGB. d) resultados FCM-GK/SVM. e) Resultados grupo uvas usando FCM sobre canales $a^* b^*$ e identificado mediante SVM. f) Resultados FCM/SVM como imagen de los grupos resultantes. g) Grupo tronco y sarmientos generado por FCM/SVM. h) Grupo hojas generado por FCM e identificado mediante SVM.	145
Figura 50. a) Imagen original en condiciones normales de iluminación. b) Grupo identificado como uvas antes de eliminar pequeñas regiones. c) Imagen de Fernández et al. (2013) en condiciones de terreno. d) Imagen de Reis et al. (2012).	147
Figura 51. a) Imagen original. b) imagen de profundidad. En rojo se observa la hilera paralela. c) imagen de profundidad con valores menores a 1,7 m. d) imagen original enmascarada usando la información de profundidad (distancia menor 1,7 m).	150
Figura 52. Validación manual de la clasificación. a) Imagen original, donde 73,1% son píxeles de uva. b) Imagen de los grupos creados. Píxeles azules corresponden a la clase Uva, y representan un 72% de la imagen. c) Imagen original, donde el 58% de los píxeles son de hojas. d) Imagen de los grupos creados. Los píxeles negros corresponden a hojas, y representan un 63% de la imagen.....	151
Figura 53. a) Detalle de la imagen original. b) Detalle de las uvas identificadas antes de las operaciones morfológicas.	152
Figura 54. Ejemplo de clasificación realizada sobre una imagen de una vid, no desojada (imagen I0) y no raleada. (a) Imagen original. (b) Imagen de grupos. (c) Clase Tronco. (d) Clase Fondo. (e) Clase Uva, sin operaciones morfológicas. (f) Clase Uva después de las operaciones morfológicas. (G) área ampliada en la parte (e) que muestra pequeños grupos de píxeles mal clasificado. (h) área ampliada en la imagen (f) que muestra cómo las	

operaciones morfológicas eliminan los pequeños grupos de píxeles. (I) Clase Hojas viejas grado 1. (j) Clase Hojas viejas grado 2. (k) Clase Hojas jóvenes grado 1. (l) Clase Hojas jóvenes grado 2. 154

Figura 55. Ejemplo de clasificación, realizado sobre una imagen de la vid después de la primera etapa de la defoliación (I1) para mostrar la capacidad del algoritmo para separar entre diferentes tipos de hojas (a) imagen original. (B) Imagen de grupos. (C) la clase de Uva, sin operaciones morfológicas. (D) Clase Tronco. (E) Clase Hojas viejas de grado 1. (f) Clase Hojas viejas grado 2. (g) Clase Hojas jóvenes grado 1. (h) Clase Hojas jóvenes grado 2. 156

Figura 56. Ejemplo de clasificación realizada sobre una imagen de vid después de la tercera etapa de defoliación (imagen I5). a) Imagen original. b) Imagen de grupos. c) Clase de Uva. d) Fondo brillante. e) Fondo oscuro. f) Clase Tronco. g) Clase Hojas jóvenes. h) Clase Hojas viejas..... 157

Figura 57. (a) Correlación entre el área foliar real ($m^2 \text{ vid}^{-1}$) de la cubierta en las imágenes del grupo de entrenamiento, y el número de píxeles correspondientes a la clases Hoja calculado por el método de clasificación y análisis de imágenes. (b) Comparación entre los valores reales de área foliar ($m^2 \text{ vid}^{-1}$) provenientes de las imágenes del grupo de validación, y los valores estimados de área de hojas calculados con la ecuación de correlación en (a).g La línea punteada es la línea 1:1..... 158

Figura 58. (a) Correlación entre el rendimiento de las vides (kg) obtenidos a partir de las imágenes del grupo de entrenamiento, y el número de píxeles correspondientes a la clase de uva, calculada por el método de clasificación y análisis de imágenes. (B) Comparación entre los valores reales de rendimiento (kg) de las vides de las imágenes del conjunto de validación, y los valores de rendimiento estimados, calculados con la ecuación de correlación en (a). La línea punteada representa la línea 1:1..... 160

INDICE DE TABLAS

Tabla 1. Velocidades reales (m s^{-1}) para cada marcha seleccionada en el tractor y tiempos de integración establecidos para ambas cámaras (- combinación no usada).....	40
Tabla 2. Descripción de las etapas de defoliación y raleo, durante la sesión de adquisición de imágenes de una vid individual	81
Tabla 3: Desempeño de K-Means y FCM como clasificadores.	100
Tabla 4. Resultados del método de los vecinos más cercanos aplicados don dos distancias.	113
Tabla 5. Resultados del método <i>spatial fuzzy clustering with level set</i>	122
Tabla 6. Exactitud de las técnicas de agrupamiento, con y sin ruido. Desempeño para $m=2$	123
Tabla 7. Tiempo de ejecución, en segundos, para las técnicas de agrupamiento empleadas, a diferentes resoluciones.	126
Tabla 8. Tiempos de ejecución a diferentes resoluciones para la estrategia de aceleración de la convergencia.....	136
Tabla 9. Porcentajes de aciertos a dos resoluciones.....	136

INDICE DE ECUACIONES

Ecuación 1. Función objetivo FCM en términos de la distancia.	50
Ecuación 2. Distancia inducida por norma en FCM.....	50
Ecuación 3. Función objetivo K-means.....	51
Ecuación 4. Índice de vegetación de diferencia normalizada.	58
Ecuación 5. Índice de vegetación ajustado al suelo.....	58
Ecuación 6. Función objetivo de FCM en términos de los vectores x y v	60
Ecuación 7. Cálculo de centros de los grupos en FCM.....	60
Ecuación 8. Cálculo del grado de pertenencia en FCM.....	60
Ecuación 9. Función objetivo en PCM.	61
Ecuación 10. Cálculo de tipicalidad en PCM	61
Ecuación 11. Cálculo de los centroides en PCM.	61

Ecuación 12. Cálculo del coeficiente γ en PCM.	61
Ecuación 13. Función objetivo para FPCM.....	62
Ecuación 14. Cálculo de la tipicidad en FPCM.	62
Ecuación 15. Cálculo de los centroides en FPCM.....	62
Ecuación 16. Cálculo de distancia euclidiana en FPCM.....	62
Ecuación 17. Producto interno expresado a través del núcleo Mercer.....	63
Ecuación 18. Distancia euclidiana expresada en el espacio de las características en RFPCM.	63
Ecuación 19. Función objetivo del algoritmo RFPCM.....	63
Ecuación 20. Cálculo de distancia en el espacio de las características en RFPCM.	64
Ecuación 21. Función objetivo expresada en términos del kernel de Mercer.....	64
Ecuación 22. Grado de pertenencia en RFPCM.	64
Ecuación 23. Cálculo de tipicidad en RFCM.....	64
Ecuación 24. Cálculo de centroides en RFPCM.....	65
Ecuación 25. Cálculo de los centroides en FCM-GK.	66
Ecuación 26. Matriz de covarianza del grupo i	66
Ecuación 27. Cálculo del cuadrado de la distancia en FCM-GK.....	66
Ecuación 28. Calculo del grado de pertenencia en FCM-GK.....	66
Ecuación 29. Función espacial añadida a la función objetivo en el cálculo del grado de pertenencia.	67
Ecuación 30. Cálculo del grado de pertenencia en FCM con restricciones espaciales.....	67
Ecuación 31. Función objetivo a minimizar en KRFCM.	68
Ecuación 32. Restricciones a la función de pertenencia en KRFCM.....	68
Ecuación 33. Grado de pertenencia en KRFCM.	69
Ecuación 34. Cálculo de los centroides en KRFCM.....	69
Ecuación 35. Distancia de Mahalanobis	83
Ecuación 36. Matriz de covarianza.	83
Ecuación 37. Porcentaje de acierto en la clasificación de área de racimos.	100

Ecuación 38. Porcentaje de racimos correctamente clasificados.	100
Ecuación 39. Cálculo de la exactitud de la clasificación de los métodos difusos.	123

I. RESUMEN

Resulta vital para los enólogos contar con información que les permita estimar la producción antes de la cosecha, así como el área foliar de las vides. Esto les permite tomar decisiones relativas al manejo del riego y de otras prácticas culturales tendientes a alcanzar los objetivos de producción y calidad deseados.

Si bien es cierto que existen técnicas que permiten hacer esto manualmente, dichas técnicas son destructivas, ya que requieren el deshoje y cosecha de todos los racimos presentes en una vid, y además tienden a ajustarse hacia rendimientos históricos (Clingeffer *et al.*, 2001). Esto es costoso, laborioso y poco representativo de la escena en su conjunto, debido a la limitada cantidad de vides que pueden ser deshojadas y cosechadas con éste propósito (Nuske *et al.* 2011a).

Como alternativa a estas técnicas destructivas, se encuentran técnicas no invasivas como por ejemplo las basadas en procesamiento de imágenes (Dunn & Martin, 2004; Rabatel *et al.*, 2007; Meunkaewjinda *et al.*, 2008). Dichas técnicas presentan ventajas tales como: bajo costo de implementación, alta velocidad en la obtención de resultados y alta representatividad ya que no solamente es posible tomar muestras puntuales, sino que se puede fotografiar el viñedo en su totalidad (Diago *et al.*, 2012).

Sin embargo, la mayor parte de los trabajos publicados no están orientados a proveer resultados en tiempo real y cuando lo hacen, se concentran en determinar una sola característica a la vez y en limitadas condiciones de aplicación.

Por estos motivos en este trabajo se desarrolló una metodología tendiente a extraer características de la cubierta vegetal del viñedo, de forma automática robusta, mediante algoritmos de clasificación difusa y técnicas de inteligencia artificial.

Para ello se desarrolló un sistema autónomo de caracterización de la cubierta vegetal en viñedos, empleando una combinación de imágenes visibles (RGB), de profundidad (D) e infrarrojas (IR). Lo que permite estimar el área foliar, estimar el rendimiento empleando el área de uva visible, además de caracterizar las hojas en función de sus distintos grados de vigor.

Esta metodología se implementa en 10 pasos: 1) Adquisición de imágenes; 2) transformación de las imágenes al espacio de color $L^*a^*b^*$; 3) Agrupamiento Etapa I: segmentación de imagen a través de *Fuzzy C-Means* con Gustafson Kessel (FCM-GK); 4) Agrupamiento Etapa II: emplea las salidas FCM- GK como valores iniciales para *K-Means*; 5) Identificación de los grupos (etiquetado) generados por *K-Means* utilizando una Máquina de Soporte Vectorial; 6) Fusión de grupos con igual etiqueta; 7) Proceso de filtrado: realiza operaciones morfológicas sobre los grupos uva y hojas con el fin de llenar agujeros y eliminar pequeños grupos de píxeles; 8) Calcular las áreas de cada grupo; 9) Crear una imagen mosaico mediante *Scale-Invariant Feature Transform* para evitar el traslape de imágenes; 10) Corrección de las áreas de las hojas y las uvas.

Esta metodología fue aplicada sobre un total de 1500 imágenes capturadas durante tres campañas de adquisición, lográndose una precisión en la clasificación de un 95 % para uva, y de un 100% para hojas.

Esta metodología permite procesar una imagen de 320x240 píxeles en menos de 100 ms, lo que combinado con los elevados porcentajes de aciertos, frente a diferentes condiciones del viñedo, constituye sin duda una mejora significativa en el estado del arte, en lo que a clasificación e identificación robusta de características en terreno se refiere.

Ventajosamente esta metodología entrega en tiempo real más de una característica de modo simultáneo, lo que permite en una sola pasada determinar el área de sarmientos, el área de hojas de diverso tipos, así como el área de uva presente, lo que abre las puertas a diversas aplicaciones como predicción de rendimiento, aplicación de plaguicidas inteligente, manejo de la poda, planificación de riego y deshoje.

II. ABSTRACT

It is vital for winemakers have information that allows them to estimate production before harvest as well as the leaf area of vines. This allows them to make decisions regarding to irrigation management and other cultural practices aimed at achieving the desired production and quality objectives.

Although it is true that there are techniques to do this manually, such techniques are destructive as they require defoliation and the harvest of all clusters present on a vine, and furthermore tend to fit to historical yields (Clingeleffer *et al.*, 2001). This is costly, laborious and unrepresentative of the scene as a whole, due to the limited number of vines that can be defoliated and harvested for this purpose (Nuske *et al.* 2011a).

As an alternative to these destructive techniques are non-invasive techniques, such as those based on image processing (Dunn & Martin, 2004; Rabatel *et al.*, 2007; Meunkaewjinda *et al.*, 2008). Such techniques have advantages such as low cost of implementation, high speed in obtaining results and high representativeness, since it is not only possible to take spot samples, but can photograph the vineyard as a whole (Diago *et al.*, 2012.).

However, most of the published works are not designed to provide real-time results and when they do, they focus on determining one feature at a time and in restricted application conditions.

For these reasons, in this paper, a methodology was developed aimed at extracting features of vegetation cover of the vineyard, automatically, robust, using fuzzy classification algorithms and artificial intelligence techniques.

For this purpose an autonomous characterization system of plant cover in vineyards was developed using a combination of visible images (RGB), depth (D) and infrared (IR). Allowing estimate leaf area, estimate yield using visible grape area, as well as leaves characterization according to their degrees of vigor.

This methodology is implemented in 10 steps: 1) Images acquisition; 2) Transformation of the images to the $L^*a^*b^*$ color space ; 3) Grouping Stage I: image segmentation by Fuzzy C-Means with Gustafson Kessel (FCM-GK) algorithm; 4) Grouping Stage II: Take FCM-GK outputs as initial values for K-Means; 5) Identification of groups (labeled) generated by K-Means using Support Vector Machine; 6) Merger of groups with the same label; 7) Filtering process: morphological operations performed on grape leaves and groups in order to fill holes and remove small groups of pixels; 8) Calculate the areas of each group; 9) Create a mosaic image using scale-invariant feature transform to avoid overlapping of images; 10) Correction of the areas of leaves and grapes.

This methodology was applied to a total of 1500 images captured during three seasons acquisition, achieving a classification accuracy of 95% for grapes, and 100% for leaves.

This methodology allows to process an image of 320x240 pixels in less than 100 ms, which combined with high percentages of correct identification, against different conditions of the vineyard, is undoubtedly a significant improvement in the state of the art, as far as, on site, robust features classification and identification refers.

Advantageously, this methodology provided, in real time, more than one feature simultaneously, allowing on a single pass to determine the area of branches, the area of different types of leaf, as well as the area of grapes present, which opens the door to various applications such as yield prediction, intelligent pesticides application, pruning management, irrigation scheduling and defoliation.

III. ESTADO DEL ARTE Y DEFINICIÓN DEL PROBLEMA.

1. Caracterización del viñedo.

Qué duda cabe acerca de la importancia económica y cultural del vino, y su repercusión en la agricultura Argentina, Australiana, Chilena, Española, Estadounidense, Italiana y Francesa entre otras (Arnó *et al.*, 2009). La viticultura es sin duda un adoptador temprano en la incorporación de nuevas prácticas y tecnologías, debido que es una industria altamente competitiva, en la que quien no innova y optimiza sus procesos pierde frente a sus competidores.

Por éste motivo prácticas como la agricultura de precisión y la incorporación de tecnologías como la teledetección (Soliman *et al.*, 2013; Hall & Wilson, 2013), redes de sensores inalámbricos (Burrell *et al.*, 2004; Matese *et al.*, 2013) o vehículos aéreos no tripulados (Mathews & Jensen, 2013) han sido empleadas de modo creciente en la última década.

En particular tecnologías que permiten caracterizar la cubierta vegetal son de gran importancia para los viticultores ya que por una parte, a través de las características geométricas y estructurales (altura, ancho, volumen, densidad de hojas, área foliar, etc.) de las plantas, es posible optimizar muchas labores agrícolas, como la aplicación de pesticidas, la determinación de la demanda hídrica (riego), la fertilización, las podas y el guiado del cultivo, entre otros (Gil *et al.*, 2014).

Por otra parte caracterizar la cubierta vegetal a través de sus características de color permite estimar la producción, a la vez que determinar el estado del viñedo, (proporción de hojas nuevas marchitas, etc.) (Diago *et al.*, 2013; Dunn & Martin, 2004, Liu *et al.*, 2013).

Otra aplicación de la caracterización de la cubierta vegetal la podemos encontrar en la determinación del área de hojas aportantes a la fotosíntesis. En efecto uno de los parámetros más ampliamente usado en los modelos dedicados a la determinación de la demanda hídrica de las plantas, es el índice de área foliar (IAF) (Williams & Ayars, 2005), el cual es usualmente estimado mediante métodos destructivos, basados en extraer las

hojas del viñedo para luego medir su área de modo local. Otra forma de medir índice de área foliar, orientado a aumentar la cobertura espacial, es a partir del procesamiento de imágenes aerotransportadas o satelitales (Johnson, 2003), pero dicha forma de medida es más bien una estimación sujeta a calibraciones en terreno, lo que la hace poco fiable.

Otra de las aplicaciones la encontramos en uno de los principales objetivos históricos de la industria del vino: la estimación precisa y objetiva del rendimiento a cosecha (Dunn & Martin, 2004; Clingeleffer & Krstic, 2001; Wolper & Vilas, 1992) y de la calidad potencial de las uvas en la viña. Más específicamente, la predicción de rendimientos ha sido identificada en los últimos años como uno de los temas más rentables para investigación científica en la viticultura (Dunstone, 2002), ya que podría dar lugar a viñedos administrados de manera más eficiente, produciendo así, vinos de mejor calidad (Dunn & Martin, 2003). Hasta ahora, la mayoría de los métodos empleados para la estimación del rendimiento son destructivos, exigentes en mano de obra y tiempo (Martin et al., 2003), o muy caros de implementar (Blom & Tarará, 2009).

Si bien es cierto que el problema de la caracterización de la producción puede ser abordado con herramientas convencionales, como el corte y posterior pesaje de los racimos en el caso de que se busque estimar rendimiento, no es menos cierto que son técnicas destructivas, y costosas desde el punto de vista del tiempo y recursos humanos empleados. Estas técnicas por otra parte sólo ofrecen proyecciones con limitados resultados, ya que provienen de muestras puntuales y no de una medición de todo el viñedo. En efecto, autores como Clingeleffer *et al.* (2001) y Nuske *et al.* (2011a) señalan que las prácticas actuales para predecir el rendimiento son inexactas debido a que los métodos de muestreo tienden a ajustarse en función de rendimientos históricos e incluyen en sus cálculos parámetros subjetivos.

Cuando se estima rendimiento en base al peso del racimo, el cálculo del peso final de éste, se realiza a partir de mediciones de peso en envero, y se aplican multiplicadores fijos a partir de mediciones históricas, (Wolpert & Vilas, 1992). El problema de este método es que los multiplicadores están sesgados hacia las vides saludables discriminando así en contra de vides débiles o faltantes y por otra parte, estos multiplicadores tienen amplias variaciones según el viñedo, la temporada, la topografía,

el manejo y variedad, entre otros factores. Adicionalmente estos métodos estimativos no se adaptan bien a los nuevos modelos de manejo del vigor vegetativo del viñedo (podas en verde, deshoje, incluso aligeramiento de racimos) en los que prima la calidad de la uva frente al peso final.

Otra técnica frecuentemente empleada es la de las “hojas de puntuación para viñedo” para la evaluación de la calidad potencial de las uvas (Smart y Robinson, 1991; Bertamani *et al.*, 1994; Gray *et al.*, 1994; Carbonneau, 1995; Tardáguila & Martines de Toda, 2008), técnica que requiere una evaluación visual de diversas variables de la parte aérea de la vid, tales como vigor, estado de las hojas, área foliar expuesta, porosidad de la canopia¹ y fruta expuesta, todos ellos intrínsecamente relacionados con la calidad y composición final de la uva y del vino y (Kliewer & Lider, 1968 ; Kliewer, 1970; Crippen & Morrison, 1986a, Crippen & Morrison, 1986b;. Reynolds *et al.*, 1986; Mabrouk & Sinoquet, 1998; Smart, 1985).

En consecuencia, existe una necesidad de evaluar el rendimiento del viñedo y la calidad potencial de la uva a través de un monitoreo objetivo de las características de la parte aérea del viñedo, pero los métodos usuales para obtener las medidas de la canopia, como el “Point Quadrant” (Smart y Robinson, 1991) o LIDAR (Garrido *et al.* 2012; Fleck *et al.*, 2004, Moorthy *et al.*, 2007), aunque cuantitativos, están limitados en su precisión

¹ NOTA ETIMOLÓGICA: el término “**canopia**” puede ser definido como el conjunto de las partes aéreas de una planta (normalmente perenne) que engloba tanto a órganos vegetativos (tronco, ramas, hojas) como a órganos fructíferos (flores, frutos). Procede del latín medieval *canoepum*, que se documenta en la *General Estoria* de Alfonso X (c1280): «& tomo aquella red que diximos que dizien canopeo que estaua colgada de los tendales. & emboluiol el cuerpo como tronco sin la cabeça». En Español moderno *canopia* todavía no está incluido en el diccionario, pero la RAE informa (consulta específica del director de tesis) de su uso frecuente y ampliamente documentado en literatura de América latina y en la base de datos IATE (“Inter-Active Terminology for Europe, iate.europa.eu). Sus principales acepciones son: ‘dosel, baldaquino, palio’ (como en una cama) y ‘cubierta’ (como la de las copas de los árboles en el bosque, o la de una máquina o aeronave) pero su uso se extiende a muchas disciplinas técnicas. En el campo semántico que nos ocupa en esta tesis (agricultura, viticultura) las traducciones más exactas podrían ser ‘cubierta vegetal’, ‘follaje’, ‘copa’ o ‘parte aérea del cultivo’. Sin embargo se ha decidido simultanear estos términos con el neologismo ‘canopia’ (del inglés ‘canopy’) dado que sólo este último incluye en su significado los órganos fructíferos, de vital importancia en la vid y en la presente tesis. Adicionalmente, en los textos consultados el uso de ‘canopia’ lleva asociada la idea de “mayor o menor densidad vegetal en su interior”, cosa que no ocurre con otras palabras como ‘copa’ o ‘cubierta’ que suelen hacer referencia solamente a las propiedades superficiales.

y practicidad, ya sea porque demandan mucho tiempo o son muy costosos. Por lo tanto, se requieren nuevos métodos para evaluar el estado de la canopia, y la captura y análisis de imágenes parece ser una técnica objetiva y potencialmente útil para reemplazar los procedimientos que requieren mucho tiempo y para proporcionar información útil para una gestión más eficiente de la canopia.

Debido a los problemas antes descritos, aproximaciones más simples, basadas en la captura y procesamiento de imágenes, orientadas a la evaluación de las características de la canopia de la vid, se describen en las obras de Dunn y Martin (2004) y de Tardáguila *et al.* (2010, 2011). En dichos trabajos técnicas de análisis de imágenes digitales, fueron aplicados para la captura de datos en un estudio de los efectos de la defoliación, revelaron descripciones cuantitativas de distribución biomasa de la canopia, fruta expuesta, compacidad de los racimos y la eficacia del tratamiento, aunque el procesamiento de las imágenes no estaba automatizado.

Por este motivo diversos autores (Rabatel *et al.*, 2007, Berenstein *et al.*, 2010, Braun *et al.*, 2010, Nuske *et al.* 2011a, Sáiz-Rubio & Rovira-Más, 2012, Wang *et al.*, 2012, Liu *et al.*, 2013, Nuske *et al.* 2014, Skrabanek & Majerík, 2016) han desarrollado metodologías automatizadas, basadas en procesamiento de imágenes, que permitan con el menor grado de intervención humana caracterizar la parte aérea del viñedo, sin embargo no es un problema resuelto.

2. Sensores y tecnologías de imagen.

La caracterización de la cubierta vegetal ha evolucionado junto con la tecnología; en décadas pasadas se empleaban técnicas destructivas basadas en quitar todas las hojas, para luego medir manualmente su área empleando papel milimetrado (Pandey & Hema, 2011). Otro enfoque se basaba en medir el área de una hoja y pesarla, para luego, a partir de su peso extrapolar el área de la vid completa. Posteriormente aparecieron en el mercado medidores de área analógicos o planímetros (Daughtry, 1990), que fueron seguidos por los digitales actualmente en uso. Estos últimos a pesar de ser rápidos,

comparados con sus predecesores, aun requieren quitar cada hoja de la planta para lograr la determinación del área completa.

Estos métodos destructivos son costosos por el intensivo uso de mano de obra, toman largo tiempo y requieren personal calificado, sin mencionar el evidente efecto destructivo que tienen para la viña. Debido a la naturaleza destructiva de estos métodos resulta inviable realizar mediciones extensivas por lo que carecen de la representatividad estadística necesaria para convertirse realmente un reflejo del estado del viñedo (Nuske *et al.*, 2011a; Correa *et al.*, 2012b).

Si el problema es abordado empleando imágenes satelitales o aéreas, estas necesariamente deben ser calibradas con mediciones en terreno ya que la visión aérea sólo muestra una parte de la superficie foliar y no permite ver los racimos. Las calibraciones no son extrapolables a todos los viñedos, ya que dependerán de las prácticas culturales empleadas y de las características propias de cada cepa entre otras (Fairlie *et al.*, 2011)

Por otra parte, los métodos basados en imágenes satelitales o aéreas, ofrecen gran cobertura espacial, pero tiene como desventajas ofrecer baja resolución temporal y entregar sólo una medida indirecta del área real debido a la posición vertical con la que éstas son adquiridas. Sin embargo en los últimos años las imágenes capturadas desde vehículos aéreos no tripulados han provisto de una solución al problema de la baja resolución espacial y temporal. Esto, se debe a que las alturas de vuelo son del orden de las decenas de metros (Burgos *et al.*, 2015) y a que por sus menores costos de operación permiten una mayor resolución temporal.

Si se compara con las imágenes aerotransportadas, las técnicas basadas en imágenes tomadas *in situ*, ofrecen interesantes características, ya que brindan una vista directa de los frutos y de una gran parte de las hojas del viñedo, por lo que correlaciona de mejor modo con el IAF o con la producción. Además, sus costes tanto de adquisición como de operación son menores si se compara con los métodos destructivos antes descritos (Nuske *et al.*, 2011a).

Por estos motivos en los últimos años se han desarrollado métodos no destructivos, basados en tierra y próximos al cultivo, que aportan una mayor resolución espacio-temporal. Métodos que se han basados principalmente en LIDAR (siglas en inglés de Light Detection and Ranging) (Fleck *et al.*, 2004; Omasa *et al.*, 2007; Garrido *et al.*, 2012; Arnó *et al.*, 2013; Correa *et al.*, 2013b) y en procesamiento de imágenes en el espectro visible (Dunn & Martin, 2004; Rabatel *et al.*, 2007; Wang *et al.*, 2012; Reis *et al.*, 2012), imágenes que, en términos generales, son tomadas con cámaras digitales convencionales.

En el primer caso, con LIDAR, es posible, aunque con baja resolución, estimar el área y volumen del viñedo. Este método permite cubrir grandes extensiones en breves periodos de tiempo, pero por otra parte tiene un elevado costo de adquisición (Rosell *et al.*, 2009).

En el segundo caso, los métodos basados en el procesamiento de imágenes, proporcionan áreas de hojas visibles, es decir una proyección bidimensional del área de superficie real de las hojas fotosintéticamente activas, lo que asociado a calibraciones locales permite realizar proyecciones de la superficie de hojas. De este modo, es posible estimar el Índice de Área Foliar (IAF) (Breda, 2003) y determinar el estado de las hojas (jóvenes, senescentes, secas, etc.) presentes en la imagen (Meunkaewjinda *et al.*, 2008)

En efecto, el estado de vigor de las plantas puede ser determinado directamente a partir del color de las hojas, ya que ciertos tonos de verde se asocian a estados fotosintéticamente activos (Merzlyak *et al.*, 1999), mientras que coloraciones amarillas o marrones corresponden a deficiencias de nitrógeno o de agua, amén del estado de senescencia natural de las hojas (Bjurström & Svensson, 2002). Este proceso de determinar el estado de vigor, normalmente lo efectúan los viticultores a partir de extensos recorridos en el viñedo empleando medidores de clorofila portables (Vibhute & Bodhe, 2013) o su conocimiento experto. Pero como esta determinación se realiza en base a la información de color, el estado de vigor puede ser directamente inferido a partir del procesamiento de imágenes en el espectro visible.

A lo anterior se le debe sumar, la existencia de índices de vigor de las plantas, como el Índice de Vegetación de Diferencia Normalizada, el Índice de Vegetación Ajustado al

Suelo, NDVI y SAVI por sus siglas en inglés respectivamente, basados en la combinación de información proveniente del espectro visible así como del infrarrojo (Fernández *et al.*, 2013; Rodríguez-Pérez *et al.*, 2007). Estas imágenes generalmente son obtenidas desde satélites o vehículos aéreos no tripulados (UAV, siglas en inglés de unmanned aerial vehicles) (Proffitt, 2015), y en contadas ocasiones en terreno (Goutouly *et al.*, 2006; Storchi *et al.*, 2013).

Por otra parte, los métodos basados en procesamiento de imágenes permiten conocer la ubicación de las hojas y densidad de la masa foliar, información a partir de la cual es posible aplicar pesticidas de forma selectiva (Fairlie *et al.*, 2011), es decir, dirigir el pesticida sólo a las hojas evitando aplicar al tronco o a espacios sin hojas (Braun *et al.*, 2010). Además, a través de éstas imágenes es posible estimar y proyectar rendimientos con base la uva visible en la imagen (Dujnn & Martin, 2004; Diago *et al.*, 2013)

Estos sistemas de caracterización generalmente no operan en tiempo real y los que sí lo hacen están diseñados para identificar sólo una característica a la vez. Por ejemplo, Reis *et al.* (2012) detecta uvas, en imágenes tomadas durante la noche, usando un algoritmo que establece rangos de color para la uva, rangos que no son aplicables a otros viñedos ya que son producto de prueba y error para un viñedo en particular y en un estado de desarrollo particular. Para otro frutal, cerezos en particular y tendiente a detectar ramas individuales, se ha recurrido a la misma estrategia de capturar imágenes durante la noche y con iluminación artificial (Amatya *et al.*, 2016) para evitar los problemas ocasionados por la variabilidad de las condiciones de iluminación dentro de la escena.

Otros ejemplos de identificación de características, se pueden encontrar en Braun *et al.* (2010) y en Geng *et al.* (2012) en los que el algoritmo propuesto únicamente detecta hojas, ya que su uso está orientado a la aplicación selectiva de pesticidas. Este enfoque de detectar sólo una característica a la vez se debe principalmente a los elevados tiempos de cómputo de los algoritmos empleados para clasificación e identificación simultánea de características presentes en las imágenes (Correa *et al.*, 2011). En efecto no existen sistemas que permitan caracterizar la parte aérea del viñedo en tiempo real y usar dicha información para simultáneamente estimar rendimientos (kilos de uva al

momento de cosecha), calcular el Índice de Área Foliar (IAF) o aplicar pesticidas de forma selectiva (Braun *et al.*, 2010).

3. Metodologías de análisis y clasificación.

Uno de los enfoques más empleados para caracterizar los objetos presentes en una imágenes es la forma de estos, así por ejemplo muchos frutos tendrán formas circulares o elípticas, susceptibles de ser identificadas mediante técnicas tales como La transformada circular de Hough (Tan *et al.*, 2011, Roscher *et al.*, 2014), la Transformada de Simetría Radial (Nuske *et al.*, 2011) o los momentos de Zernike (Chamelat *et al.*, 2006).

Otro enfoque para extraer características se basa en las características de textura presentes en la imagen (Yan *et al.*, 2009). Aquí el supuesto es que objetos de similares características tendrán igual textura, sin embargo estas técnicas se enfrentan al problema de la distancia al foco de la cámara, es decir si un objeto está más distante en la escena, aunque su textura es la misma su escala habrá variado y por ende la premisa deja de ser verdadera. Ejemplo de esta técnica se pueden encontrar en Meunkaewjinda *et al.* (2008) y en Yan *et al.* (2009).

Dentro de las técnicas para la extracción de características, estas pueden ser clasificadas en las de umbralización y en las de separación de regiones. En las de umbralización el espacio de color es separado en planos y se intenta separar la región de interés, dentro del espacio, en base a valores umbral. Estas técnicas permiten separar zonas con formas rectangulares, lo que limita las características de color a extraer. Por otra parte, las técnicas de separación de regiones, intentan separar regiones del espacio empleando separadores lineales (Zheng *et al.*, 2010) o no lineales (Skrabanek & Majerík, 2016). En el primer caso el grado de complejidad de los algoritmos es menor si se compara con las técnicas de separación lo lineal.

Las técnicas de clasificación, independiente del espacio de color que se emplee, pueden ser divididas en supervisadas y no supervisadas (Kotsiantis, 2007). En métodos supervisados, el número de características a identificar o clases es determinado *a priori*

y el supervisor es quien selecciona el prototipo de estas clases. Por el contrario, en los métodos no supervisados, se desconocen las características de las clases, y el algoritmo de clasificación atribuye pertenencia de tal manera que los elementos en cada clase exhiban características similares y sean más similares entre sí, que con respecto a los elementos de otras clases.

Los métodos supervisados y no supervisados se han utilizado en exteriores con propósitos agrícolas (Tian & Slaughter, 1998) y específicamente en viñedo, para la extracción de características de la cubierta orientadas a caracterizar su vigor (Bjurström & Svensson, 2002); detectar racimos y hojas (Berenstein *et al.*, 2010); uvas individuales (Chamelat *et al.*, 2006); contar 'píxeles' de frutas para la estimación del rendimiento (Dunn & Martin, 2004), o separar las uvas, hojas y sarmientos (Correa *et al.*, 2011a, Correa *et al.*, 2011b).

En entornos no estructurados, tales como un campo agrícola, las condiciones son variables, por lo que la robustez de los algoritmos no supervisados puede estar en riesgo (Kotsiantis, 2007). Por lo tanto las técnicas de clasificación supervisadas son de especial interés en esta área, ya que el grupo de entrenamiento puede ser preparado a priori estableciendo qué características corresponden a los elementos de una clase (Gonzales, 2010), que, a su vez, reduce la incertidumbre y conduce a la soluciones posibles.

Pero las técnicas de agrupamiento supervisadas tienen escasa capacidad de generalización y por ende requieren un entrenamiento previo para cada conjunto de imágenes a tratar, por ello los esfuerzos han sido dirigidos a adoptar técnicas no supervisadas que se adapten a las condiciones locales de vegetación e iluminación (Correa *et al.*, 2012, Liu *et al.*, 2013).

En definitiva la extracción de características del viñedo empleando técnicas de procesamiento de imágenes ha probado su viabilidad, pero está lejos de haber resuelto el problema debido principalmente a la variabilidad de las condiciones de terreno, así como por los elevados tiempos de cómputo requerido por la mayoría de la técnicas empleadas.

De ahí surge la necesidad por desarrollar una metodología que permita determinar: el área de hojas y sus características, el área de racimos para estimar rendimiento a cosecha y claro está, hacerlo en tiempo real, para de este modo proveer de una herramienta robusta y fiable, que le permita a los viticultores tomar decisiones oportunas e informadas.

En sucesivos epígrafes de esta tesis (Metodología) se discutirán en profundidad un gran número de técnicas de análisis de imagen aplicadas por otros autores a imágenes de viñedo, comparando su efectividad o proponiendo modificaciones novedosas. Se ha creído conveniente mantener las descripciones de dichas técnicas matemáticas en la Metodología pues, aunque provienen de la literatura, su aplicación, modificación y discusión son parte fundamental de esta tesis.

IV. HIPÓTESIS

- No existen sistemas que permitan caracterizar en tiempo real la parte aérea del viñedo.
- Es posible caracterizar la parte aérea del cultivo mediante la captura de imágenes Visibles, Infrarrojas y de Profundidad.

V. OBJETIVOS.

1. Objetivo Generales

Desarrollar una nueva metodología de análisis de imagen para de caracterización de la cubierta vegetal en viñedos en tiempo real, empleando un sistema de adquisición de imágenes visibles (RGB), de profundidad (D) e infrarrojas IR.

2. Objetivos Específicos

- Obtener información del cultivo mediante el análisis de las imágenes tomadas para:
 - Estimar el área foliar.
 - Estimar el rendimiento empleando el área de uva visible.
 - Caracterizar las hojas entre distintos grados de vigor.

VI. METODOLOGIA.

Síntesis metodológica y relación con trabajos publicados.

Debido al desfase temporal de las campañas de adquisición de imágenes y las disímiles condiciones de terreno en la que las imágenes fueron adquiridas, se exploró el uso de diversas técnicas tendientes a resolver problemas producidos por diferencias en iluminación y producto de la variabilidad propia del viñedo. Dichos resultados parciales arrojaron luz acerca del camino a seguir, lo que permitió, en cada campaña de adquisición, seleccionar las técnicas más adecuadas de entre todas las evaluadas. Las técnicas con resultados promisorios fueron publicadas en revistas científicas y en congresos de la especialidad, generando una serie de artículos que se resumen a continuación.

En el artículo “Extracción de Características de la Cubierta Vegetal del Viñedo Mediante Imágenes RGB y RGIR Obtenidas de Forma Dinámica” se muestra la selección del espacio de color y canales adecuados para la extracción de características del viñedo mediante algoritmos de agrupamiento convencionales y difusos.

Se investiga además la detección de racimos empleando descriptores de forma, mediante la transformada de Hough, además del uso de imágenes infrarrojas e índices de vegetación con el mismo propósito.

Una vez definido el espacio de color y canales adecuados para la caracterización del viñedo se evaluaron diversas técnicas de agrupamientos difusas, ya que estas mostraron mejores resultados que las convencionales. Se seleccionaron algunos de estos algoritmos y su comparativa puede ser observada en el artículo “A Comparison of Fuzzy Clustering Algorithms Applied to Feature Extraction on Vineyard”.

Una vez definido el algoritmo de agrupamiento difuso de mejor desempeño en clasificación, se diseñó una estrategia de aceleración de convergencia, ya que estos algoritmos resultaron no ser adecuados para ser ejecutados en tiempo real. Dicha estrategia de aceleración de convergencia puede observarse en “Feature Extraction on Vineyard by Gustafson Kessel FCM and K-Means.”. En el mismo artículo se implementó una red neuronal con el propósito de identificar los grupos de píxeles generados.

Ya establecida la estrategia de aceleración de convergencia, en el artículo “Real-Time Characterization of Vineyard's Canopy Through Fuzzy Clustering and SVM Over Color Images”, se evaluó otra técnica de identificación denominada *Support Vector Machine* SVM, esta técnica requiere menos tiempo de ejecución si se compara con las redes neuronales y además requiere grupo de entrenamiento de menor tamaño, lo que la hace más adecuada para operar en tiempo real. Por otra parte se evaluó la generación de un mosaico con las imágenes de cada hilera, a través de la técnica denominada *Scale-Invariant Feature Transform* SIFT, con el fin de evitar errores producidos por el solapamiento de imágenes.

Ya definidos los espacios de color y algoritmos de clasificación no supervisados, se evaluó el uso de algoritmos de clasificación supervisados con el propósito específico de determinar la viabilidad de proyectar rendimiento a cosecha. Los resultados de esta investigación pueden ser observados en “Grapevine's Yield and Leaf Area Estimation Using Supervised Classification Methodology on RGB Images Taken Under Field Conditions”.

Debido a que los algoritmos de agrupamiento e identificación empleados pueden operar en tiempo real, se evaluó la posibilidad de emplearlos en la aplicación selectiva o localizada de pesticidas. Dicha evaluación puede ser observada en el artículo “Usando Kinect como Sensor para Pulverización Inteligente”.

Además en el artículo “Comparación de Desempeño de una Cámara de Profundidad Comercial y un LIDAR en Explotaciones Agrícolas” se realizó una comparativa entre una cámara RGB-D y un LIDAR convencional con el propósito de evaluar las ventajas de uno y otro dispositivo en la aplicación selectiva de pesticidas.

La metodología presentada a continuación será separada en tres secciones. En la primera se describirán los procedimientos de adquisición de imágenes utilizados, en la segunda se describirán las técnicas de procesamiento empleada y en la tercera se describirá la metodología de predicción de rendimiento y de área foliar.

Pero debido a que la metodología fue evolucionando en cada campaña de adquisición de imágenes, cada sección será separada a su vez en campañas de adquisición de imágenes.

Trabajos publicados:

- Extracción de Características de la Cubierta Vegetal del Viñedo Mediante Imágenes RGB y RGIR Obtenidas de Forma Dinámica. Congreso: VI Congreso Ibérico de AgroIngeniería (2011)
- A Comparison of Fuzzy Clustering Algorithms Applied to Feature Extraction on Vineyard. Iberoamerican Journal of Artificial Intelligence. (2011).
- Feature Extraction on Vineyard by Gustafson Kessel FCM and K-Means. Congreso: 16th IEEE Mediterranean Electrotechnical Conference. Tunicia. (2012)
- Real-Time Characterization of Vineyard's Canopy through Fuzzy Clustering and SVM over Color Images. Congreso 4th International Workshop on Computer Image Analysis in Agriculture, held at CIGR-AgEng. Valencia, España. Julio de 2012
- Grapevine's Yield and Leaf Area Estimation Using Supervised Classification Methodology on RGB Images Taken under Field Conditions. Journal: Sensors 12 (2012)
- Usando Kinect como Sensor para Pulverización Inteligente. Congreso Ibérico de Agroingeniería, Madrid, España. Agosto de 2013.
- Comparación de Desempeño de una Cámara de Profundidad Comercial y un LIDAR en Explotaciones Agrícolas. Congreso Ibérico de Agroingeniería. Madrid, España. Agosto de 2013.

1. PROCEDIMIENTOS DE ADQUISICIÓN DE IMÁGENES

1.1. Primera campaña de adquisición

1.1.1. Adquiridas de modo estático.

Como primera aproximación y con el propósito de generar condiciones favorables para el procesamiento, en una primera temporada se capturaron imágenes de modo estático con una cámara RGB convencional (Pentax modelo K200D, sensor CCD de 23,5 x 15,7 mm, Japón) montada en un trípode fijado normal a la cubierta a 2 m del eje de las hileras y 1,05 m sobre el suelo.

Se dispuso una pantalla blanca detrás de la vid para evitar los efectos de confundir la vid de interés con la vegetación de fondo, y no se empleó iluminación artificial. Se capturó un conjunto de 400 imágenes las que fueron capturadas con una resolución de 3504 x 2336 píxeles y luego reducidas a 800 x 600 píxeles con el fin de acelerar el tiempo de procesamiento.

Un segundo conjunto de 400 imágenes fueron adquiridas siguiendo el mismo procedimiento experimental antes descrito, pero esta vez en el campo prácticas de la ETSIA. Dichas imágenes fueron reducidas de tamaño sistemáticamente (800x600, 640x480, 320x240 píxeles) para determinar la mínima resolución a la que es posible extraer e identificar características de las vides.

1.2. Segunda campaña de adquisición

1.2.1. Adquiridas de modo dinámico.

Debido a que las imágenes visibles podrían no expresar todas las características del viñedo, en la segunda campaña de adquisición se incluyó una cámara RGIR, además de una cámara de video estándar.

Debido al alto volumen de imágenes y distancias recorridas para capturar un número de imágenes representativo del viñedo, además del peso que significa el sistema de captura en su conjunto, se decidió instalar las cámaras en un carro remolcado.

De este modo, para la adquisición de imágenes, se construyó un vehículo remolcable con una estructura modular adecuada para la implementación de cámaras y sensores en campo. El remolque fue construido con unas dimensiones y peso adecuados para ser arrastrado mediante 'quads' o tractores viñeros sin pérdida de capacidad de maniobra (**Figura 1**).

La estructura superior emplea perfiles de aluminio ranurados que permiten la fijación rápida de soportes en todas sus caras mediante pernos de fijación. Esto permite la rápida adaptación en campo a las condiciones del cultivo y de los sensores a emplear (altura, distancia al cultivo).

Figura 1. Remolque empleado para la adquisición de las imágenes, dispone de una cama de video, una cámara IR, un DGPS y un computador portátil para el almacenamiento de la información.

En la parte superior del remolque se dispuso de una plataforma con elementos de sujeción para portar los PCs empleados para el ajuste, control y almacenamiento de datos procedentes de los distintos sensores. El remolque dispuso también de un alojamiento para una batería de 12 VCC la cual fue conectada a un inversor de 12 VCC/220VCA con protección por diodos, para proveer la alimentación eléctrica a los PCs.

Tabla 1. Velocidades reales (m s^{-1}) para cada marcha seleccionada en el tractor y tiempos de integración establecidos para ambas cámaras (- combinación no usada).

Marcha seleccionada	Tiempo de integración (ms)					
	50	30	20	10	5	3
4HT	-	0,91	-	0,84	0,81	0,83
2HT	0,42	0,43	0,42	0,41	0,41	0,40
4HC	0,19	0,20	0,19	0,18	-	-
Paradas	0,00	0,00	-	0,00	-	-

Luego de optimizar algunos parámetros como la frecuencia de grabación de imágenes, campo de visión en función de la óptica disponible y las distancias al cultivo, se establecieron los parámetros y variables a ensayar en las pruebas de campo sobre un viñedo (V. vinífera L. cv. Tempranillo) situado en Ayegui (lat: 42° 39'N; long 2° 03'W, 540 m) La Rioja (España) durante el mes de octubre de 2010, una semana antes de la vendimia.

La Tabla 1 muestra las velocidades reales (m s^{-1}) para los ensayos realizados a distintas marchas del tractor empleado para arrastrar el vehículo remolcable y para los distintos tiempos de integración seleccionados para las cámaras RGB y RGIR. Para la realización de los ensayos se implementaron en el vehículo remolcable los siguientes equipos de medida: cámara RGB: DuncanTech MS3100-RGB, sensor CCD de 7,6 x 6,2 mm (VIS), cámara RGIR: DuncanTech MS3100-CIR, sensor CCD de 7,6 x 6,2 mm (VIS-IR), cámara de vídeo: Sony Handycam digital, GPS diferencial: Garmin 17xHVS y GPS RTK: Leica Zeno 10.

1.2.2. Tercera campaña de adquisición

1.2.3. Adquiridas de modo dinámico.

De los ensayos realizados con la plataforma surgieron una serie de necesidades que conducen a la conclusión de que se requería una plataforma más estable, que mantuviese las cámaras paralelas a la hilera, que redujese el efecto de la vibración sobre las cámaras y que eliminase el efecto de la incidencia directa de los rayos solares sobre las cámaras, por lo que se diseñó y construyó la plataforma y el sistema de estabilización de la **Figura 2**.

En las pruebas de terreno, llevadas a cabo empleando la plataforma remolcada descrita en Correa *et al.* (2011a), la plataforma mostró que a pesar de sus reducidas dimensiones (1,5x 1,2 m) estaba afectada fuertemente por el espacio disponible entre hileras y las vibraciones. Además mostró baja maniobrabilidad en las cabeceras de las hileras.

Figura 2. Esquema de la plataforma de adquisición empleada en los ensayos de terreno. Dispone, además de las cubiertas para evitar los efectos de la luz, de un sistema de estabilización mecánico y electrónico diseñado para mantener la cámara de forma paralela a la hilera.

La baja maniobrabilidad se debió en parte a que se tuvo que añadir una cubierta superior y una trasera, adicional a la dispuesta inicialmente, para evitar las fuertes variaciones de iluminación que se producen a lo largo del día en los viñedos.

Estas cubiertas añadidas aumentan significativamente la envergadura de la plataforma, llevándola hasta dimensiones de 3 m de largo por 2,5 de ancho por 2,2 m de alto, lo que explica la dificultad en la maniobrabilidad, amén de aumentar el movimiento y las vibraciones en la plataforma en su conjunto.

Debido a ello se diseñó e implementó un sistema de estabilización mecánico y otro electrónico. El sistema mecánico se basa en un brazo basculante que cuenta con un contrapeso para mantener vertical la cámara (Figuras 2, 3 y 4).

Para los movimientos en los que la velocidad del brazo basculante es insuficiente se diseñó un sistema electrónico basado en una unidad de medición inercial (IMU del inglés *Inertial Measurement Unit*) de 6 grados de libertad, modelo MiniIMU-9 de Pololu Corporation y un microcontrolador modelo 328 de Atmel. El microcontrolador se encarga de procesar datos provenientes de la IMU, empleando un filtro de Kalman extendido. Este filtro fusiona datos provenientes de los giróscopos, acelerómetros y magnetómetro, para generar una orientación estable y libre de ruido. El microcontrolador se encarga además de enviar las señales de posicionamiento a los servomotores que accionan el sistema de estabilización (**Figura 3**).

Figura 3. Sistema de posicionamiento y estabilización de la cámara, accionado por dos servo motores y controlados desde un microcontrolador.

Por otra parte se determinó, a partir de la experiencia de campo obtenida en la segunda campaña de adquisición, que el largo debía ser de al menos 3 m para conseguir eliminar completamente la incidencia directa de los rayos del sol, sin embargo su distancia entre ejes debería ser de 2 m, ya que esta configuración de ruedas equidistantes permite una mejor maniobrabilidad.

a)

b)

Figura 4. Plataforma empleada en los ensayos de terreno, en la que se aprecia parcialmente el sistema de cobertura para evitar los efectos de las variaciones en iluminación. a) Vista frontal b) vista en perspectiva.

2. TÉCNICAS DE PROCESAMIENTO EMPLEADAS.

Se evaluaron diversos algoritmos diseñados para la identificación y extracción de características de objetos de interés en la escena, en éste caso uva y hojas. La identificación se basa en agrupar píxeles que posean características similares de color, textura o que describan una forma circular en el caso de la uva. Una vez creados los grupos estos son etiquetados, en primera instancia de forma manual y posteriormente de modo automático, para más tarde extraer características como área, color, ubicación.

Aunque las técnicas empleadas varían según la campaña de adquisición de imágenes estas pueden ser agrupadas del siguiente modo: A. Algoritmos de Clasificación Supervisados, B. Algoritmos de Clasificación No Supervisados y C. Métodos de identificación o etiquetado de grupos.

A. Algoritmos de Clasificación Supervisados.

Los algoritmos de clasificación supervisada operan sobre la información proporcionada por un conjunto de muestras, patrones o prototipos que se consideran representativos de las clases previamente definidas. Se asume que se conoce la clase a la que pertenece cada muestra y que la etiqueta de clase es correcta. Se evaluará el método de agrupamiento basado en la distancia de Mahalanobis y el Método de los Vecinos más Cercanos.

B. Algoritmos de Clasificación No Supervisados.

Los algoritmos no supervisados, a diferencia de los supervisados, no emplean un conjunto de datos para entrenamiento y, mediante algoritmos de agrupamiento, generan su propio conjunto de entrenamiento (González, 2010). Los algoritmos no supervisados basados en color que se presentarán serán: K-means, K-means++, K-medoids, Fuzzy C-Means, Possibilistic C-Means, Fuzzy Possibilistic C-Means, Robust Fuzzy Possibilistic C-Means, Fuzzy C-Means con el algoritmo de Gustafson Kessel, Kernelld Robust Fuzzy Possibilistic C-Means, Robust Fuzzy Possibilistic C-Means con restricciones espaciales y triangulación de Delaunay.

Por otra parte se evaluará métodos no supervisados basados en la textura tales como Wavelets y segmentación de texturas (Brox & Weickert, 2004) y por forma como la Transformada de Hough y la Transformada Rápida de Simetría Radial.

C. Métodos de identificación o etiquetado de grupos.

Debido a que una vez generados los grupos es necesario etiquetarlos, es decir, identificar a que clase pertenecen, se requiere un método de caracterización de los grupos de píxeles, si bien es cierto que es posible caracterizarlos desde el punto de vista estadístico, mediante técnicas como análisis discriminante o a través de distancias como la de Mahalanobis. Estos métodos estadísticos no poseen la capacidad de generalización que ofrecen otras técnicas del área de la inteligencia artificial como por ejemplo las redes neuronales.

En esta etapa se evaluarán dos técnicas, la primera basada en redes neuronales, la *Feed-Forward Back Propagation* y la otra basada en la técnica de identificación denominada Máquinas de Soporte Vectorial (SVM, *support vector machines*).

2.1. Primera campaña

2.1.1. Procedimiento aplicado.

En esta primera campaña se recurrió a las técnicas básicas de segmentación de imágenes como el método de Otsu, método que no generó resultados viables ya que se basa en la separación lineal del histograma de la imagen a partir de un sólo canal de color o de la combinación lineal de estos. Por tal motivo se enfocó el problema desde las perspectivas de clasificación por color, clasificación por forma y clasificación por textura. Las premisas subyacentes son: que los objetos presentes en una escena presentan características homogéneas en su interior, en este caso, textura y/o color y por otra parte que objetos de similar forma pertenecen a una misma clase, en este caso, las figuras de forma circular pertenecen a la clase uvas, ya que las bayas presentan formas circulares. Los objetos a identificar dentro de la escena son: Hojas, Tronco Sarmientos, Racimos (bayas) y Fondo.

Así en este apartado se describirá que espacios de color y canales son más adecuados para los métodos evaluados. Se describirá además métodos de agrupamiento por color (supervisado y no supervisado), métodos de segmentación de texturas y métodos de identificación por forma y su procedimiento de validación.

En el caso de la identificación por forma se empleó la transformada circular de Hough, como la describe Fernández *et al.*, (2008) y en el caso de la clasificación por color se empleó las técnicas de clasificación no supervisada denominadas *K-Means* y *Fuzzy C-Means* descritas por Ray *et al.*, (1999) y Chuang *et al.*, (2006) respectivamente. Estas técnicas básicamente agrupan píxeles de similares características (mismo color), en base a un número preestablecido de grupos o clases.

Estas técnicas fueron aplicadas sobre un conjunto de 400 imágenes RGB, capturadas a una resolución de 3504 x 2336 píxeles. Luego estas imágenes fueron procesadas, mediante rutinas desarrolladas en Matlab 7.5, en un ordenador con procesador Pentium 4 ® de 1,6 MHz y 1Gb de RAM.

2.1.2. Método de Validación.

Con el propósito de evaluar el desempeño de los algoritmos de segmentación, de un conjunto de 400 imágenes de vides, un subconjunto representativo de 20 imágenes (con una resolución de 640x480 píxeles en primera instancia y 320x240 píxeles en segunda instancia) fueron manualmente segmentadas. De éste modo se contabilizó el número de píxeles de cada clase.

2.1.3. Selección del espacio de color.

Como primera etapa en el procesamiento de las imágenes RGB en el espectro visible, éstas fueron transformadas a los espacios de color HSV, CMYK, $L^*a^*b^*$, XYZ y Ohta, (Ohta *et al.*, 1980) para determinar así cuál de ellos o su combinación es el más adecuado para el proceso de identificación de distintos tipos hojas, racimos, sarmientos y fondo. Todo ello bajo las premisas de que píxeles de un mismo color pertenecen al mismo tipo de objetos y de que algunos espacios de color permiten diferenciar mejor los objetos que otros.

El espacio de color RGB no es una buena opción para la segmentación de color, debido a la alta correlación entre los canales R, G y B y también debido a que las mediciones de distancia no representan las diferencias de color en el espacio RGB. A diferencia del espacio RGB, el espacio de color $L^*a^*b^*$ es capaz de proporcionar la información de intensidad y de color de forma independiente. Además expresa las diferencias de color, percibidas por los seres humanos, a través de la distancia euclidiana. En este espacio de color, L^* canal representa la intensidad (luminosidad) y a^* b^* son los canales cromáticos, por lo que es posible eliminar el canal de L^* y dar cuenta de las diferencias de color simplemente usando los canales a^* y b^* (Cheng *et al.*, 2001). La eliminación de canal L^* reduce el espacio de color a un plano, lo que permite realizar la segmentación (agrupamiento) de un modo más rápido.

2.1.4. Métodos de Agrupamiento no supervisados

Aquí se describen dos aproximaciones seguidas para agrupar píxeles de similares características. Dichas aproximaciones son las difusas (Fuzzy c-Partition) y las no difusas o clásicas (Hard c-Partition), tratando en ambos casos de agrupar objetos hasta en “c” clases o grupos.

2.1.4.1. Algoritmo de agrupamiento difuso

En los métodos de agrupamiento difuso (FCM, Fuzzy c-Means) objetos (píxeles) pueden pertenecer al mismo tiempo a varios grupos con distintos grados de pertenencia, mientras que los métodos de agrupación determinista asignan cada objeto a un grupo único. Esta es una característica importante porque algunos píxeles pueden pertenecer en diferentes grados, por ejemplo, al grupo racimos de uva y al mismo tiempo al grupo tronco. Un observador humano puede distinguir entre ellos debido al contexto, aunque el color sea en apariencia el mismo. El agrupamiento difuso resuelve el problema de dar una respuesta a la pregunta: ¿a qué clase pertenece más un píxel?

El número de grupos debe ser cuidadosamente seleccionado, siendo esencial tener tantos grupos como características en la imagen. Esto permite crear un contexto para el proceso de agrupamiento.

Se han propuesto muchos algoritmos de agrupamiento difuso, sin embargo el algoritmo Gustafson-Kessel con la estimación de covarianza mejorada (Babuska *et al.*, 2002) ha demostrado ser suficientemente robusto y tener altos niveles de desempeño (Meyer *et al.*, 2004, Correa *et al.*, 2011). Las principales características del algoritmo de Gustafson-Kessel (GK) son su tolerancia al ruido y la adaptación local, de la métrica de distancia, a la forma del grupo, mediante la estimación de la matriz de covarianza del grupo y la adaptación correspondiente de la matriz de distancia inducida por norma (Graves y Pedrycz, 2007).

El algoritmo FCM-GK se basa en la optimización iterativa de un funcional² objetivo del tipo c-means (Ecuación 1):

$$J_m(U, V) = \sum_{i=1}^c \sum_{k=1}^n u_{ik}^m d_{ik}^2$$

Ecuación 1. Función objetivo FCM en términos de la distancia.

Donde c es el número de grupos (i es el grupo actual) ; $U=[u_{ik}] \in [0,1]^{c \times n}$ es una matriz de partición difusa de los n datos $X=[x_k]$ donde u_{ik} es el grado de pertenencia de x_k al i -ésimo grupo.

La distancia inducida por norma (Ecuación 2) d_{ikA_i} puede dar cuenta de grupos de diferentes formas en un conjunto de datos:

$$d_{ikA_i}^2 = (x_k - v_i)^T (x_k - v_i)$$

Ecuación 2. Distancia inducida por norma en FCM.

Siendo $V=[v_i] \in \mathbb{R}^n$ los centros de grupo (centroides) y $m=[1, \infty]$ es el exponente ponderado, que determina cómo de difuso será el la funcional.

La métrica de cada grupo d_{ikA_i} se define por una matriz de norma inducida local A_i , que se utiliza como una de las variables de optimización del funcional. Esto permite que la distancia inducida por norma se adapte a la estructura topológica local de los datos

El inconveniente principal de este algoritmo es que necesita tiempos de ejecución largos en comparación con K-means.

² Función cuyo dominio es un conjunto de funciones.

2.1.4.2. Agrupamiento mediante el algoritmo K-means:

K-means es el algoritmo de agrupamiento más popular utilizado en minería de datos, bioinformática, procesamiento de imágenes y reconocimiento de patrones (Berkhin, 2002; Amorim y Mirkin, 2012). Esta técnica de agrupación divide un determinado conjunto de datos a través de un cierto número de grupos 'K' conocido *a priori*. Busca minimizar la distancia entre puntos pertenecientes al mismo grupo. K-means es significativamente sensible a los valores iniciales de los centros de grupos, generalmente seleccionados al azar, no tiene defensa contra el ruido o características irrelevantes y además no posee cohesión espacial.

Así, aunque no ofrece ninguna garantía de exactitud, su sencillez y rapidez son muy interesantes en aplicaciones del mundo real. El funcional a minimizar, se puede expresar como en la Ecuación 3.

$$J_m(V) = \sum_{i=1}^c \sum_{k=1}^n \|x_k - v_i\|^2$$

Ecuación 3. Función objetivo K-means.

Donde $\|x_k - v_i\|$ es distancia inducida por norma, también expresada como d_{ik} . En este caso fue elegida la distancia coseno porque tiene una particularidad: es directamente una métrica de similitud entre dos vectores. La salida de las distancias varía de 0 a 1.

A pesar de que K-means es un método de agrupación determinista, cuando se introduce la distancia coseno, la noción de similitud aparece entre el centro del grupo y los píxeles a ser clasificados. También cuando este algoritmo se inicializa con un buen punto de partida, es capaz de evitar los mínimos locales y su desempeño es más rápido si se compara con la inicialización aleatoria (Correa *et al.*, 2012). Un buen punto de partida puede ser definido como un punto cercano al centro del grupo generado como resultado de la ejecución del algoritmo. El marco metodológico propuesto en este trabajo se utiliza los centros de los grupos generados por FCM-GK como valores iniciales para K-means.

Esto permite obtener algunas de las ventajas de la FCM-GK (evitar los mínimos locales, la independencia de la selección inicial del grupo) y utilizarlos para mejorar el rendimiento K-means, pero manteniendo y mejorando la ventaja principal de K-means: su velocidad.

2.1.4.3. Segmentación basada en Delaunay.

Se aplicó la triangulación de Delaunay sobre el histograma de la imagen como propone Cheddad *et al.*, (2003). A diferencia de otros algoritmos como K-means, este método no necesita ningún parámetro para efectuar la segmentación; es más bien un proceso totalmente automatizado. En otras palabras es totalmente no supervisado. Todo lo que necesita para funcionar es un histograma de la imagen.

2.1.5. Métodos de Agrupamiento supervisados.

2.1.5.1. Agrupamiento mediante el algoritmo Vecinos más Cercanos:

El algoritmo de los k-Vecinos más cercanos (kNN por sus siglas en Inglés, de *k nearest neighbours*) es un método de clasificación supervisado no paramétrico, no hace supuestos acerca de la distribución de los datos, no requiere distribución gaussiana o que los datos sean separables linealmente. No requiere una fase de entrenamiento, sólo un conjunto de referencia, pero requiere mayores tiempos de cómputo en la etapa de prueba ($O(n*d)^3$) y es altamente sensible al ruido.

El algoritmo de los kNN estima el valor de la función de densidad de probabilidad de que un punto pertenezca a la clase a partir de un conjunto de prototipos definidos a priori.

³ Orden de complejidad del problema, donde n es el número de datos y d la dimensión del vector.

Dado un conjunto X de n puntos y una función de distancia, el algoritmo de los k -vecinos más cercanos (kNN) permite encontrar los k puntos más cercanos en X , a un punto o conjunto de puntos de referencia.

En esta tesis se evaluó el uso de la distancia Euclidiana y del coseno como medidas de distancia para kNN. Por otra parte se implementó una interfaz para el usuario, que le permita a partir de la imagen seleccionar los píxeles de referencia para las clases Uva1, Uva 2, Hoja secas, Hojas jóvenes, Hojas Viejas Fondo, Sarmientos y Tronco.

2.1.5.2. Agrupamiento mediante la distancia de Mahalanobis.

Es un método que permite determinar la similitud entre dos variables aleatorias multidimensionales y tiene como característica principal que le asigna el mismo peso a cada dimensión al ponderarlas según su varianza. Este método será descrito in extenso en la sección 3.4

2.1.6. Detección a través de la forma.

En fases posteriores del análisis se evaluó la identificación de las uvas basados en la forma casi circular de las bayas. Para ello se empleó la transformada circular de Hough como en Murillo *et al.* (2012), algoritmo diseñado para buscar círculos de un radio determinado o segmentos de estos dentro de una imagen.

Además se evaluó la Transformada rápida de simetría radial, propuesta por Loy & Zelinsky, (2003). Esta transformada busca simetría radial local para resaltar los puntos de interés dentro de una imagen.

2.1.6.1. Transformada circular de Hough.

La transformada de Hough comienza obteniendo un mapa de bordes sobre una imagen en escala de grises, en este caso, el mapa fue generado empleando las técnicas de Sobel, Prewitt, Roberts y Canny (Heath *et al.*, 1998), sobre los canales a^* y b^* del espacio de color $L^*a^*b^*$ y sobre el canal R del espacio RGB.

La información obtenida es utilizada para deducir la ubicación de los centros (x, y) y los valores de los radios (r) de los posibles círculos, por último se lleva a cabo la detección de máximos. La desventaja de este método es la alta demanda de recursos computacionales, ya que debe realizar múltiples operaciones en forma recursiva y almacenar los resultados parciales de dichas operaciones (Fernández *et al.*, 2008).

2.1.6.2. Transformada de Simetría Radial.

Se evaluó la técnica denominada Transformada de Simetría Radial (como lo proponen Nuske *et al.* (2011)). El algoritmo es robusto a los problemas de iluminación y de bajo contraste de color, lo que causa problemas a las técnicas de detección de cultivos existentes que se basan en la discriminación de color simple. El método detecta los centros de bayas de todos los colores, incluso los que son del mismo color de las hojas de fondo

La TSR requiere conocer de antemano los radios de las bayas que aparecerán en la imagen, pero los radios de bayas (en píxeles) dependen de la longitud focal de la cámara, el tamaño real de la baya y la distancia desde la cámara. Entonces para determinar los radios se debe mantener fija la longitud focal, así como la distancia de la cámara a las vides.

2.1.7. Detección a través de Análisis de textura.

La textura de una imagen es definida como una región macroscópica estructurada de la imagen que presenta propiedades locales constantes, lentamente variables o aproximadamente periódicas. A partir de la textura se intentan distinguir diferentes patrones de imagen mediante la extracción de la dependencia de la intensidad entre los píxeles y sus píxeles vecinos (Kartikeyan & Sarkar, 1991).

De este modo bajo la premisa de que zonas con igual textura representan objetos con características similares, se realizó un análisis de textura, sometiendo a las imágenes a

distintos filtros i.e. filtro de entropía local, filtro entropía global, filtro del rango y filtro de la media, sobre diversos espacios de color.

Además se evaluó el uso de la metodología de Wavelets sobre imágenes RGB y el método propuesto por Brox *et al.* (2004), sobre imágenes en escala de grises.

Las imágenes sometidas a los filtros de textura fueron convertidas previamente a los espacios de color, L^*a^*b , RGB, HSV, XYZ.

2.1.7.1. Wavelet de Haar.

La transformada *Wavelet* mapea la imagen en una sub-imagen de baja resolución y una serie de imágenes de detalles. La imagen baja resolución se obtiene enturbiando iterativamente la imagen, mientras que las imágenes de detalles contienen la información perdida durante dicha operación.

Esto permite que las imágenes de menor resolución generadas en cada iteración conserven la información más significativa y por ende más representativa de las características remanentes.

Para su evaluación se empleó el Wavelet Toolbox de Matlab, con las imágenes en su espacio de color original, HSV y La^*b^* , a resoluciones de 800x600, 640x480 y 320x240 píxeles.

2.1.7.2. Segmentación de texturas

En cuanto al método propuesto por Brox & Weickert, (2004) de “Segmentación de texturas sobre un espacio de características de difusión no supervisada”, presenta interesantes características ya que permite identificar racimos, hojas, tronco etc., pero requiere un proceso posterior de identificación y fusión de grupos de texturas generados.

Todas estas características en su conjunto permiten medir la disimilaridad entre las texturas y luego agrupar las texturas iguales, que posteriormente podemos asociar a la clase de interés.

2.2. Segunda Campaña de adquisición

2.2.1. Procedimiento Aplicado.

En esta campaña se expandió los espectros de color capturados al añadir imágenes infrarrojas. Además se añadió imágenes capturadas en movimiento para aumentar la cantidad y por ende la representatividad de los datos generados.

Debido a los promisorios resultados obtenidos, en la primera campaña de adquisición, con los métodos de agrupamiento, se expandió la evaluación a otros métodos de agrupamiento que serán descritos in extenso en el apartado 2.2.3.3

2.2.2. Imágenes obtenidas de forma estática

Los experimentos de campo fueron llevados a cabo en un viñedo (*V. vinifera* L. cv. Tempranillo) ubicado en in Ollauri, La Rioja (España) en octubre de 2010. Las vides fueron fotografiadas usando una cámara digital montada en un trípode dispuesto en forma perpendicular a la espaldera a 2 m desde el eje de la hilera y 1.05 m sobre el suelo. Una pantalla de color blanco fue dispuesta detrás de la hilera para evitar confundir la hilera en estudio con las ubicadas detrás. Las imágenes fueron capturadas con una resolución de 3504x2336 píxeles y luego reducidas a 640x480 píxeles y 320x240 píxeles.

2.2.3. Imágenes obtenidas de forma Dinámica.

2.2.3.1. Análisis Imágenes Infrarrojas

A partir de análisis preliminares se descartó el uso de las imágenes RGIR de modo directo, debido a la dificultad de identificar correctamente las estructuras presentes en el viñedo. Por ello un conjunto de 200 imágenes RGIR fueron transformadas a imágenes en las que cada píxel representa un índice espectral.

Índices Multiespectrales. Los índices empleados fueron: el Índice de Vegetación de Diferencia Normalizada NDVI (Ecuación 4) y el Índice de Vegetación Ajustado al Suelo (SAVI) (Ecuación 5; Chuvienco *et al.*, 2002) ya que estos índices permiten separar vegetación del fondo o suelo.

$$NDVI = \frac{(IR - R)}{(IR + R)}$$

Ecuación 4. Índice de vegetación de diferencia normalizada.

$$SAVI = \frac{(IR - R)}{(IR + R + L)}(1 - L)$$

Ecuación 5. Índice de vegetación ajustado al suelo.

Donde:

IR es el canal infrarrojo, R el canal rojo.

L corresponde a un factor de corrección para evitar el efecto del brillo del suelo.

2.2.3.2. Análisis Imágenes visibles.

Como se demostró en la primera campaña de adquisición de imágenes el algoritmo Fuzzy C-Means, posee un buen desempeño como clasificador, sin embargo este algoritmo no es adecuado para imágenes con ruido, como las obtenidas en algunas condiciones de terreno; por otra parte tiene tiempos de ejecución mayores, si se le compara con otros algoritmos de agrupamiento como K-means.

Por estos motivos diversos métodos han sido propuestos para mejorar estas debilidades. Métodos como Possibilistic C-Means (PCM), Fuzzy Possibilistic C-Means (FPCM), Robust Fuzzy Possibilistic C-Means (RFCM) y Fuzzy C-Means con el algoritmo de Gustafson-Kessel (FCM-GK), Fuzzy c-Mean con Restricciones Espaciales (FCMSC por sus siglas en inglés), Kernelled Robust Fuzzy C-means (KRFCM). Todos ellos fueron evaluados en este trabajo para determinar su viabilidad de uso en tiempo real.

2.2.3.3. Selección de algoritmos de clasificación difusos

Las imágenes segmentadas con los métodos antes descritos fueron evaluadas usando varios parámetros de calidad como tasa de áreas correctamente clasificadas y tiempo de ejecución.

Como primer paso, las imágenes fueron transformadas a los espacios de color, HSI, CMYK, $L^*a^*b^*$, XYZ y Ohta (Busin et al., 2009). Este paso es llevado a cabo para determinar qué espacio de color y que canales dentro de estos espacios son los más adecuados para la segmentación de imágenes de vides empleando métodos de clasificación difusos.

Como segundo paso, los siguientes algoritmos de clasificación difusos fueron aplicados sobre las imágenes: Fuzzy C-Means (FCM), Possibilistic C-Means (PCM), Fuzzy Possibilistic C-Means (FPCM), Robust Fuzzy Possibilistic C-Means (RFCM) y Fuzzy C-Means con el algoritmo de Gustafson-Kessel (FCM-GK) y comparados. Estos algoritmos fueron seleccionados debido a ventajas teóricas así como para determinar sus desventajas cuando son aplicados bajo condiciones de campo. Para esta comparación se configuró los algoritmos con ocho grupos objetivo.

Como tercer paso, estos algoritmos fueron aplicados sobre las mismas imágenes, pero esta vez se le añadió un 5% de ruido (tipo “sal y pimienta”). Esta modificación fue realizada para simular imágenes adquiridas en condiciones de campo, condiciones como cámaras de baja resolución, vibración y polvo.

Como cuarto paso, los algoritmos fueron comparados contra una segmentación realizada manualmente. Se emplearon como parámetros de desempeño: exactitud en área y tiempo de ejecución.

2.2.3.4. Algoritmos de agrupamiento difuso evaluados

2.2.3.4.1. Fuzzy C-Means (FCM).

El algoritmo FCM asigna valores de grado de pertenencia a cada punto. Valores que están inversamente relacionados con la distancia relativa del punto a los prototipos o centros de las clases en el modelo FCM (Ganesan & Rajini, 2010). En FCM, dado un conjunto de datos $X = (x_1, x_2, \dots, x_n)$, la cercanía de cada dato, x_k , al centro del grupo, v_i , es definido como el grado de pertenencia (u_{ik}) de x_k , al grupo i de X , minimizando la siguiente función objetivo:

$$J_m(U, V) = \sum_{i=1}^c \sum_{k=1}^n \mu_{ik}^m \|x_k - v_i\|^2$$

Ecuación 6. Función objetivo de FCM en términos de los vectores x y v .

Donde X_k pertenece a un conjunto de N datos no clasificados; V_i son los centros de los grupos y $m = [1, \infty]$ es el exponente de pesos el cual determina que tan difusos serán los grupos resultantes. $U = [u_{ik}]$ es una matriz de $c \times n$, donde u_{ik} es el grado de pertenencia de x_k al grupo i , $\sum_{i=1}^c u_{ik} = 1, \forall k = 1, 2, \dots, n$ y los centros de los grupos son calculados como:

$$V_i = \sum_{k=1}^n \mu_{ik}^m X_k / \sum_{k=1}^n \mu_{ik}^m$$

Ecuación 7. Cálculo de centros de los grupos en FCM.

Y el grado de pertenencia:

$$u_{ik} = 1 / \sum_{j=1}^c \left(\frac{\|x_k - v_i\|}{\|x_k - v_j\|} \right)^{2/(m-1)}$$

Ecuación 8. Cálculo del grado de pertenencia en FCM

2.2.3.4.2. Algoritmo Possibilistic C-Means (PCM)

El algoritmo PCM considera el problema de agrupamiento desde el punto de vista de la teoría de las posibilidades (Krishnapuram & Keller, 1993). La aproximación adoptada en PCM difiere de la del algoritmo FCM, ya que el grado de pertenencia resultante puede ser interpretado como el grado de posibilidad (o compatibilidad) de los puntos, de pertenecer a una clase dada, en lugar de la probabilidad como en FCM. El algoritmo PCM, produce simultáneamente los valores de y grado de pertenencia. Los puntos atípicos (Outliers) tienen bajos valores de compatibilidad tipicidad y son eliminados automáticamente por el algoritmo.

La función objetivo en PCM es:

$$P_m(T, V; X, \gamma) = \sum_{i=1}^n \sum_{k=1}^c t_{ik}^m d_{ki}^2 + \sum_{i=1}^c \gamma_i \sum_{k=1}^n (1 - t_{ki})^m$$

Ecuación 9. Función objetivo en PCM.

Donde t_{ki} es la tipicalidad de x_k respecto al grupo i , T es la matriz de tipicalidad, $T = [t_{ik}]_{NC}$, d_{ik} es la distancia medida entre x_k y c_i , y γ_i denota una constante definida por el usuario $\gamma_i > 0, 1 < i < c$. Usando una optimización aproximada (AO) de P_m , el algoritmo PCM-AO encuentra las condiciones necesarias para una solución de la Ecuación 9, para $1 \leq i \leq c, 1 \leq k \leq N$, como sigue:

$$t_{ki} = \frac{1}{\left(1 + \frac{d_{ik}}{\gamma_i}\right)^{\frac{1}{m-1}}}, \forall i, k$$

Ecuación 10. Cálculo de tipicalidad en PCM

$$v_i = \frac{\sum_{k=1}^n t_{ki}^m x_k}{\sum_{k=1}^n t_{ki}^m}, \forall i$$

Ecuación 11. Cálculo de los centroides en PCM.

El algoritmo PCM resuelve la **Ecuación 9** con la **Ecuación 11** y añade las siguientes condiciones sobre $\{\gamma_i\}$:

$$\gamma_i = K \frac{\sum_{k=1}^n u_{ki}^m d_{ki}^2}{\sum_{k=1}^n u_{ki}^m}, K > 0$$

Ecuación 12. Cálculo del coeficiente γ en PCM.

Donde u_{ki} son los grados de pertenencia obtenidos al usar el algoritmo FCM y donde el valor más comúnmente usado es $K = 1$.

2.2.3.4.3. Fuzzy Possibilistic C-Means.

Empleando la misma notación que la usada en FCM, u_{ik} es el grado de pertenencia del punto x_k , al grupo i , mientras que t_{ik} es el valor de tipicalidad compatibilidad de x_k , en

el grupo i . Así, el objetivo del modelo FPCM es encontrar una partición de X , en c subgrupos difusos, mediante la minimización de la siguiente función objetivo (Pal *et al.*, 1997):

$$J_{m,\eta}(U, T, V) = \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik}^m + t_{ik}^\eta) \|x_k - v_i\|^2$$

Ecuación 13. Función objetivo para FPCM.

Sujeto a las siguientes restricciones $m > 1, \eta > 1, 0 \leq u_{ik}, t_{ik} \leq 1, \sum_{i=1}^c u_{ik} = 1, \forall k = 1, 2, \dots, n$ y $\sum_{k=1}^n t_{ik} = 1, \forall i = 1, 2, \dots, c$. donde m y η son exponentes de peso.

Bajo las condiciones y restricciones antes establecidas en los problemas de optimización c-means, obtendremos las condiciones, de primer orden, necesarias para extremar $J_{m,\eta}(U, T, V)$ en términos del teorema de los multiplicadores de Lagrange, como sigue:

$$t_{ik} = \frac{1}{\sum_{j=1}^n \left(\frac{d_{ik}}{d_{jk}} \right)^{2/m-1}}, \forall i, k$$

Ecuación 14. Cálculo de la tipicalidad en FPCM.

$$v_i = \frac{\sum_{k=1}^n (\mu_{ik}^m + t_{ik}^\eta) x_k}{\sum_{k=1}^n (\mu_{ik}^m + t_{ik}^\eta)}, \forall i$$

Ecuación 15. Cálculo de los centroides en FPCM.

donde d_{ik} es la distancia del punto x_k al prototipo v_i , i.e.:

$$d_{ik} = \|x_k - v_i\| = (x_k - v_i)^T A (x_k - v_i)$$

Ecuación 16. Cálculo de distancia euclidiana en FPCM.

donde A es la matriz identidad, d_{ik} , representa la distancia euclidiana, que a su vez representa la similitud entre puntos y prototipos (centros de grupos).

2.2.3.4.4. Robust Fuzzy-Possibilistic C-Means.

El algoritmo FPCM, intenta resolver el problema de sensibilidad al ruido que afecta al algoritmo FCM. Por otra parte resuelve el problema de grupos coincidentes que afecta a PCM, ya que mantiene la restricción de que la suma de los grados de pertenencia de un punto, a todos los grupos, sumen uno. Pero el algoritmo FPCM mantiene el uso la distancia inducida por norma, por lo que no posee suficiente robustez.

Con el propósito de mejorar el desempeño sobre datos ruidosos del tradicional algoritmo FPCM, se considerará un tipo de distancia núcleo-inducida. Basado en estas modificaciones, Yong *et al.* (2008) proponen un nuevo algoritmo denominado algoritmo Robust Fuzzy-Possibilistic C-Means (RFPCM).

Suponga que un punto $x_k \in R^S$, $k = 1, 2, \dots, n$, que es transformado desde el espacio original, a un espacio de características H con un tipo de mapeo no lineal, Φ , se convierte a la siguiente forma $\Phi(x_1), \Phi(x_2), \dots, \Phi(x_n)$. Así el producto interno, en el espacio original, puede ser expresado a través del núcleo Mercer (Girolami, 2002).

$$K(x_k, x_j) = (\Phi(x_k) \cdot \Phi(x_j))$$

Ecuación 17. Producto interno expresado a través del núcleo Mercer.

La distancia euclidiana en el espacio de las características puede ser expresada como sigue:

$$d_H(x, y) = \sqrt{\|\Phi(x) - \Phi(y)\|^2} = \sqrt{\Phi(x) \cdot \Phi(x) - 2\Phi(x) \cdot \Phi(y) + \Phi(y) \cdot \Phi(y)}$$

Ecuación 18. Distancia euclidiana expresada en el espacio de las características en RFPCM.

Así, la función objetivo del algoritmo FPCM puede ser modificada del siguiente modo (Yang & Tsai, 2008):

$$J_{m,\eta}(U, T, V) = 2 \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik}^m + t_{ik}^\eta) (\|\Phi(x_k) - \Phi(v_i)\|^2)$$

Ecuación 19. Función objetivo del algoritmo RFPCM.

Sujeto a las restricciones $m > 1, \eta > 1, 0 \leq u_{ik}, t_{ik} \leq 1, \sum_{i=1}^c \mu_{ik} = 1, \forall k, \sum_{k=1}^n t_{ik} = 1, \forall i$.

Donde m y η son ambos exponentes de pesos. Combinado las ecuaciones (18) y (19), se obtiene:

$$\|\Phi(x_k) - \Phi(v_i)\|^2 = K(x_k, x_k) + K(v_i, v_i) - 2K(x_k, v_i)$$

Ecuación 20. Cálculo de distancia en el espacio de las características en RFPCM.

Usando la función gaussiana $K(x, y) = e^{-\|x-y\|^2/\sigma^2}$, como núcleo, dado $K(x_k, x_k) = 1$ y $K(v_k, v_k) = 1$.

Por tanto, la Ecuación 19 puede ser transformada en la siguiente forma a través de su nucleación (Kernelization).

$$J_{m,\eta}(U, T, V) = 2 \sum_{i=1}^c \sum_{k=1}^n (\mu_{ik}^m + t_{ik}^n)(1 - K(x_k, v_i))$$

Ecuación 21. Función objetivo expresada en términos del kernel de Mercer.

Bajo las mismas condiciones del algoritmo FPCM, se tendrán las condiciones de primer orden necesarias para optimizar $J_{m,\eta}(U, T, V)$ en términos del algoritmo de los multiplicadores de Lagrange.

$$u_{ik} = 1 / \sum_{j=1}^c \left(\frac{1 - K(x_k, v_i)}{1 - K(x_k, v_j)} \right)^{1/m-1}, \forall i, k$$

Ecuación 22. Grado de pertenencia en RFPCM.

$$t_{ik} = 1 / \sum_{j=1}^n \left(\frac{1 - K(x_k, v_i)}{1 - K(x_j, v_i)} \right)^{1/m-1}, \forall i, k$$

Ecuación 23. Cálculo de tipicidad en RFCM.

$$v_i = \frac{\sum_{k=1}^n (\mu_{ik}^m + t_{ik}^\eta) K(x_k, v_i) x_k}{\sum_{k=1}^n (\mu_{ik}^m + t_{ik}^\eta) K(x_k, v_i)}, \forall i$$

Ecuación 24. Cálculo de centroides en RFPCM.

Para este análisis se usó σ , como el máximo σ , de los canales a^* y b^* del espacio de color CIE 1976 $L^*a^*b^*$. Se emplearon los parámetros $m = 2$ y $\eta = 2$, como los usados por Wu & Zhou, (2006).

2.2.3.4.5. Gustafson-Kessel Clustering

La principal característica del algoritmo de Gustafson-Kessel (GK) es la adaptación local de la métrica de la distancia a la forma del grupo, a través de la estimación de la matriz de covarianza del grupo y adaptando, en consecuencia, la matriz de distancia-inducida (Kotsiantis, 2007). El Algoritmo FCM-GK se basa en la optimización iterativa de un funcional objetivo del tipo c-means:

$$J_m(U, V, \{A_i\}) = \sum_{i=1}^c \sum_{k=1}^n \mu_{ik}^m d_{ik}^2$$

La distancia inducida por norma d_{ikA_i} como en la Ecuación. 16.

La métrica de cada grupo está definida por una matriz, A_i , de norma-inducida local, la cual es usada como una de las variables de optimización del funcional. Esto le permite a la distancia inducida por norma adaptarse a la estructura topológica local de los datos. La minimización del funcional objetivo de GK es lograda mediante el uso del método de optimización alterna (*Alternating optimization*, AO) según el conocido algoritmo propuesto por Babuka *et al.*, (2002), expresado como sigue:

Para $l = 1, 2, \dots, s$ donde l representa cada paso en el proceso de iteración y s es el máximo número de iteraciones permitido (criterio de parada)

$$v_i^{(l)} = \frac{\sum_{k=1}^n (u_{ki}^{(l-1)})^m x_k}{\sum_{k=1}^n (u_{ki}^{(l-1)})^m}, 1 \leq i \leq k$$

Ecuación 25. Cálculo de los centroides en FCM-GK.

$$F_i = \frac{\sum_{k=1}^n \left(u_{ik}^{(l-1)}\right)^m \left(x_k - v_i^{(l)}\right) \left(x_k - v_i^{(l)}\right)^T}{\sum_{k=1}^n \left(u_{ik}^{(l-1)}\right)^m}, 1 \leq i \leq k$$

Ecuación 26. Matriz de covarianza del grupo i .

$$d_{ikA_i}^2 = \left(x_k - v_i^{(l)}\right)^T \left[\rho_i \det \left(F_i^{1/n} F_i^{-1}\right)\right] \left(x_k - v_i^{(l)}\right), 1 \leq i \leq k$$

Ecuación 27. Cálculo del cuadrado de la distancia en FCM-GK.

para $1 \leq k \leq N$

si $d_{ik} > 0$

para $1 \leq i \leq K$,

$$u_{ik}^{(l)} = \sum_{j=1}^k \left(\frac{d_{ik}}{d_{jkA_i}}\right)^{-2/(m-1)}$$

Ecuación 28. Calculo del grado de pertenencia en FCM-GK.

De otro modo $u_{ik}^{(l)} = 0$, si $d_{ikA_i} > 0$, y $u_{ik}^{(l)} \in [0,1]$ con $\sum_{i=1}^k u_{ik}^{(l)} = 1$.

Hasta que $\|U^{(l)} - U^{(l-1)}\| < \epsilon$.

Este algoritmo fue implementado en esta tesis usando la estimación de covarianza mejorada propuesta por Babuska *et al.*, (2002). Para los efectos de este análisis se usó $m = 2$ y $\epsilon = 0.0001$.

2.2.3.4.6. Fuzzy c-Mean con Restricciones Espaciales y Contornos de Regiones.

Se evaluó el uso del método *denominado* Fuzzy c-mean con restricciones espaciales (FCMSC, por sus siglas en inglés) propuesto por Chuang *et al.*, (2006) e implementado con las mejoras propuestas por Li *et al.*, (2011) denominadas “Integrating spatial fuzzy clustering with level set methods”. Estos autores señalan que una de las características

más relevantes en las imágenes es que los píxeles vecinos están altamente correlacionados. Es decir, píxeles vecinos poseen valores de características similares y por tanto existe una alta probabilidad de que pertenecen al mismo grupo. Esta relación espacial es importante para los métodos de agrupamiento, pero no se utiliza en el algoritmo estándar FCM.

Así, para explotar la información espacial, se define una función espacial como:

$$h_{ij} = \sum_{k \in NB(x_j)} u_{ik}$$

Ecuación 29. Función espacial añadida a la función objetivo en el cálculo del grado de pertenencia.

Donde $NB(x_j)$ representa una ventana cuadrada centrada en el pixel x_j en el dominio espacial. Al igual que la función de pertenencia de FCM, la función espacial h_{ij} representa la probabilidad de que x_j píxel pertenezca a i -ésimo grupo. La función espacial de un pixel en un grupo es grande si la mayoría de su vecindad pertenece al mismo grupo. La función espacial se incorpora en la función de pertenencia de FCM de la siguiente manera:

$$u'_{ij} = \frac{u_{ij}^p h_{ij}^q}{\sum_{k=1}^c u_{kj}^p h_{kj}^q}$$

Ecuación 30. Cálculo del grado de pertenencia en FCM con restricciones espaciales.

donde p y q son parámetros para controlar la importancia relativa de las dos funciones. En las pruebas realizadas se evaluó las imágenes con $p=2$ y $q=0,5$.

El agrupamiento es un proceso de dos pasos en cada iteración. El primer paso es el mismo que en FCM estándar para calcular la función de pertenencia en el dominio espectral. En el segundo paso, la información de pertenencia de cada pixel es mapeada al dominio espacial, y la función espacial se calcula a partir de eso. Una vez terminado el agrupamiento espacial las mejoras propuestas por Li *et al.*, (2011) incluyen un paso adicional que consiste en generar regiones de contorno que encierren las regiones creadas en alguna clases que sea de interés, en este caso la detección de contorno

llevada a cabo mediante el método “Level Set”, configurado con un patrón de búsqueda de 5x5 píxeles y una constante de regulación de la función de Dirac $\epsilon=1,5$, un factor de ponderación de suavizado Gaussiano $\sigma=2$ y un criterio de detención de 200 iteraciones.

2.2.3.4.7. Kernelled Robust Fuzzy C-means

En cuanto al algoritmo Kernelled Robust Fuzzy C-means (KRFCM), propuesto por Li & Shen, (2008), éste adopta una nueva métrica inducida mediante núcleos (kernel) en el espacio de datos para reemplazar la métrica original de norma euclidiana utilizada por FCM, con el objetivo de realizar agrupamiento de regiones no esféricas. Esto se consigue al reemplazar el producto interior de FCM con una función 'kernel' apropiada, lo que posibilita realizar implícitamente un mapeo no lineal, a un espacio de características de alta dimensión, sin aumentar el número de parámetros.

La función objetivo a minimizar es:

$$J = 2 \sum_{j=1}^n \sum_{i=1}^c u_{ij}^2 (1 - K(x_j - v_i)) + \sum_{j=1}^n a_j \sum_{i=1}^c u_{ij} (1 - u_{ij})$$

Ecuación 31. Función objetivo a minimizar en KRFCM.

Donde $K(x, y)$ es tomada como la función generalizada de base radial (GRBF generalized radial basis function). J puede ser minimizado del mismo modo que FCM.

Si $X = \{x_1, x_2, \dots, x_n\}$ denota una imagen con N píxeles a ser dividida en c ($2 \leq c \leq n$) grupos, donde x_j representa el valor de la característica, el algoritmo es una optimización iterativa que minimiza la función de coste y se define por la Ecuación 31, con las siguientes restricciones:

$$\sum_{i=1}^c u_{ij} = 1, u_{ij} \in [0,1], 1 \leq j \leq i \leq c$$

Ecuación 32. Restricciones a la función de pertenencia en KRFCM

Donde:

$$u_{ij} = \frac{1}{\sum_{k=1}^c \frac{2(1 - K(x_j - v_i)) - a_j}{2(1 - K(x_j - v_k)) - a_j}}$$

Ecuación 33. Grado de pertenencia en KRFCM.

$$v_i = \frac{\sum_{j=1}^n u_{ij}^2 K(x_j, v_i) x_j}{\sum_{j=1}^n u_{ij}^2 K(x_j, v_i)}$$

Ecuación 34. Cálculo de los centroides en KRFCM.

2.3. Tercera campaña de adquisición

2.3.1. Estrategia de Aceleración de Convergencia e Identificación de Grupos.

En la segunda campaña de adquisición de imágenes se determinó que los algoritmos con mejor desempeño como clasificadores eran FCM-GK, seguido de FCM, pero con una velocidad de ejecución no adecuada para tiempo real, mientras que K-means demostró tener un desempeño aceptable (85-88%) pero además mostró una mayor velocidad de ejecución (del orden de 6 veces), por lo que en esta etapa de la investigación se exploró la idea de aprovechar las mejores características de cada algoritmo para lograr la mejor relación desempeño/velocidad. Además se exploró una vía de identificar los grupos generados mediante máquinas de vectores de soporte (SVM) y se pretende resolver el problema del doble conteo de características mediante la creación de una imagen mosaico empleando SIFT.

Por ello en esta sección se describe una metodología de 10 pasos, basada en procesamiento de imágenes visibles, orientada a la cuantificación en tiempo real de uvas y hojas. El objetivo último de esta metodología es demostrar la viabilidad de caracterizar (identificar y cuantificar) la cubierta del viñedo, en tiempo real bajo condiciones de terreno, combinando técnicas de agrupamiento como Fuzzy C-Means con el algoritmo

de Gustafson Kessel (FCM-GK) con K-Means para agrupar píxeles y Máquinas de Vectores de Soporte para identificar los grupos creados.

Con el fin de crear un marco metodológico suficientemente flexible y robusto, ese llevó a cabo la segmentación de imágenes usando técnicas de agrupamiento no supervisadas, para posteriormente identificar los grupos creados a través de un clasificador basado en SVM. Sin embargo, también se han implementado dos características adicionales tendientes a mejorar los resultados. La primera es un procedimiento de inicialización para acelerar el proceso de agrupamiento y la segunda es la creación de un mosaico (por unión de imágenes) de cada hilera de viñedo fotografiada, a fin de evitar el doble conteo de racimos de uva.

Bajo la premisa de desarrollar un marco metodológico que permita realizar varias tareas, algunas de ellas en tiempo real, se debe efectuar una cuidadosa selección del espacio de color, resolución de las imágenes y de las técnicas de clasificación e identificación.

En esta sección se presenta una visión general acerca de cómo, en este marco metodológico, se realiza la selección del espacio de color, se implementan los algoritmos y procedimientos para clasificación e identificación, junto con presentar además sus ventajas, desventajas y sinergias cuando son combinados.

Casi finalizada la campaña de adquisición se incorporó una cámara de profundidad la que permite a través de la distancia a la hilera filtrar que objetos pertenecen a la hilera de interés y de este modo eliminar la pantalla dispuesta para tal propósito. La metodología específica para aplicar esta cámara se describe al final de éste capítulo en el apartado 2.4, mientras que una síntesis metodológica de la solución finalmente propuesta se detalla en el apartado 2.4.1.

2.3.1.1. Clasificador basado en Máquinas de Vectores de Soporte

Los sistemas de aprendizaje clásicos como las redes neuronales artificiales (ANN, por sus siglas en inglés) tienen algunas debilidades teóricas, por ejemplo, las metodologías denominadas "back-propagation" por lo general sólo convergen a una solución óptima local. Para resolver este problema se desarrolló la metodología basada en máquinas de vectores de soporte (SVM, por sus siglas en inglés), las que se caracterizan por siempre encontrar un mínimo global, y por ser menos propensa a sobre-ajustar el resultado final (Kotsiantis, 2007). Las SVM requieren grupos pequeños de datos para ser entrenadas y también tienen una buena capacidad de generalización.

SVM puede descartar todos los datos que no provean información válida, a diferencia por, ejemplo, de kNN, que requiere todo el conjunto de datos de entrenamiento (referencia) para realizar la clasificación.

En el caso de un espacio de entrada no separable linealmente (que es comúnmente el caso de las complejas imágenes de viñedo), los datos de entrada pueden ser mapeados a otro espacio de características que posea una dimensión mayor, en la que los puntos (datos) sean linealmente separables

Las SVM proporcionan un mecanismo genérico para adaptar la superficie del hiper plano a los datos a través del uso de una función *kernel*. Las funciones del *kernel* más comunes son: lineal, cuadrática, gaussiana de base radial, polinómicas y de "perceptrón multicapa". En este trabajo se seleccionó un *kernel* polinómico (orden 5), ya que aporta flexibilidad suficiente al clasificador. La flexibilidad del clasificador aumenta con el grado del polinomio del *kernel*. Por otra parte, las SVM son un método de clasificación binario que determina si un dato pertenece a una clase o no. Sin embargo, en este caso se enfrenta a un problema de clasificación multiclase. Para resolver este problema, el mejor enfoque consiste en utilizar una combinación de varios clasificadores SVM binarios (Duan y Keerthi, 2005).

2.3.1.2. Creación de Mosaicos Mediante Transformada de Características de Escala Invariante (SIFT).

La transformada de características de escala-invariante (SIFT) es un método para la extracción de características invariantes distintivas de las imágenes utilizada para realizar correspondencia fiable entre diferentes vistas de un objeto o escena (Lowe, 2004). Las características son invariantes al escalado y rotación de la imagen, y han mostrado proveer correspondencias robustas en un amplio rango de distorsiones afines: cambio de punto de vista en 3D, adición de ruido, y cambio en la iluminación. Por lo tanto, SIFT es muy adecuado para que solapen imágenes de viñedo adquiridas en condiciones de campo, y de esta manera crear mosaicos de la totalidad de las hileras.

2.3.1.3. Descripción del marco metodológico final.

Los pasos necesarios para caracterizar el viñedo pueden ser sintetizados en el siguiente marco metodológico para el análisis de las imágenes RGB mostrado en la **Figura 5** y descritos en detalle como sigue:

Figura 5. Esquema del marco metodológico propuesto para el análisis de imágenes RGB.

Paso 1. Adquisición de imagen: Un conjunto de 40 imágenes de la parte aérea del viñedo fueron tomadas cada 0,4 m con una cámara digital (modelo Canon 550D). La viña es un campo experimental, propiedad de la Universidad Politécnica de Madrid, situado a 40 ° 26'31.33 "de latitud N, 3 ° 44'8.70" O de longitud, con un espacio entre hileras de 2 m, y 1,5 m entre vides sobre la hilera.

El sistema de conducción es un alambre doble estándar. Las imágenes fueron tomadas de forma normal a la cubierta, a de 1 m de distancia desde el eje de hilera y 1,1 m sobre el suelo. Las imágenes fueron adquiridas en diferentes horas para evaluar el efecto de iluminación. Un telón blanco se colocó detrás de la cubierta para evitar los efectos producidos por la confusión con la vegetación de fondo. Las imágenes fueron capturadas a una resolución de 2592x1728 píxeles y después de eso reducida a 800x600, 480x320 y 240x320 para evaluar la precisión y tiempos de ejecución.

Paso 2. Selección de canales y Espacio de color: Las imágenes fueron transformadas al espacio de color $L^*a^*b^*$, pero sólo los canales a^* y b^* fueron utilizados para el proceso de agrupamiento, ya que mantienen la información cromática y ofrecen un mejor rendimiento en segmentación de imágenes viña (Correa *et al.*, 2011). El canal L^* fue eliminado para evitar los efectos de producidos por las diferencias de iluminación.

Paso 3. Proceso de agrupamiento Etapa I: Se aplicó un algoritmo de agrupamiento difuso sobre los canales a^* y b^* . Específicamente se aplicó Fuzzy C-Means con algoritmo de Gustafson-Kessel (FCM-GK), el que fue seleccionado debido a sus ventajas cuando se aplican bajo condiciones de terreno encontradas en la agricultura (Correa *et al.*, 2011).

El número de grupos se estableció en ocho, y corresponde grupos de uvas, tronco, fondo oscuro, fondo brillante y cuatro clases de hojas de diversas edades.

Esta rutina lee la primera imagen adquirida y genera ocho centros de grupos como salida. Este proceso también se denomina "creación de semillas", en referencia a que los resultados de este proceso de agrupamiento (ocho centros de clases) serán usados

como semillas para el algoritmo K-means, en la Etapa II. Este proceso se llevó a cabo utilizando la herramienta de computación paralela de Matlab.

Paso 4. Proceso de agrupamiento Etapa II: Los centros de grupos generados por FCM-GK (Paso 3) se utilizan como semillas para la clasificación con K-mean, con el fin de acelerar la convergencia del algoritmo K-means y también para introducir un sesgo en las nuevas imágenes procesadas.

La inicialización se actualiza cada cuatro imágenes con el fin de evitar cualquier desviación producida por las variaciones de las vides. Para lograr esto, el proceso de inicialización se ejecuta en paralelo con el proceso principal (**Figura 5**).

Paso 5. Proceso de Identificación: Con el fin de identificar los grupos generados por K-means se implementaron dos clasificadores, uno basado en redes neuronales artificiales y otro basado en SVM. La red implementada es una Feed-forward Back-propagation Neural Network (FFBP NN). Que fue implementada con 8 neuronas en la capa de entrada, 10 en la capa oculta, y 8 para la capa de salida (una neurona por grupo), como función de entrenamiento se empleó Levenberg-Marquardt. Esta red fue entrenada con 10 imágenes y validada usando otras 40.

Las SVM son clasificadores binarios; por tanto, se implementó una SVM por clase utilizando el método “uno contra todos” y utilizando la estrategia “el ganador se lleva todo”. La SVM fue entrenada usando 10 imágenes y probada con 40. Esta metodología de aprendizaje efectuada por una rutina creada utilizando la caja de herramientas de Bioinformática de Matlab.

Paso 6. Fusión de grupos: Una vez identificados los grupos, si alguna clase no está presente en la imagen, por ejemplo no hay uvas, al menos un grupo tendrá una etiqueta duplicada. En ese caso, estos grupos se fusionan y son considerados como un solo clúster con el fin de llevar a cabo las operaciones morfológicas para luego sumar sus áreas.

Paso 7. Proceso de filtrado: Con el fin de eliminar los errores de clasificación en los racimos de uvas, se realizaron diversas operaciones.

En primer lugar, se aplicó una operación de cierre de la imagen para llenar "huecos" en el interior del Grupo Uvas; lo que ayuda a reducir el problema de la cohesión espacial de K-Means, ya que re-etiqueta los píxeles que se encontraban en los "hoyos".

En segundo lugar, los grupos con áreas de menos de 10 píxeles fueron eliminados, ya que se les consideró falsos positivos.

En tercer lugar, se calculó la *excentricidad* y *orientación* (de la elipse que tiene los mismos segundos-momentos que la región estudiada) utilizando la función REGIONPROPS perteneciente a la caja de herramientas "Image Processing Toolbox" de Matlab.

Se eliminaron los grupos de píxeles con orientaciones entre -30 a 30 grados y excentricidad entre [0,7-1]. La razón para aplicar este criterio se debe a que los racimos de uvas tienen formas predominantemente verticales (más largos que anchos).

Paso 8. Cálculo de las áreas. En este paso se calcula de modo preliminar las áreas de hojas y de racimos y se lleva a cabo con el fin de proporcionar la estimación de área instantánea. Estas áreas son útiles para hacer calibraciones locales cuando se realizan proyecciones de rendimiento. El número de píxeles de cada grupo fue contado y dividido por el número total de píxeles en la imagen para proveer porcentajes de cobertura, para posteriormente transformarlos en áreas expresadas en m².

Paso 9. Proceso de unión de imágenes: Se creó una imagen de mosaico, por hilera, usando SIFT con el fin de evitar traslape entre las imágenes y por tanto doble cómputo de áreas.

Paso 10. Corrección de áreas: La imagen de mosaico creada en el paso 7 se dividió en imágenes de 1m de ancho. Cada imagen se procesa para determinar el área de hojas y uvas como porcentaje del área total de la imagen; el número de píxeles de cada grupo

es contabilizado y dividido por el número total de píxeles en la imagen para expresarlos como porcentaje.

Esta colección de rutinas puede resolver varios problemas simultáneamente, esto es identificar y la cuantificar uvas, hojas y troncos. Sin embargo, no es factible realizar el proceso de unión de imágenes en tiempo real ya que es un proceso demandante de tiempo. Es por ello que, para efectos de evaluación del rendimiento, el proceso de unión se realiza mediante un pos-proceso.

2.4. Cámara de Profundidad.

Para eliminar el error producido por el efecto de confundir la hilera de interés con las hileras paralelas se abordó el problema de dos formas. En primera instancia se empleó una pantalla a modo de fondo para la hilera fotografiada, de esta manera se proveyó de un fondo homogéneo, que impedía visualizar las hileras paralelas.

De modo alternativo se empleó la información de profundidad proveniente de la cámara Kinect para Xbox 360 de Microsoft. A partir de la información provista por la cámara es posible eliminar de la imagen todos aquellos píxeles que se encuentren a una distancia superior a la distancia entre hileras.

Para ello se desarrolló un software en Microsoft Visual Studio 2013 empleando el "Kinect for Windows Software Development Kit". Este software permite almacenar y fusionar información de profundidad y de color proveniente de la cámara, a la vez que reducir la resolución de las imágenes para procesarlas en tiempo real, con el marco metodológico planteado en la sección anterior con la adición de la información de profundidad.

2.4.1. Síntesis de la solución final planteada.

La solución final planteada, que incluye el uso de la información de profundidad, puede ser sintetizada de la siguiente forma:

1. Se captura en paralelo una imagen de profundidad y una RGB empleando una cámara Kinect.

2. Se transforma la imagen RGB al espacio de color L^*a^*b y se reduce su tamaño a 340x420 píxeles.
3. Se usa la imagen de profundidad como filtro por distancia eliminando las hileras paralelas
4. Se realiza un agrupamiento con FCM-GK sobre los canales a^*b^* de la imagen filtrada.
5. Se realiza un agrupamiento con K-means usando como valores de inicialización los centros de los grupos creados por FCM-GK.
6. Se identifican los grupos Uva y Hojas mediante SVM.
7. Se fusionan grupos perteneciente a la misma clase.
8. Se filtra mediante operaciones morfológicas y mediante factores de forma, el grupo Uvas, para eliminar falsos positivos.
9. Se realiza la unión de imágenes mediante SIFT para evitar el doble cómputo de área de uvas.
10. Se calculan las áreas de uva y hojas cada 420 píxeles de las imágenes fusionadas mediante SIFT.

Un diagrama de flujo del algoritmo propuesto se presenta en la **Figura 6**, en la que se pueden apreciar en cuadros blanco dos de las aplicaciones propuestas.

Figura 6. Diagrama de flujo final de la metodología de análisis de imágenes RGB y de Profundidad propuesta.

2.4.1.1. Validación del Algoritmo.

Con el fin de validar los algoritmos de agrupamiento, diez regiones de interés (30x30 píxeles) obtenidas de cinco imágenes diferentes, fueron segmentados de forma manual y comparadas con el algoritmo de agrupamiento a diferentes resoluciones.

El criterio de selección de las imágenes fue escoger imágenes tan disimiles entre ellas como sea posible, para garantizar la capacidad de adaptación del algoritmo. Luego, la SVM fue validada contrastando las etiquetas asignadas, en contra de las etiquetas de los grupos etiquetados manualmente.

3. PREDICCIÓN DE RENDIMIENTO Y DE ÁREA FOLIAR.

3.1. Sección experimental

Para correlacionar los datos de área foliar y rendimiento estimado con mediciones reales de la planta, se desarrolló una detallada configuración experimental para la adquisición de imágenes, basado en defoliaciones sucesivas y raleo de racimos por etapas en vides individuales.

3.2. Sitio experimental

Los experimentos se llevaron a cabo en 2010 en un viñedo comercial de secano cv. Tempranillo (*Vitis vinifera* L.), situado en Casas Blancas, Cidamón (lat.42° 29 '8.83' 'N; lon. 2° 50' 22.57 " O; 181 m snm), La Rioja, España. Las vides de Tempranillo estaban injertadas en porta injertos 41B y plantadas en 2005 con una separación entre hileras y dentro de la hilera de 2,70 m x 1,15 m respectivamente.

Las viñas recibieron una poda de vigorización (12 yemas por vid) en cordón bilateral y guiado mediante alambrado en sistema VSP. El sistema de guía contaba con un alambre de soporte a 0,70 m, dos alambres a 1,00 m por encima del suelo para la protección contra daños por el viento, y un par de alambres móviles para posicionado de sarmientos a 1,45 m.

3.3. Defoliación, raleo de racimos y la evaluación del área foliar y fruta eliminados

Con el fin de proporcionar una correcta validación del método de clasificación de las imágenes, al momento de cosecha, se eligieron 10 vides al azar, y cada una de ellas fue individual y sucesivamente defoliada y los racimos cortados (raleados) en varias etapas como se muestra en la Tabla 5.

Después de cada etapa, el área de hojas y/o frutos cortados también fue registrada. De esta manera, se creó una gama de diferentes condiciones de área foliar y de racimos expuestos, para proporcionar una mejor validación de la metodología de análisis de imágenes.

Todas las hojas de cada vid son cortadas, sucesivamente: en primer lugar mediante la eliminación de las primeros 6 principales hojas basales (paso 1), a continuación, otras 6 (un total de 12 hojas) hojas basales principales (paso 3), y luego el resto de las hojas principales y laterales (defoliación completa, o paso 5).

El número de hojas retiradas en cada paso fue registrado y medido utilizando un medidor de área foliar (LI-3100C; Li-Cor, Lincoln, NE, EE.UU.).

Del mismo modo, cada vid fue sucesivamente cosechada mediante la remoción de algunos racimos: en primer lugar mediante la eliminación de todo tercer racimo (paso 2), luego cada segundo racimo restante (paso 4) y luego los restantes racimos (paso 6). El número de racimos cortados y su peso combinado se registró después de cada evento de remoción de racimo

Tabla 2. Descripción de las etapas de defoliación y raleo, durante la sesión de adquisición de imágenes de una vid individual

Imagen #	Eventos de manipulación de la cubierta	Imagen
I0	Estado inicial de la vid (Etapa 0)
I1	Remoción de las 6 primeras hojas basales por cada sarmiento (Etapa 1).
I2	Remoción de uno de cada tres racimos (Etapa 2).
I3	Remoción de 6 hojas basales adicionales por sarmiento (12 hojas removidas en total) (Etapa 3)

I4 Remoción de uno de cada tres racimos (Etapa 4)

I5 Remoción de las restantes hojas principales y laterales (Etapa 5)

I6 Remoción de los racimos restantes (Etapa 6)

Antes de cualquier defoliación o remoción de racimos, y después de cada etapa de manipulación de la cubierta, cada vid (10 viñas en total) fue. Téngase en cuenta que, cuando el proceso de defoliación se realizó en vides con alta densidad foliar, la distancia al follaje remanente aumenta y por consiguiente el tamaño de los objetos parece reducirse. Para corregir este problema, las imágenes fueron escaladas para adaptarse a las imágenes adquiridas a 2 m. De esta manera todas las imágenes representan la misma área.

3.4. Procesamiento de Imágenes (Algoritmo de clústeres)

Diversas mediciones de similitud entre los grupos en términos de múltiples características se han propuesto en la literatura (Correa et al., 2011), pero se ha encontrado que la distancia de Mahalanobis es la más adecuada en la mayoría de aplicaciones, y es ampliamente utilizada para el reconocimiento de patrones y análisis de datos (Son et al., 2010). De hecho es bien sabido que muchas medidas de distancia estándar, tales como la distancia variacional de Kolmogorov, la distancia Hellinger, la distancia de Rao, etc., son funciones aumentadas de la distancia de Mahalanobis bajo supuestos de normalidad y homocedasticidad (McLachlan, 1999).

Mahalanobis mide la similitud entre un grupo de muestra desconocida y uno conocido; toma en cuenta las correlaciones del conjunto de datos, y es invariante en escala. También considera el hecho de que las varianzas en cada dirección son diferentes, así como la covarianza entre variables.

La distancia de Mahalanobis entre dos vectores aleatorios $(\vec{x}, \vec{y})$ con la misma distribución, y matriz de covarianza S , se puede definir como

$$d(\vec{x}, \vec{y}) = \sqrt{(\vec{x} - \vec{y})^T S^{-1} (\vec{x} - \vec{y})}$$

Ecuación 35. Distancia de Mahalanobis.

La distancia de Mahalanobis aplicada a color estandariza la influencia de la distribución de cada característica teniendo en cuenta la correlación entre cada par de términos (Al-Otum, 2003).

En el caso de imágenes en color RGB S se calcula como:

$$S = \begin{bmatrix} \sigma_{RR} & \sigma_{RG} & \sigma_{RB} \\ \sigma_{GR} & \sigma_{GG} & \sigma_{GB} \\ \sigma_{BR} & \sigma_{BG} & \sigma_{BB} \end{bmatrix}$$

Ecuación 36. Matriz de covarianza.

Para nuestros propósitos $\vec{x}$ es un vector de tres dimensiones (R, G, B), que representa los píxeles de la imagen para ser procesados, $\vec{y}$ también es un vector tridimensional (R, G, B), que representa los píxeles de referencia (grupo referencia) para cada clase a ser identificada.

El efecto de S fue escalar la distancia a lo largo de cada característica. Para los propósitos de procesamiento de imágenes, cada canal R, G, B se considera como una característica.

La **Figura 6**, muestra una selección aleatoria de 2000 píxeles extraídos de una imagen de una vid típica (espacio de color RGB). Es necesario tener en cuenta que las varianzas

en cada dirección (R, G, B) son diferentes y por ello fue necesario implementar un clasificador que considerase tales diferencias, como Mahalanobis. La región de donde la distancia de Mahalanobis es constante alrededor de la media forma una elipse en el espacio 2D como se muestra en la **Figura 7 b)**.

a)

b)

c)

Figura 7. a) Selección aleatoria de 2000 píxeles de una imagen típica de una vid (imagen I0 en el espacio de color RGB). b) plano RG. Elipses representan puntos equidistantes. c) plano RB. d) Plano GB.

3.5. Implementación del software

En la metodología propuesta en este trabajo siete grupos de píxeles de referencia fueron seleccionados con el fin de generar la clasificación. Cada grupo representa una característica relevante de la cubierta de la vid. Las siete clases identificadas fueron: Hojas jóvenes (dos clases), las Hojas viejas (dos clases), Tronco (incluyendo sarmientos y tronco), Uvas y Fondo (o huecos en la cubierta). Sin embargo, si cualquiera de estas clases no está presente (dependiendo de la etapa de crecimiento), o una nueva clase aparece en la imagen, el número y/o las etiquetas de los grupos podrían ser modificados.

De esta manera, cada grupo de píxeles fue seleccionado manualmente desde un conjunto de 10 imágenes representativas. Para cada grupo de referencia se eligió un conjunto de 40 píxeles. Una vez seleccionados los píxeles de referencia, se calculó la distancia de Mahalanobis entre ellos y los píxeles de las imágenes a analizar. Dichos cálculos se efectuaron sobre un conjunto de 70 imágenes y sus píxeles asignados a la clase a la que tengan menor distancia. Del conjunto de 70 imágenes (7 imágenes por vid), las imágenes I0, I2 e I4 fueron utilizados para entrenar el modelo para la estimación del área foliar, ya que estas imágenes corresponden a los eventos de defoliación. Del mismo modo, I1, I3 e I5 fueron utilizadas para construir el modelo para la estimación del rendimiento.

Con el fin de implementar el algoritmo de clasificación y proveer una interfaz gráfica para el usuario, se desarrolló un software basado en Matlab 7.11. La **Figura 8** muestra una captura de pantalla del software, en el momento en que los píxeles de referencia son manualmente seleccionados y representados como un conjunto de barras en la parte inferior de la imagen. Estas barras se muestran para proveer información visual relativa a los píxeles seleccionados como referencia.

El número de píxeles utilizados como referencia, así como el nombre y el número de grupos pueden ser modificados por los usuarios, con el fin de ajustar el algoritmo para diferentes condiciones, tales como cambios de iluminación o nuevas características que no se habían observado previamente, por ejemplo, la hoja con coloración/decoloración producida por enfermedades.

Figura 8. Captura de pantalla del proceso de selección manual de los píxeles de referencia.

3.6. Etapas de la metodología de extracción de características.

La metodología implementada consta de 12 pasos que pueden ser resumidos como sigue:

- Etapa 1. Selección de imágenes representativas, que contengan tanta variabilidad (dentro de cada clase) como sea posible, y lectura de las imágenes seleccionadas.
- Etapa 2. Selección de 40 píxeles de referencia para cada clase de interés.
- Etapa 3. Lectura del directorio que contenga las imágenes a procesar
- Etapa 4. Procesamiento de una sección de una imagen (región de interés, ROI) o de la imagen completa. Selección de la ROI a procesar si es aplicable.
- Etapa 5. Para cada clase, calcular la distancia de Mahalanobis entre los píxeles de referencia y los de la imágenes/sección a analizar.
- Etapa 6. Asignar la clase a la que pertenece cada pixel basado en la regla de la menor distancia, es decir, se asigna el pixel a la clase cuya distancia sea menor.
- Etapa 7. Efectuar operaciones morfológicas sobre la clase Uva. Remover pequeños grupos de píxeles y relleno de “hoyos” dentro de los racimos de uva usando operaciones morfológicas de erosión y dilatación respectivamente
- Etapa 8. Asignación de píxeles a la clase Uva solo si se ubican en la mitad inferior de la imagen e.g., si la resolución de la imagen es de 800x600, los píxeles solo se consideraran como pertenecientes a la clase Uva si se encuentran entre las posiciones 400 y 800 en el eje vertical.

Etapa 9. Cálculo del número de píxeles de cada clase.

Etapa 10. Almacenamiento de los resultados numéricos en una hoja de cálculo.

Etapa 11. Almacenamiento de las imágenes de clases en un directorio preestablecido.

Etapa 12. Despliegue de las imágenes de clases en pantalla.

3.7. Validación del algoritmo

Se llevó a cabo un proceso de validación de estas imágenes específicas de la cubierta de la vid. Esta validación se realizó de forma manual, seleccionando algunas regiones de interés (ROI por sus siglas en inglés) en las imágenes que mostraban condiciones de colores y de iluminación representativas. Una vez seleccionada la ROI, el número de píxeles identificados para cada clase fue contado manualmente por un experto, tanto en el de la imagen original y como en la imagen resultante del proceso de agrupamiento.

3.8. Correlaciones, estimación del área de hojas y rendimiento

Para la estimación de área foliar y de rendimiento, el conjunto de imágenes fue dividido en dos grupos: el grupo de entrenamiento, en el que se utilizó dos tercios de los datos para generar el modelo, y el grupo de validación, al que se le asignó el tercio restante de los datos para propósitos de validación.

Para el grupo de entrenamiento, se generaron, para cada imagen, correlaciones lineales entre el número de píxeles de las clases identificados como hojas y el área foliar real de la vid y entre la clase uva y el rendimiento real en la vid, (SPSS (v15.0, IBM, NY, USA)).

Luego, estas correlaciones fueron utilizadas para estimar el área foliar y el rendimiento en el conjunto de imágenes de validación. Se calcularon correlaciones entre los valores estimados y observados (reales) para el área de hojas y rendimiento, para luego calcular los coeficientes de determinación (R^2) y raíz media del error cuadrático medio (RMSE).

VII. RESULTADOS

1. Primera campaña de adquisición

1.1. Métodos de clasificación por color

1.1.1. Selección del espacio de color.

Los resultados de convertir las imágenes a diferentes espacios de color se ejemplifican en las Figuras 9, 10, 11, 12 y 13.

De los espacios de color empleados los que mejores resultados muestran (como se demostrará en las secciones siguientes) son el $L^*a^*b^*$ y el CMYK. En el $L^*a^*b^*$ los canales a^* y b^* (magenta-verde y amarillo-azul) muestran mejores resultados como se puede apreciar en la **Figura 12 b)** y 12 c), en las que mediante una simple inspección visual es posible distinguir hojas en el primer caso y racimos en el segundo.

El canal H del espacio de color HSV permite identificar claramente los racimos, pero tiende fuertemente a confundirlos con tronco. El canal a^* presenta buen contraste entre hojas y el resto de las clases, mientras que el b^* entre racimos y el resto de clases.

En el espacio de color CMYK, el canal Y (amarillo) presenta mejor contraste entre racimos y hojas, como se puede observar en la **Figura 13 c)**, ya que dichos canales permiten discriminar de mejor modo vegetación, uvas y fondo.

a) Canal R.

b) Canal G.

Figura 9. Imagen de una vid en el espacio de color RGB y sus canales.

Figura 10. Imagen de una vid en el espacio de color HSV y sus canales.

c) Canal 3 $((2 \cdot G - R - B)/4)$.

d) Imagen en Otha.

Figura 11. Imagen de una vid en el espacio de color Otha y sus canales de color.

a) Canal L^* .

b) Canal a^* .

c) Canal b^* .

d) Imagen $L^*a^*b^*$.

Figura 12. Imagen de una vid en el espacio de color $L^*a^*b^*$ y sus canales.

a) Canal C

b) Canal M

Figura 13. Imagen de una vid descompuesta en los cuatro canales del espacio de color CMYK.

Merecen especial atención los canales L (luminosidad) del espacio $L^*a^*b^*$ y K (negro) del espacio de color CMYK, ya que estos canales guardan relación con el nivel de intensidad lumínica de los píxeles, lo que resulta útil a la hora de separar o unir conjuntos de píxeles de igual color pero distinta intensidad. Por otra parte el segundo canal del espacio de color de Ohta y el canal a^* del espacio de color $L^*a^*b^*$ permiten separar sarmientos del resto de los componentes de la imagen

1.1.2. Métodos de agrupamiento por color.

1.1.2.1. K-Means

En lo que se refiere a los algoritmos de clasificación, aplicados sobre imágenes en el espectro visible, K-Means, usando inicialización al azar, ocho clases y aplicado sobre el espacio $L^*a^*b^*$, es capaz de separar adecuadamente las uvas del resto de los componentes de la imagen -véase por ejemplo la **Figura 14 d)**- alcanzando un nivel de aciertos de un 85,1%. Sin embargo las hojas, en algunos de los escenarios evaluados, requieren un pos tratamiento para aumentar el porcentaje de aciertos, como se aprecia en la **Figuras 14 b)** y su posterior tratamiento en la **Figura 15**.

En las pruebas se evaluaron desde 4 a 20 grupos, siendo 8 la mínima cantidad de grupos necesaria para separar la clase uva. El aumento hasta 12 grupos no mostró una mejoría significativa en la separación de la clase Uva al compararla con los resultados obtenidos con 8 clases. En efecto, al evaluar con 12 grupos, el porcentaje de aciertos subió a 86%, manteniéndose constante hasta los 20 grupos. Mientras que los tiempos

de ejecución aumentaron un 14% al aumentar a 12 grupos y un 41% al aumentar a 20 grupos.

Sin embargo para determinar correctamente el área de hojas o de sarmientos, se debe aumentar la cantidad de clases, ya que como se observa en la **Figura 14 b)**, en la Clase 2 se mezclan píxeles que corresponden a fondo, hojas y sarmientos. Este efecto se reduce, como se señaló, al aumentar a 12 clases o al repetir la clasificación sobre la imagen generada por la Clase 2.

a) Clase 1

b) Clase 2

c) Clase 3

d) Clase 4

e) Clase 5

f) Clase 6

Figura 14. Clases generadas por K-Means usando la Distancia del Coseno sobre canales a^*b^* .

Figura 15. Resultado de aplicar K-Means sobre la Clase 2 de la Figura 14. Nótese que las imágenes b), c) y d) contienen hojas, mientras que la imagen e) contiene segmentos de sarmientos.

La estrategia de volver a aplicar el algoritmo de agrupamiento sobre alguna de las clases generadas resulta efectiva, como lo muestra la **Figura 15** en la que sarmientos y hojas son separados del fondo, aunque requiere de la identificación manual de la clase que requiere volver a ser agrupada.

Por otra parte esta estrategia de agrupamiento recursivo, aunque exitosa, requiere aumentar los tiempos de cómputo.

De las distancias usadas para evaluar K-Means, (coseno y euclidiana) la que mejor desempeño mostró fue la denominada del coseno, ya que las otras tienden a incluir el fondo en la clase que contiene a la uva. Ejemplo de dicho error de clasificación se puede apreciar en la **Figura 16**, en la que la clase uva tiene apenas un 61% de aciertos.

Figura 16. a) Imagen original. b) Grupo racimos generado por K-Means empleando la distancia euclidiana y ocho grupos.

Adicionalmente a K-means, se evaluaron los métodos denominados K-medoids, K-medians y la estrategia de aceleración de convergencia denominada K-means++.

En el caso de K-medoids y K-medians estas técnicas fueron seleccionadas por su mayor inmunidad al ruido, pero sin embargo K-medoids no presenta diferencias significativas en el desempeño como clasificador al compararla con K-means, en efecto para el grupo de imágenes evaluadas la diferencia es de un 0.5%, el que puede ser atribuido a la inicialización aleatoria de los algoritmos.

Además, debido a que sus tiempos de ejecución fueron mayores, en un 600% a los de K-means, esta técnica fue descartada como alternativa. K-medians por su parte, muestra el mismo desempeño que K-means configurado con la distancia euclidiana y sus tiempos de ejecución en promedio son levemente inferiores. Por ejemplo, con un conjunto de 50 imágenes de 320x240 píxeles K-means tarda 380 ms mientras que para K-medians 360ms.

Por otra parte la estrategia de inicialización K-means++ no presenta mejoras significativas en los tiempos de convergencia cuando se compara con la estrategia de inicialización al azar que se emplea normalmente en K-means, en efecto, al compararla, en promedio es un 29% más lenta (configuradas a ocho grupos y distancia euclidiana).

1.1.2.2. FCM

En cuanto a Fuzzy C-Means, esta técnica de clasificación mostró mejores resultados si se compara con K-Means, como se puede observar en los resultados de la Tabla 3. En efecto, a partir de una simple inspección visual es posible comparar los resultados de ambas técnicas en las **Figura 17 d)** y **17 e)**, en donde FCM incluso identifica racimos difíciles de visualizar en la imagen original.

Sin embargo esta mejora en el desempeño como clasificador de Fuzzy C-Means, se obtiene a costa de aumentar el tiempo de ejecución, de hecho le toma 55 segundos (en un computador con un procesador Pentium 4 de 1.6 Ghz) para clasificar una imagen de

320x240 píxeles, en comparación con los 9 segundos que emplea K-Means con inicialización al azar y en ambos casos con un criterio de para de $\varepsilon = 10^{-6}$.

Por otra parte los resultados de FCM son constantes, es decir cada vez que el algoritmo es ejecutado el resultado es el mismo, a diferencia de K-Means que depende fuertemente de los valores de inicialización del algoritmo y en cada ejecución el resultado es diferente como puede verse entre la **Figuras 17 b)** y 17 c) en las que para la misma imagen de entrada, las imágenes de salida del grupos Uva, son diferentes.

a) Imagen original en el espacio de color RGB

b) Clasificación por K-Means, primera ejecución, clase 2.

c) Clasificación por K-Means, segunda ejecución, clase 2

d) Clasificación por K-Means, ocho clases

e) Clasificación por Fuzzy C-Means, ocho clases.

Figura 17. Resultados obtenidos con los diversos métodos de clasificación. Encerrado en los círculos se observan zonas mal clasificadas por Fuzzy C-Means.

1.1.2.3. Desempeño de los clasificadores.

Para evaluar estos métodos se empleó como referencia el porcentaje de las áreas de los racimos clasificadas correctamente, Ecuación 37, así como el porcentaje del número de racimos detectados correctamente, Ecuación 38. En este último caso se eliminaron todos aquellos grupos menores a 10 píxeles, ya que pueden ser considerados como ruido.

$$P_A = \frac{\text{Área Real}}{\text{Área detectada}} 100$$

Ecuación 37. Porcentaje de acierto en la clasificación de área de racimos.

$$P_N = \frac{\text{Número de Racimos Real}}{\text{Número de Racimos detectadas}} 100$$

Ecuación 38. Porcentaje de racimos correctamente clasificados.

Los resultados pueden ser apreciados en la Tabla 3. Nótese que el porcentaje de racimos identificados correctamente en ambos métodos es igual, ya que una vez eliminados los grupos menores a 10 píxeles, sólo quedan los racimos y parte del tronco que han sido erróneamente clasificados.

Tabla 3: Desempeño de K-Means y FCM como clasificadores.

		K-Means	FCM
Porcentaje de área correctamente clasificada	P_A	85,1	90,0
Porcentaje de racimos identificados correctamente	P_N	87,5	87,5

Ambos clasificadores sobreestiman la cantidad de racimos, ya que el tronco tiende a ser confundido con los racimos debido a que ambos poseen píxeles de igual color. Esto puede ser subsanado al eliminar regiones que sean más anchas que largas, ya que los troncos generalmente tienen una posición horizontal, mientras que los racimos son por norma más largos que anchos.

Realizando esta modificación, el porcentaje de racimos detectados correctamente P_N sube a un 100%, pero no así el área clasificada correctamente P_A , la que sube apenas de un 85,1% a un 85,8% para K-Means y de un 90% a un 90,7% para FCM.

De ser factible, se deberían emplear cuatro bandas para una clasificación robusta, ya sea mediante imágenes visibles o de imágenes basadas en los índices de vegetación. En otras palabras, se debería emplear imágenes RGB convertidos al espacio L^*a^*b y clasificar empleando los tres canales del espacio L^*a^*b junto con el canal IR de la imágenes RGIR.

1.2. Métodos de clasificación por forma.

1.2.1. Transformada circular de Hough

Con respecto a la identificación por forma, la transformada circular de Hough no mostró resultados viables que puedan ser empleados para la identificación de bayas, debido al bajo número de aciertos (10,6%) y de falsas detecciones (51%). Como se puede apreciar en la Transformada circular de Hough sobre el canal b^* . Busca círculos de radio entre 3 y 9 píxeles, sin tolerancia a incluir círculos concéntricos. **Figura 18**, la frontera entre bayas es difusa, porque comparte el color con sus vecinas y a que la sombra que proyectan entre sí hace esa zona difusa sin bordes definidos.

Figura 18. Transformada circular de Hough sobre el canal b^* . Busca círculos de radio entre 3 y 9 píxeles, sin tolerancia a incluir círculos concéntricos.

Estos resultados no mejoran significativamente al aplicar la transformada circular de Hough sobre la imagen de la clase Uva generada por K-Means o FCM, ni al aumentar la resolución.

1.2.2. Transformada de Simetría Radial (TSR)

Pero el tamaño de las bayas que aparecen en la imagen es variable debido a su propia naturaleza y a que existe una variación en la ubicación dentro de la vid, es decir se encuentran más o menos alejadas de la cámara. Por lo tanto se deben tomar en cuenta

estas variaciones en la búsqueda de puntos radialmente simétricos e introducir un rango de posibles radios N.

Los resultados de aplicar esta técnica de detección pueden ser observados en la **Figura 19 b)**. La búsqueda se realizó sobre el canal a^* del espacio de color $L^*a^*b^*$ y además sobre las imágenes convertidas a escala de grises utilizando la expresión $I=0.2989 R + 0.5870 G + 0.1140 B$.

Figura 19. a) Imagen original en escala de grises. b) Mapa de simetría para búsqueda de simetría en dos radios (4 y 5 píxeles). c) Zonas con mayor índice de simetría. d) mapa de simetría para búsqueda de un sólo radio (5 píxeles).

Debido a su forma las bayas deberían mostrar un alto grado de simetría pero producto de lo compacto de los racimos y a lo difuso de la frontera entre bayas, no es posible identificar claramente sus bordes y por tanto detectar simetrías. Además se debe considerar que durante la adquisición de las imágenes la cámara no mantiene una distancia constante a los racimos, por lo que las bayas quedan fuera de los rangos de búsqueda aptos para operar en tiempo real. A mayor rango de búsqueda mayor tiempo de ejecución.

En cuanto a los tiempos de ejecución, en un ordenador Intel i7, estos están en torno a los 700 ms para imágenes de 240x320 pixeles, y dos radios de búsqueda. Cuando se reduce a un sólo radio de búsqueda el tiempo de ejecución cae a 450 ms, lo que lo convertiría en una técnica válida para operar en tiempo real. Pero como se ha señalado se requiere un rango amplio de radios para obtener resultados viables.

1.3. Análisis de textura.

1.3.1. Estadísticos de segundo orden para caracterizar textura.

Al someter las imágenes a los filtros de: entropía global, entropía local, de rango y de la media, sobre los espacios de color, L^*a^*b , RGB, HSV, XYZ, se generan las imágenes de textura que se observan en la **Figura 20 b)** a **Figura 20 k)**.

a)

b)

c)

d)

e)

f)

g)

h)

Figura 20. Imágenes de textura generadas. a) Imagen original. b) Espacio de color $L^*a^*b^*$, filtro entropía local. c) Espacio de color $L^*a^*b^*$, filtro de desviación estándar local. d) Espacio de color $L^*a^*b^*$, filtro del rango. e) Espacio de color RGB, filtro entropía local. f) Espacio de color RGB, filtro de desviación estándar local. g) Espacio de color RGB, filtro del rango. h) Espacio de color HSV, filtro de entropía local. i) Espacio de color HSV, filtro de desviación estándar. j) Espacio de color HSV, filtro del rango. k) Espacio de color XYZ, filtro de entropía local.

Los análisis de textura muestran que es posible separar el fondo de la vegetación, como lo muestran las figuras 20 b), c), d), f), g), j) e i). Sin embargo no representa un aporte significativo al desarrollo del sistema, ya que las características que se desea extraer no están claramente identificadas, a diferencia de lo que ocurre con la separación por color, que obtiene mejores resultados, al separar características de un modo más claro.

Por otra parte los tiempos de ejecución son de entre 1,6 y 2,2 s en un ordenador Intel i7, por lo que esta estrategia fue descartada para aplicarse en tiempo real.

1.3.2. Wavelet de Haar.

La transformada *Wavelet* mapea la imagen en una sub-imagen de baja resolución y una serie de imágenes de detalles, como se ha dicho. La imagen de baja resolución se obtiene enturbiando iterativamente la imagen, mientras que las imágenes de detalles contienen la información perdida durante dicha operación.

En la **Figura 21** se puede observar el resultado de aplicar Wavelet de Haar de tres niveles de descomposición. Mientras que en la **Figura 22** se puede observar la descomposición y energía (información) remante luego de la descomposición.

Figura 21. a) Imagen original, b) Imagen comprimida mediante Wavelet de Haar de tres niveles aplicada sobre la imagen original a). c) Píxeles identificables como uva.

Figura 22. Tres niveles de descomposición e imágenes de detalles en tres orientaciones.

Figura 23. A la izquierda, imagen de un racimo en envero espacio de color RGB. A la derecha imagen comprimida a nivel de descomposición 5, mediante Wavelt de Haar.

En las imagens de racimos en envero, el algortimo Wavelt de Haar permite separar zonas que contienen bayas maduras de la que contienen bayas verdes, como se puede apreciar en las figuras 23, 24, 25 y 26 en las que se emplearon distintas configuraciones y espacios de color. Pero tanto en el espacio de color HSV como en el $L^*a^*b^*$ se pueden apreciar, enmarcados en elipses amarillas, algunos errores de clasificación.

Figura 24. Imagen de un racimo en envero, el espacio de color RGB, con distintos niveles de compresión mediante Wavelt de Haar.

Figura 25. A la Izquierda, Imagen de un racimo en envero, el espacio de color HSV. A la derecha imagen después de la compresión mediante Wavelt de Haar.

Figura 26. A la Izquierda, Imagen de un racimo en envero, el espacio de color $L^*a^*b^*$. A la derecha imagen después de la compresión mediante Wavelet de Haar.

En la **Figura 26**, se puede apreciar el resultado de la clasificación sobre el espacio de color $L^*a^*b^*$, aunque con algunos errores de clasificación importantes, sobre todo entre tronco y racimos.

Por otra parte, aunque los resultados mostrados en la imagen de baja resolución creada mediante Wavelet (**Figura 21 b**) y **Figura 21 c**) son utilizables para una posterior identificación, estas no son aptas para procesamiento en tiempo real debido a que el algoritmo tarda 21s por imagen de 640x480 píxeles y 7 s a 320x240 píxeles, por lo que las técnicas basadas en Wavelet fueron descartadas.

1.3.3. Segmentación de texturas.

En cuanto al método propuesto por Brox & Weickert, (2004) de “segmentación de texturas sobre un espacio de características de difusión no supervisada”, en principio presenta interesantes características ya que permite identificar racimos, hojas, tronco etc., pero requiere un proceso posterior de identificación y fusión de grupos de texturas generados. Además este método requiere elevados tiempos de ejecución, del orden de 1 a 2,5 minutos por imagen de 320x240 píxeles.

Figura 27. a) Imagen original en falso color de un segmento de la Figura 20 a). b) Vector de características para escala de la textura. (Imagen segmentada). c) Vector de características para la orientación de la textura en el eje x d) Vector de características para la orientación de la textura en el eje y. e) Vector de características para la magnitud de la textura. f) Vector de características para la imagen de intensidad.

En la **Figura 27** se puede observar un ejemplo de la segmentación de texturas sobre una imagen en escala de grises. En la **Figura 27 b)** se aprecia la segmentación en grupos de igual textura. En la **Figura 27 c)** y **d)** se puede observar las características de las texturas presentes en la dirección del eje x e y respectivamente. En la **Figura 27 e)** se observa el vector de características para la magnitud de la textura expresado como imagen. En la **Figura 27 f)** se muestra el vector de características para la imagen de intensidad.

Todas estas características en su conjunto permiten medir la disimilaridad entre las texturas y luego agrupar las texturas iguales, las que posteriormente podemos asociar a la clase de interés. Sin embargo se descartó su uso por las razones citadas.

1.4. Métodos de agrupamiento supervisados.

1.4.1. Vecinos más cercanos.

Para evaluar el uso del algoritmo de los vecinos más cercanos, se desarrolló una interface de usuario como la que se muestra en la **Figura 28**. Dicha interface solicita al usuario que seleccione áreas representativas de cada una de las clases de interés y usa los píxeles contenidos en dichas regiones como referencia o grupo de entrenamiento para la posterior clasificación.

Se evaluó la distancia euclidiana y la del coseno para determinar la pertenencia uno u otro grupo; por otra parte se empleó la regla de “consenso” para determinar a qué clase pertenece cada pixel.

Los resultados de la clasificación de la imagen presentada en la **Figura 28**, mediante el método de los Vecinos más Cercanos, pueden ser observados en las **Figura 29 a)** a **h)**.

El software desarrollado, genera resultados visuales y cuantitativos, en particular muestra las imágenes de cada clase creada y muestra además los porcentajes de

píxeles pertenecientes a cada clase identificada por el usuario. Un ejemplo de resultados se presenta en la **Figura 29** y **Figura 30**.

Una vez aplicado a un conjunto de 20 imágenes con una resolución de 480x320 píxeles sobre los canales a^*b^* del espacio de color $L^*a^*b^*$ se obtuvo los resultados de la Tabla 4. Los principales errores en clasificación provienen de píxeles que podemos considerar ruido, por ejemplo lo pertenecientes al fondo, los que sistemáticamente son incluidos en la clase Uva así como en la clase tronco (véase **Figura 29** f) y g).

Tabla 4. Resultados del método de los vecinos más cercanos aplicados con dos distancias.

Distancia	Aciertos clase Uva	Aciertos clase hojas
Euclidiana	78%	92%
Coseno	85%	95%

Figura 28. Ejemplo de la interface de usuario creada para capturar áreas con píxeles representativos para cada clase. Encerrado en líneas rojas, las áreas seleccionadas por el usuario como píxeles de referencia para cada clase.

a)

b)

c)

d)

e)

f)

Figura 29. Resultados de la clasificación mediante el método los vecinos más cercanos. a) Píxeles pertenecientes a la clase Hojas Viejas. b) Píxeles pertenecientes a la clase Sarmientos. c) Píxeles pertenecientes a la clase Fondo 1. d) Píxeles pertenecientes a la clase Fondo 2. e) Píxeles pertenecientes a la clase Racimo. f) Píxeles pertenecientes a la clase Racimo 2. g) Píxeles pertenecientes a la clase Tronco. h) Píxeles pertenecientes a la clase Hojas Nuevas.

Area de Hojas Viejas	47.21%
Area de Sarmientos	2,73 %
Area de Fondo 1	3.84 %
Area de Fondo 2	14.18%
Area de Racimo 1	1.78 %
Area de Racimo 2	1.56 %
Area de Tronco	7.55 %
Area de Hojas Nuevas	21.21%

Figura 30. Impresión de pantalla del software donde se muestra el resumen de los porcentajes de cada clase generada.

1.5. Métodos de agrupamiento no supervisados.

1.5.1. Segmentación basada en Delaunay.

Al aplicar la triangulación de Delaunay sobre el histograma de la imagen mostrada en la **Figura 30 a)**, como propone Cheddad et al., (2003), se producen agrupamientos como los mostrados en la **Figura 30 b)**. Este método resulta ventajoso ya que no necesita inicializar el número de clases a generar, pero su tiempo de ejecución, de alrededor de 1,5 s, no lo hace apto para procesamiento en tiempo real.

Por otra parte al no conocerse el número de clases a generar, si la imagen presenta una gran variabilidad en el color, el método que se emplee para identificar que clase pertenece a cada categoría buscada (hojas, racimos etc.) tardará más si se compara con los métodos en donde el número de clases generadas es reducido y conocido *a priori*.

En la **Figura 31 d)** se muestra el resultado de fusionar manualmente grupos similares. Nótese en la **Figura 31 c)** encerrado en un círculo de color amarillo, se visualiza un racimo que contiene tres tonos de azules, los que para mayor claridad pueden ser identificados en la **Figura 31 d)** en colores amarillo, azul oscuro y azul claro. La fusión muestra que finalmente que el algoritmo genera importantes errores de clasificación ya que se suman píxeles de sarmientos a la clase uva, así como píxeles de tronco. En un

conjunto de 20 imágenes de 480x320, los resultados alcanzan apenas un 57 % de aciertos para la clase Uva.

Figura 31. a) Imagen Original. b) Imagen en escala de grises agrupada mediante el método de Delaunay. c) Imagen en color como la original con los 64 grupos generados por el método de Delaunay. d) Imagen falso color con grupos fusionados manualmente.

2. Segunda campaña de adquisición.

2.1. Índices multiespectrales.

En cuanto a los índices de vegetación NDVI y SAVI, el que mejores resultados entregó fue SAVI con un valor para L de 0,5, ya que permite eliminar el fondo de las imágenes y conservar sólo la vegetación. Por su parte NDVI tiende a confundir píxeles de vegetación con el fondo como se puede apreciar en la **Figura 32 c)**.

Luego empleando la imagen SAVI con valores menores a 0 es posible crear dos grupos, vegetación y no vegetación lo que resulta útil al momento de corregir errores en la clasificación de las imágenes RGB.

Figura 32. Índices espectrales umbralizados generados partir de la imagen infrarroja.

Esta estrategia evita el uso de iluminación artificial y reduce el efecto de las variaciones en la iluminación. El uso de imágenes infrarrojas y visibles de modo conjunto permite lograr mayor robustez en la clasificación, al proveer una mayor cantidad de información

a los algoritmos de agrupamiento. Los resultados señalan que este método es viable, (sobre todo si se desea remover el fondo si es que se usa una pantalla posterior) siempre que se emplee una cámara RGB-IR, ya que esto evita las correcciones que se deben hacer al emplear dos cámaras diferentes.

2.2. Métodos de clasificación de agrupamiento difusos.

Se realizó una extensa evaluación cuantitativa de métodos de agrupamiento difuso como se puede ver en Correa *et al.*, (2011c), tanto para imágenes en escala de grises como para imágenes en color. Además de las evoluciones descritas en Correa *et al.*, (2011c), se evaluó Fuzzy c-mean con Restricciones Espaciales y Kernelled Robust Fuzzy C-means. Esta última de modo cualitativo, ya que en las evaluaciones preliminares no presentó desempeños aceptables.

2.2.1. Fuzzy c-mean con Restricciones Espaciales y Contornos de Regiones.

Un ejemplo de clasificación difusa aplicando el método denominado *Fuzzy c-mean with spatial constraints*, FCMSC, cuyos resultados se pueden observar en la **Figura 33**.

Figura 33. En a) se observa la imagen original, de b) a h) corresponden a las diversas clases generadas.

El algoritmo de agrupamiento difuso es ejecutado sobre el canal b^* y configurado con 8 clases sobre una imagen de 340x240 píxeles.

Luego al aplicar la detección de regiones sobre el grupo 2 de la **Figura 33 d)** como lo proponen Li *et al.* (2011) con el método “spatial fuzzy clustering with level set methods”, se obtienen resultados como los de la **Figura 34 a)**, en la que, encerrado en magenta, aparecen los contornos de las regiones creadas por FCM y en verde aparecen las regiones creadas después de 200 iteraciones. En la **Figura 34 b)** se muestra el resultado después de 400 iteraciones, en el que se aprecia encerrado en amarillo un crecimiento excesivo de los contornos de la región, lo que implicó la inclusión de píxeles de sarmiento. En la **Figura 34 c)** en rojo racimos no detectados.

Figura 34. a) Algoritmo ejecutado con 200 iteraciones. b) Algoritmo ejecutado con 400 iteraciones. Encerrado en amarillo una zona de mal clasificada. c) Imagen con racimos no detectados encerrado en rojo y en amarillo zonas mal clasificadas.

Este método, a pesar de operar sobre imágenes en escala de grises, permite discriminar correctamente la clase Uva. Sin embargo requiere identificar manualmente la clase Uva y además requiere tiempos de ejecución de 8 s para imágenes de 320x240 píxeles adicionales a los 14 s requeridos por la clasificación difusa. Por dichos motivos fue descartado para operar en tiempo real, a pesar de que la coherencia espacial de la información permite reducir los errores de clasificación. Con frecuencia, en FCM y métodos derivados, píxeles al interior de los racimos son asignados a otra clase, error que es corregido por FCMSC.

Los resultados de aplicar este método a un pequeño conjunto de 8 imágenes se pueden apreciar en la Tabla 5.

Tabla 5. Resultados del método *spatial fuzzy clustering with level set*.

Iteraciones	Aciertos clase Uva	Aciertos clase hojas
200	91%	92%
400	89%	93%

2.2.2. Kernelled Robust Fuzzy C-means.

En cuanto Kernelled Robust Fuzzy C-means (KRFCM), este método no mostró buenos resultados en la clasificación, como se observa en la **Figura 35**, ya que tiende a generar grupos coincidentes, en este caso genera 4 de las 8 clases buscadas, razón por la cual no fue posteriormente evaluado.

Figura 35. a) Imagen original de 320x240 píxeles. b) Resultado de la clasificación mediante Kernelled Robust Fuzzy C-Means.

2.3. Evaluación cuantitativa de la Segmentación

Con el propósito de evaluar el desempeño de las técnicas de segmentación, sobre un grupo de 20 imágenes se segmentaron manualmente las áreas que contenían hojas y racimos, para luego calcular el área de cada una de ellas. Estas áreas fueron comparadas con las áreas entregadas por los cinco métodos, usando la Ecuación 39 y sus resultados pueden ser observados en la Tabla 6.

$$Exactitud = \sum_{k=1}^n \frac{Area \text{ en clase } k \text{ correctamente segmentada}}{Area \text{ total método en clase } k}, \forall k$$

Ecuación 39. Cálculo de la exactitud de la clasificación de los métodos difusos.

El espacio de color L*a*b* confirmó ser el espacio con mejor desempeño, en específico los canales de color a* y b*. Estos canales permitieron el mejor agrupamiento de píxeles, debido a que solo contienen información cromática, lo que contribuye a la eliminación de los efectos producidos por el oscurecimiento o iluminación en las imágenes. Esto se debe a que el canal L es el que contiene la información de oscurecimiento/iluminación en el espacio de color L*a*b*.

Tabla 6. Exactitud de las técnicas de agrupamiento, con y sin ruido. Desempeño para m=2.

Método de agrupamiento difuso	FCM	PCM	FPCM	RFPCM	FCM-GK
Exactitud sin ruido	90%	2%	88%	5%	91%
Exactitud con ruido	87%	3%	85%	5%	88%

Diversas pruebas fueron llevadas a cabo con el propósito de determinar el número adecuado de grupos a emplear, en particular se probó entre 6 y 20 grupos con incrementos de uno. Este parámetro es de importancia capital para el desempeño de los algoritmos de agrupamiento, ya que el número de grupos está directamente relacionado con las características de color de los objetos a identificar dentro de una escena. Si se establece un número pequeño de grupos, los algoritmos tenderán a unir en un mismo grupo píxeles de diversas clases.

Un bajo número de grupos tiende a generar grupos poco densos, es decir con valores de similaridad baja. Por otra parte un elevado número de grupos generará grupos densos pero con una mayor costo computacional, lo que lo hace inviable para procesamiento en tiempo real.

Como se ha mencionado, el número de grupos debe obedecer a las características de color de los objetos a identificar en vides, esto es: frutos en diversos estados de madurez, hojas secas, hojas senescentes, hojas jóvenes, tronco, sarmientos, fondo y eventualmente cielo.

De este modo, el número mínimo de grupos seleccionado fue ocho. Al emplear menos de ocho grupos, píxeles de tronco son agrupados dentro del grupo de racimos, incluso racimos y fondo son agrupados en los mismos grupos. Por otra parte, como se esperaba, a mayor número de grupos mejor el desempeño de los algoritmos, pero empeoran significativamente los tiempos de ejecución.

Debido a que las imágenes originales no son ruidosas, el desempeño de los algoritmos orientados a la reducción del ruido no muestran un mejor desempeño si se les compara con el algoritmo FCM estándar, como se puede apreciar en la Tabla 6.

En las **Figura 36 b) a 35 k)** se muestran diferentes segmentaciones producidas por los grupos creados al emplear dos resoluciones 640x480 píxeles y 320x240 píxeles. Los algoritmos no muestran una diferencia significativa en su desempeño, cuando la resolución es reducida a 320x240 píxeles.

El efecto sobre los grupos creados, al aplicar artificialmente un 5% de ruido del tipo “Sal y Pimienta” se puede apreciar en la **Figura 36**. Por otra parte el efecto del ruido, sobre las imágenes segmentadas con las diversas técnicas evaluadas, se puede apreciar en las **Figura 36 l) a 35 o)**.

Nótese que en la **Figura 36 i)** el algoritmo RFPCM genera cuatro grupos, los grupos faltantes son grupos coincidentes. En la **Figura 37 c)** PCM genera dos y en la **Figura 37 e)** FPCM genera seis de los ocho requeridos, también como consecuencia del efecto de los grupos coincidentes.

Figura 36. Imágenes generadas por la segmentación por agrupamiento, empleando ocho clases sobre los canales a^* , b^* del espacio de color $L^*a^*b^*$, con resoluciones de 640x480 y 320x240 píxeles.

Figura 37. Ejemplos de gráficos de los grupos y centros de grupos generados a partir del agrupamiento de píxeles sobre el plano a^*b^* usando ocho grupos.

Tabla 7. Tiempo de ejecución, en segundos, para las técnicas de agrupamiento empleadas, a diferentes resoluciones.

Método de agrupamiento difuso.	FCM	PCM	FPCM	RFPCM	FCM- GK
Tiempo a 640x480	50	56	100	5	600
Tiempo a 320x240	13	19	39	1	210

En definitiva podemos señalar que la reducción de la resolución llevada a cabo no afecta el desempeño de los algoritmos de agrupamiento, pero mejora significativamente los tiempos de ejecución (Tabla 7).

Los algoritmos FPCM y FCM resuelven el problema de los datos atípicos, pero este punto no resultó relevante en el conjunto de imágenes empleadas debido a que no

tienen ruido. El algoritmo FPCM entrega buenos resultados, pero genera dos grupos coincidentes, es decir solo entrega 6 de los 8 grupos deseados.

En cuanto a los tiempos de ejecución (Tabla 7), el mejor desempeño lo obtuvo el algoritmo RFPCM, pero a expensas de un pobre desempeño como clasificador. FCM permanece como el algoritmo más rápido, después de RFPCM, ya que los restantes algoritmos, usan como punto de partida, en el proceso iterativo, los centros de los grupos generados por FCM.

Debido a que los algoritmos RFPCM, KRFCM y PCM generan grupos coincidentes no deberían ser considerados como técnica de segmentación en viñedos. Aunque, RFPCM considera a los grupos pequeños como ruido, y entrega solo dos grupos, este algoritmo no debe ser desechado, ya que ha demostrado ser suficientemente rápido como para implementarse en forma recursiva, como clasificador binario

Finalmente, FCM-GK es el clasificador de mejor desempeño pero a costa de un mayor tiempo de ejecución, mientras que FCM muestra el mejor compromiso entre velocidad y desempeño como clasificador, lo que lo convierte en la mejor opción para ser implementado en procesos en tiempo real.

2.3.1. Errores de Clasificación del método seleccionado, FCM-GK.

Un problema que aparece cuando en la imagen existe una cantidad pequeña de uva, como la que se muestra en la **Figura 38 a)**, es que algoritmos basados en FCM tienden a incluir pixeles no deseados en el grupo Uva, como se puede apreciar en la **Figura 38 b)**, en donde se confunde Tronco y parte del suelo con Uva.

De este modo, si al inicio de la operación de captura de imágenes el tamaño de los racimos de uva es muy pequeño, entonces el grupo identificado como Uva vuelve a ser procesado mediante el algoritmo de clasificación, obteniendo resultados como los que se muestran en la **Figura 38 c)**.

La segunda clasificación se realiza con los mismos parámetros que la primera, es decir, aplicada sobre los canales a^* , b^* , con 8 grupos.

a)

b)

c)

Figura 38. a) Imagen original, con baja presencia de uva. c) Grupo Uva creado al aplicar el algoritmo de clasificación una vez. c) Grupo Uva al aplicar el clasificador sobre la imagen b).

3. Tercera campaña de adquisición.

3.1. Resultados de la estrategia de aceleración de convergencia

En esta sección se presentan los resultados en cuanto a: desempeño, clasificaciones erróneas y tiempos de ejecución del algoritmo, cuando K-means es inicializado con los centros de los grupos generados por FCM-GK. Además se muestra el desempeño de la SVM como clasificador, y el rendimiento de SIFT como algoritmo de unión.

3.1.1. Desempeño de la clasificación

La **Figura 39** muestra un ejemplo de la salida del proceso de agrupamiento representado en un gráfico en los que sus ejes son los canales de color a^* y b^* respectivamente. Los grupos de pixeles similares se representan con el mismo color y los centros o prototipos de cada grupo se representan con un marcador en negro de mayor tamaño.

Como se puede ver, los grupos generados tienen diferentes tamaños y formas y son separables de forma no lineal. Este primer resultado también fue corroborado en el resto de las imágenes procesadas.

Figura 39. Partición por agrupamiento, realizada sobre el plano a^*b^* del espacio de color $L^*a^*b^*$. Las marcas grandes en negro marcan los centros de cada grupo o prototipos de los grupos.

Los grupos se pueden representar en imágenes de falso color, en las cuales se puede observar gráficamente la diferencia en desempeño de los clasificadores. En la **Figura 40 a)** se muestra una imagen de una vid en el espacio de color RGB y la **Figura 40 b)** muestra el desempeño de los grupos generadas por FCM-GK. Como se puede apreciar, todas las características de interés de la imagen original pueden ser distinguidas claramente.

Sin embargo, cuando se muestra el rendimiento de K-means con inicialización aleatoria (**Figura 40 c)**) se puede observar que el desempeño se ve disminuido. Por el contrario, cuando se observa en la **Figura 40 d)** el rendimiento de K-means inicializado con los centroides de FCM-GK, se observa una mejora significativa.

Cuando se cuantifica esta impresión gráfica, FCM-GK obtiene un rendimiento de un 97%, elevada en comparación con el 88% de K-means con inicialización aleatoria. Por otra parte, cuando se aplica la estrategia de inicialización para K-means, el desempeño se eleva a un 95%, lo que representa una mejora significativa en comparación con la inicialización aleatoria estándar.

a)

b)

c)

d)

Figura 40. a) Imagen original. b) Grupos generados por FCM-GK. c) Imagen de grupos generado por K-means con inicialización aleatoria. d) Imagen de grupos generado por K-means mediante la estrategia de inicialización.

Para analizar en detalle los resultados de clasificación de K-means realizados con la estrategia de inicialización propuesta, se presentan en la **Figura 41 b)** a f) un conjunto de grupos, uno por imagen. La **Figura 41 a)** muestra la imagen original (en el espacio de color RGB), mientras que en la **Figura 41 b)**, c) y d) se presentan los grupos con hojas viejas, nuevas y secas, respectivamente. La **Figura 41 e)** muestra el grupo Fondo, pero muestra además una clasificación errónea: en esta imagen aparece también el tronco de la vid.

Esta clasificación errónea se debe a la acumulación de polvo sobre la pantalla blanca, utilizada durante la temporada de muestreo, para evitar confundir las hileras paralelas de fondo con la hilera de interés. Por último, en la **Figura 41 f)** se presenta el grupo Uva. Nótese que este grupo tiene pequeños errores de clasificación expresados como agujeros en el interior de las uvas y como pequeños grupos de píxeles. Estos errores de clasificación se reducen drásticamente cuando se aplica el proceso de filtrado.

(a)

(b)

(c)

d)

(e)

(f)

(g)

Figura 41. a) Imágenes originales en RGB. b) Grupo Hojas viejas. c) Grupo Hojas Nuevas. d) Grupo Hojas Viejas II. e) Grupo Hojas Secas. f) Grupo Fondo. g) Grupo Uva.

Para comparar visualmente el desempeño de FCM-GK, K-means sin la estrategia de inicialización y K-means con la estrategia de inicialización, se presenta en la **Figura 42** el grupo Uva generado a partir de estos tres clasificadores. En la **Figura 42 a)** y la **Figura 42 b)** se puede notar que los resultados son diferentes simplemente debido a la inicialización aleatoria estándar en K-means.

Por otra parte, la **Figura 42 c)** muestra el elevado rendimiento de FCM-GK y en la **Figura 42 d)** con un rendimiento comparable se muestra el desempeño de K-means con la estrategia de siembra propuesta. Nótese que los desempeños mostrados son previos al proceso de filtrado, donde se eliminan pequeños grupos de píxeles y se rellenan “hoyos”.

(a)

(b)

(c)

(d)

Figura 42. a) Grupo Uva proveniente de K-means con inicialización aleatoria. b) Grupo Uva proveniente de K-means con inicialización al azar cuando se ejecuta el algoritmo por segunda vez. c) Grupo Uva proveniente de FCM-GK. d) Grupo Uva proveniente de K-means inicializado con los centroides FCM-GK.

3.1.2. Tiempos de ejecución

Como se mencionó, los K-means con la estrategia de inicialización tienen un rendimiento similar al de FCM-GK pero con la ventaja de un tiempo de ejecución inferior. De hecho, la diferencia es de un orden de magnitud, siendo de la reducción de 7 s a 0,7 s a 480x320, 3,4 s a 0,3 s en 240x320, como se aprecia en la Tabla 8.

Tabla 8. Tiempos de ejecución a diferentes resoluciones para la estrategia de aceleración de la convergencia.

Resolución	FCM-GK	K-means	K-means con inicialización
Tiempo a 480x320	7,2 $\pm$ 0.2	2,7 $\pm$ 0,2	0,72 $\pm$ 0,02
Tiempo a 240x320	3,4 $\pm$ 0.2	1,2 $\pm$ 0,1	0,29 $\pm$ 0,02

En consecuencia, es posible utilizar bajas resoluciones manteniendo el rendimiento y reduciendo el tiempo de ejecución.

3.1.3. Resolución

No hay diferencia en el desempeño de los clasificadores cuando la resolución se reduce de 800x600 píxeles a 480x320 píxeles o incluso cuando se reduce a 240x320 píxeles (Tabla 9). Esto se debe a que, para la distancia de adquisición empleada, cada píxel representa un cuadrado de lado 4 mm a la resolución más baja, lo que indica una alta densidad de información disponible, incluso para las imágenes de menor resolución.

Tabla 9. Porcentajes de aciertos a dos resoluciones.

Resolución	FCM-GK	K-means	K-means con inicialización
480x320	97%	88%	95%
240x320	97%	88%	95%

Otra razón para afirmar que la resolución es suficiente, es que el procedimiento de validación manual tiene una precisión de 2% cuando se compara con una imagen creada sintéticamente.

3.1.4. SIFT

En la **Figura 43** se muestra un mosaico creado mediante SIFT, utilizando imágenes adquiridas cada 0,5 m con una resolución de 800x600 píxeles. Cuando se utilizan resoluciones menores a 800x600 píxeles no es posible hacer una correspondencia entre imágenes con las características detectadas. En efecto la cantidad de características, como esquinas, decrece considerablemente al reducir la resolución y las que se mantienen no son suficientes para lograr correspondencias y por ende la unión entre las dos imágenes.

a)

Figura 43. a) Mosaico creado por el algoritmo basado en SIFT. La imagen muestra la unión mediante la superposición de tres imágenes que producen un mosaico 3 m de longitud. b) Regiones de interés usadas para acelerar la velocidad de convergencia de SIFT.

Cada imagen captura un segmento de hilera de 1,5 m y cuando se fusiona con la imagen siguiente, el algoritmo crea una nueva imagen de 2 m. Diversas pruebas se llevaron a cabo con el fin de determinar el solapamiento mínimo necesario para crear un mosaico usando SIFT. Debido a que las imágenes fueron adquiridas en paralelo a la viña, hay un gran cambio en las características presentes en la imagen a lo largo de la hilera. Por lo tanto, fue necesario tomar una imagen cada 0,5 m para garantizar la eficacia de algoritmo SIFT.

Con el fin de acelerar el rendimiento de los algoritmos, las imágenes fueron divididas en tres secciones, (regiones de interés ROI, por sus siglas en inglés), de 0,5 m de ancho por 1 m de altura y numeradas de izquierda a derecha. SIFT es aplicado a la ROI 2 de la imagen 1 y a la ROI 1 de la imagen 2. Si no hay características coincidentes entre estas regiones, SIFT se aplica a la ROI 3 de la imagen 1 y a la ROI 2 de la imagen 2 como se muestra en la **Figura 43 b)**.

3.1.5. Identificación de grupos mediante redes neuronales.

Los centroides o prototipos de grupo, identificados mediante redes neuronales permitieron identificar correctamente el 100 % de los grupos Uva. Mientras que para el Grupo Hojas secas tuvieron un rendimiento promedio del 95%. Por otra parte para los grupos hojas nuevas el rendimiento fue de un 100% y para el grupo Hojas Viejas también alcanzó un 100 %.

Estos resultados son alcanzados en promedio en 30 ms por imagen, lo que los hace un resultado completamente viable para ser implementado en tiempo real.

3.1.6. Identificación de grupos mediante SVM

Las SVM han mostrado un rendimiento de detección del 100% para el grupo Uva y en menos de 2 ms. Sin embargo, debido al hecho de que las SVMs son clasificadores binarios, fue necesario crear otra SVM específicamente dedicada a la identificación de las hojas, obteniéndose finalmente el mismo rendimiento que en el grupo Uva. Los bajos tiempos de cálculo requerido se deben a que la SVM solo debe clasificar 8 pares de datos (centroides) por imagen.

De esta manera SVM ha demostrado ser un clasificador de alto rendimiento que permite la separación de espacios de entrada no linealmente separables.

La diferencia entre ambas técnicas evaluadas es el tiempo de ejecución y la extensión del conjunto de imágenes empleadas para el entrenamiento de dichos métodos. En el caso de las redes neuronales artificiales el conjunto de imágenes debe ser más amplio si se compara con el requerido por SVM y si nos referimos a los tiempos de ejecución, aun siendo ambos viables para ser utilizados en tiempo real, SVM muestra una velocidad mayor de ejecución.

3.2. Cámara de profundidad

Dentro de los principales resultados de esta investigación podemos señalar que los métodos que mejor se adaptan a la caracterización del viñedo en tiempo real, son

aquellos basados en las características cromáticas del viñedo. Esto es válido tanto en el espectro visible como en el infrarrojo, ya que, como se muestra en Correa *et al.* (2011a), las imágenes en el espacio de color L^*a^*b y las imágenes en IR permiten separar e identificar hojas, tronco, sarmientos, racimos etc.

Para la adquisición de imágenes visibles y de profundidad se desarrolló un software *ad hoc* que puede observarse en la **Figura 44**.

Figura 44. Captura de pantalla del software de adquisición de imágenes desarrollado para la cámara Kinect. En rojo objetos más cercanos.

En la **Figura 45** se muestra el resultado de filtrar la imagen usando los datos de profundidad y conservando solo los píxeles que se encuentren a una distancia menor a 2 m y mayor a 0,6 m. Nótese que desaparecen las hojas pertenecientes a las hileras posteriores.

Figura 45. Imagen usando como mascara los objetos más cercanos, marcados en rojo en la Figura 44 derecha.

En la **Figura 46**, se muestran los resultados de aplicar la estrategia de aceleración, es decir, procesar inicialmente con FCM-GK, y usar los centroides generados como semillas para K-Means.

En la Figura 46 a) se muestra el grupo identificado como tronco, en las que el color blanco indica regiones filtradas por distancia, en negro zonas no pertenecientes al grupo. En la **Figura 48 a)** se muestran además algunos errores de clasificación, en los que se confunde tronco con algunos de los postes metálicos instalados como soporte del viñedo.

En las imágenes en la **Figura 46 b), c), d), y f)** se muestran grupos identificados como hojas verdes, mientras que en la Figura 46 h) muestra un grupo de hojas secas.

Por otra parte en la **Figura 46 g)** se muestra la clase identificada como Uva. Dicha imagen es producto de la agrupación por color, en la que claramente hay errores de clasificación. En efecto, se confunde parte de la manguera perteneciente al sistema de riego por goteo y parte del poste metálico de sujeción del viñedo con los racimos de uva.

Esto se debe a que los postes metálicos tienen un recubrimiento galvánico de color gris, que fácilmente se confunde con el gris del recubrimiento ceroso de las bayas. Asimismo, la manguera de plástico, originalmente de color negro, sufre un proceso de degradación producto de su exposición al sol y al polvo lo que la deja de un color gris similar al de la cubierta cerosa de las bayas.

Figura 46. Grupos resultantes del proceso de agrupamiento con K-menass inicializado con FCM-GK.

En la **Figura 47** se muestra el resultado del proceso de fusión, que genera el grupo hojas verdes, a partir del cual se calcula el área de hojas.

Figura 47. Grupo Hojas resultantes del proceso de fusión de grupos realizado por SVM.

En la **Figura 48** se muestra el proceso de filtrado mediante operaciones morfológicas y mediante filtrado por electricidad y ángulo de inclinación del elipsoide que contiene los objetos en este grupo.

Figura 48. a) Grupo identificado como Uva. b) imagen binaria del grupo Uva antes de las operaciones morfológicas y de filtrado. c) Imagen binaria después de las operaciones morfológicas y del proceso de filtrado (excentricidad entre 0.70-0.98). d) Grupo Uva enmascarado con la imagen binaria filtrada.

Como se puede observar después del proceso de filtrado es posible separar exitosamente los racimos del resto de los componentes de la imagen.

3.3. Discusión comparada de resultados en clasificación e identificación

Otros investigadores, como Nuske et al. (2011a) han obtenido desempeños de clasificación de uva cercanos al 98%, pero con un 25% de falsos negativos y con elevados tiempos de ejecución. Berenstein et al. (2010) lograron desempeños cercanos al 90% en la identificación de uvas, pero no abordan los problemas asociados a la superposición de imágenes. Braun et al. (2010) enfrenta la cuestión de la superposición de imágenes, pero su investigación está dirigida únicamente a evaluar la ubicación de las hojas. Finalmente, Dey et al., (2012) identifica tronco, hojas y uvas simultáneamente y tiene un rendimiento cercano al 98%, pero con elevados tiempos de ejecución.

Por ejemplo en Fernández *et al.* (2013), que utiliza una metodología similar a la aquí propuesta, basada en aplicar K-means sobre los canales a^* y b^* del espacio de color $L^*a^*b^*$, alcanza un 89,7% de uvas correctamente clasificadas, valor similar al 90% y 91% obtenidos en Correa *et al.* (2011a) y Correa *et al.* (2011b) respectivamente.

Si se compara esta investigación con otras basadas en clasificación por color podemos observar diferencias en los resultados obtenidos con diferentes metodologías. Una publicación realizada por Liu et al. (2013) compara su propuesta de clasificación de píxeles mediante histogramas 3D, con la propuesta de umbralización RGB de Reis *et al.* (2012) y con la propuesta de FCM-GK/SVM de esta tesis (Correa *et al.*, 2012), aunque solo emplean un conjunto de 16 imágenes.

Los resultados muestran que la propuesta de Liu *et al.* tiene una tasa de aciertos (verdaderos positivos) de un 95% con un 2% de falsos positivos, la umbralización RGB alcanza una tasa de verdaderos positivos de un 87%, con un 5% de falsos positivos, mientras que FCM-GK/SVM alcanza una tasa de verdaderos positivos de un 97%, pero con un 16% de falsos positivos.

Figura 49. Resultados de clasificación para diversos métodos basados en color publicados por Liu et al. (2013). a) Imagen original. b) Resultados histograma 3D. c) resultados umbralización RGB. d) resultados FCM-GK/SVM. e) Resultados grupo uvas usando FCM sobre canales a^*b^* e identificado mediante SVM. f) Resultados FCM/SVM como imagen de los grupos resultantes. g) Grupo tronco y sarmientos generado por FCM/SVM. h) Grupo hojas generado por FCM e identificado mediante SVM.

Los resultados obtenidos de FCM-GK/SVM según Liu *et al.*, tienen un desempeño significativamente menor al encontrado en esta investigación. Estos resultados pueden ser explicados por el hecho de que en la comparativa no se usó el espacio de color L*a*b, una de las bases del algoritmo propuesto en esta tesis.

De hecho al aplicar FCM-GK/SVM sobre la imagen publicada por Liu *et al.*, se obtiene un 96,5 % de verdaderos positivos y sólo un 1,5 % de falsos positivos. Estos resultados pueden ser observados en la **Figuras 49 e)** y **49 f)** en las que queda patente la escasa cantidad de falsos positivos a diferencia de los que se ha informado en la comparativa presentada por Liu *et al.*

Además si se observa la **Figura 49 g)** no hay confusión entre el grupo uvas y el grupo tronco y sarmientos, como aparece en los resultados de la comparativa en la **Figura 49 d)**. Por otra parte la técnica FCM-GK/SVM permite simultáneamente generar otros grupos de interés como las hojas (**Figura 49 h)**) o tronco y sarmiento (**Figura 49 g)**), lo que como se ha señalado ofrece importantes ventajas.

Aunque se ha demostrado la factibilidad de clasificar e identificar imágenes de viñas, mediante métodos de clasificación no supervisados como FCM-GK en tiempo real, con porcentajes de aciertos superiores al de otras investigaciones basadas en procesamiento de imágenes, no es menos cierto que las condiciones en las que las imágenes son adquiridas deben ser, en cierto modo, controladas para conseguir estos niveles de aciertos (superiores al 90%), ya que la iluminación juega un papel preponderante en cualquier técnica basada en imágenes.

Véase por ejemplo la **Figura 50 a)**, en la que difícilmente se pueden identificar los racimos ocultos entre las sombras, que tienen un color azul muy oscuro o por otra parte resulta difícil identificar algunos racimos como tales, ya que aparecen de un color celeste brillante, por efecto de la pruina y la luz incidente.

Figura 50. a) Imagen original en condiciones normales de iluminación. b) Grupo identificado como uvas antes de eliminar pequeñas regiones. c) Imagen de Fernández et al. (2013) en condiciones de terreno. d) Imagen de Reis et al. (2012).

En esta situación el clasificador debe incluir en una misma clase píxeles de un azul oscuro intenso junto con píxeles celeste brillante, tarea imposible si no se usa un espacio de color y canales adecuados. En la **Figura 50 b)** se puede apreciar el resultado de clasificar una imagen, con los extremos de iluminación antes señalados, empleando la metodología propuesta en esta tesis.

Nótese que la metodología propuesta es capaz de identificar correctamente uvas oscuras y claras poniéndolas a ambas en la misma clase. El porcentaje de aciertos para esta imagen es del 86,3%, con una tasa de 2,5% de falsos positivos, que principalmente corresponden a hojas con tonalidades azules inducidas por la aplicación de sulfato de cobre como fungicida.

Por otra parte, algunos autores, como Fernández *et al.* (2013), señalan que sus algoritmos han sido validado bajo condiciones no controladas de iluminación (**Figura 50 c)**) en viñedos bajo prácticas de manejo convencionales. Sin embargo como lo muestra la **Figura 50 c)** los racimos están completamente expuestos y la iluminación corresponde a un día nublado, lo que reduce al mínimo el efecto de la luz sobre la imagen, además de no presentar oclusiones por hojas, lo que por ende facilita la labor del clasificador.

Aun así, si se compara los resultados de ésta investigación, llevada a cabo sin control de la iluminación y en condiciones de manejo normal (sin desoje), con los de otras donde se hayan usado condiciones controladas de iluminación, el porcentaje de aciertos en clasificación es superior. Por ejemplo Reis *et al.* (2012) alcanzan un 97 % de aciertos en uvas rojas, adquiriendo las imágenes durante la noche y fotografiando directamente el racimo como se puede ver en la **Figura 50 d)**. Debe notarse que Liu *et al.* (2013) informan que la técnica propuesta por Reis *et al.* (2012) alcanza un 87% de aciertos y un 5 % de falsos positivos.

En el trabajo realizado por Nuske *et al.* (2011b) se alcanza un 75 % de aciertos, en uvas blancas, empleando un flash de alta intensidad. Por otra parte Grocholsky *et al.* (2011),

Nuske *et al.* (2014) emplean la técnica propuesta por Nuske *et al.* (2011b) con imágenes nocturnas iluminadas con una lámpara alógena.

Otro ejemplo lo encontramos en Casser, (2016) quien también emplea imágenes nocturnas, las que clasifica empleando SVM sobre el espacio de color $L^*a^*b^*$ con resultados de 99 % pero solo aplicadas a un conjunto de 12 imágenes. Este pequeño conjunto no permite garantizar capacidad de generalización. Kicherer *et al.* (2015) recurre a la misma estrategia de captura de imágenes nocturnas, lo que da cuenta acerca de la importancia del problema de la iluminación.

Debe notarse que las estrategias de procesamiento en paralelo y de inicialización inteligente y adaptiva propuestas en Correa *et al.* (2012a) resultan esenciales a la hora de proporcionar clasificación en tiempo real. Si a lo anterior se le suma el mínimo entrenamiento requerido por las SVM, hace que la propuesta en su conjunto sea viable en condiciones de terreno ya que se auto calibra, se adapta a las cambiantes condiciones del viñedo e incluso de la iluminación, asigna etiquetas de modo autónomo, calcula las aéreas y finalmente geo-referencia la información generada.

Por otro lado se debe señalar que ha quedado de manifiesto la habilidad de las cámaras de profundidad para separar la hilera de interés de la hilera paralela (ver **Figura 45**), lo que resuelve el problema en cuestión, a diferencia de lo que ocurre en las investigaciones en la que no se utiliza una pantalla o cubierta detrás la hilera a analizar. Esta estrategia se ve viable como lo muestran los trabajos de Nguyen *et al.*, (2016) y de Vitzrabin & Edan (2016) que emplean cámaras RGB-D (visible + profundidad) para ayudar al proceso de clasificación, aunque en dichos caso la estrategia está destinada a detectar manzanas y pimientos, respectivamente.

Por ejemplo si se analiza la imagen de la **Figura 49 a**), podemos notar que resulta casi imposible separar la hilera de interés de las paralelas a ésta. Sin embargo, al utilizar la información de profundidad, ésta queda claramente delimitada (**Figura 51 d**)), lo que reduce los errores de clasificación, al evitar confundir hojas de hileras paralelas con la de interés y al evitar confundir malezas con hojas, debido a su similitud en el color.

Figura 51. a) Imagen original. b) imagen de profundidad. En rojo se observa la hilera paralela. c) imagen de profundidad con valores menores a 1,7 m. d) imagen original enmascarada usando la información de profundidad (distancia menor 1,7 m).

Por otra parte, resulta atractiva la idea de combinar información de ubicación espacial del pixel, para en conjunto con la información de color realizar una clasificación mediante alguno de los métodos de agrupamiento probados, como se puede observar en el apartado 3.1 de los resultados. Esto debido a que la coherencia espacial y tamaño de los grupos de píxeles obtenidos ayudan sin lugar a dudas, como ya se ha probado, a determinar si ciertos grupos de píxeles pertenecen a la clase uvas o troncos, clases normalmente confundidas por los algoritmos debido a su similitud en color. Esto, ya que la pruina que recubre las uvas maduras, sumadas al polvo y al efecto de estar cubiertas o parcialmente cubiertas por sombras, hace que estas dos clases que *a priori* parecen diferentes sean a veces confundidas por los clasificadores. En efecto en el trabajo de Roscher *et al.* (2015) se llega a la misma conclusión acerca de la utilidad de la

información 3D en la identificación de las estructuras del viñedo así como en la eliminación de la hilera paralela.

3.4. Resultados predicción de rendimiento y área foliar

En la siguiente sección se describirá la relación existente entre área expuesta de uva y el rendimiento a cosecha, así como la relación entre área de hojas expuestas y su relación con el área total de hojas contenidas en una vid.

3.4.1. Validación del algoritmo

En la **Figura 52** se muestra un ejemplo de ROI (30x30 píxeles) seleccionada para el proceso de validación manual del algoritmo

Figura 52. Validación manual de la clasificación. a) Imagen original, donde 73,1% son píxeles de uva. b) Imagen de los grupos creados. Píxeles azules corresponden a la clase Uva, y representan un 72% de la imagen. c) Imagen original, donde el 58% de los píxeles son de hojas. d) Imagen de los grupos creados. Los píxeles negros corresponden a hojas, y representan un 63% de la imagen.

La validación manual presenta un 98% de clasificación correcta para la clase Uva y un 92% para hojas (suma de los grupos Hojas jóvenes y Hojas viejas). La mayor parte de

los errores de clasificación en los grupos de hojas se debieron a sarmientos jóvenes y brotes laterales más pequeños, ya que exhiben casi el mismo color verde de las hojas. Como se muestra en la **Figura 53** y mediante el proceso de validación manual, los clasificadores se desarrollaron correctamente sin aplicar ningún pre-tratamiento a las imágenes, como podría haber sido mejora de contraste, brillo o ajuste de color. Este es un resultado importante, ya que hace más simple el proceso, si se compara con trabajos anteriores en el que las imágenes debían ser cortadas, reorientadas o fuertemente pre-procesadas (Herrero *et al.*, 2010), y especialmente interesante, dado el hecho de que no se utilizó iluminación artificial para la adquisición de imágenes en el viñedo.

a)

b)

Figura 53. a) Detalle de la imagen original. b) Detalle de las uvas identificadas antes de las operaciones morfológicas.

3.4.2. Desempeño del clasificador

Con el fin de ilustrar el desempeño del clasificador, se seleccionaron tres imágenes que contenían errores representativos de clasificaciones incorrectas, para luego detallar todas las clases generadas. Estas imágenes corresponden a tres etapas de defoliación diferentes y se presentan en las Figuras 54, 55, 56.

Del mismo modo, en la **Figura 54**, se eligió una imagen de vid no deshojada, no raleada en la Etapa 0 ilustrando pixeles de uva erróneamente clasificados. A continuación, se seleccionaron dos imágenes que muestran otros errores comunes de clasificación, tales como los producidos por coloración atípica de hojas (**Figura 55**) o por la influencia del cielo (**Figura 56**).

En la **Figura 54** se generaron siete clases: Uva, Tronco (sarmientos y tronco), Fondo y cuatro tipos diferentes de hojas, que incluye Hojas jóvenes de grado 1, Hojas jóvenes de grado 2, Hojas viejas grado 1 y Hojas viejas de grado 2, siendo la numeración de los grados equivale a aumentar etapas de madurez hojas.

La diferenciación de varios tipos de hojas (**Figura 54** i), j), k), l)) tiene importantes implicaciones en el desempeño de la evaluación de la fisiología de la vid, debido a que las hojas más jóvenes son fotosintéticamente más activas que las hojas más viejas (Pallioti *et al.*, 2000) y revelan diferencias en el comportamiento fisiológico de la planta, que pueden alertar de la presencia de estrés abiótico tales como el producido por deficiencias de nutrientes o de agua.

Figura 54. Ejemplo de clasificación realizada sobre una imagen de una vid, no desojada (imagen I0) y no raleada. (a) Imagen original. (b) Imagen de grupos. (c) Clase Tronco. (d) Clase Fondo. (e) Clase Uva, sin operaciones morfológicas. (f) Clase Uva después de las operaciones morfológicas. (G) área ampliada en la parte (e) que muestra pequeños grupos de píxeles mal clasificado. (h) área ampliada en la imagen (f) que muestra cómo las operaciones morfológicas eliminan los pequeños grupos de píxeles. (I) Clase Hojas viejas grado 1. (j) Clase Hojas viejas grado 2. (k) Clase Hojas jóvenes grado 1. (l) Clase Hojas jóvenes grado 2.

La clase Uva tuvo errores de estimación en una primera etapa (**Figura 54 e**)) debido a dos efectos combinados. En primer lugar, las regiones de uva también incluyen pequeños píxeles brillantes blancos asignados erróneamente a la clase de Fondo (**Figura 54 d**)), píxeles que en realidad corresponden a la capa de cera blanca visible en la piel de las uvas, y conocida como “bloom” o pruina, lo que lleva a una subestimación de los racimos de la clase Uva. Por otra parte, los píxeles azulados en la parte superior de la imagen corresponden a algunas áreas de hojas (**Figura 54 e**) y detalle en la **Figura 54 g**)) píxeles que fueron mal clasificados en la clase de Uva, lo que aporta significativamente a su posterior sobreestimación. Estos errores se redujeron drásticamente cuando estos pequeños grupos de píxeles fueron eliminados mediante un conjunto de operaciones morfológicas. Específicamente, se realizaron operaciones consecutivas de erosión y dilatación sobre la clase de Uva para eliminar estos píxeles como se muestra en la **Figura 54 f**) (detallado en (h)).

Cuando se realizó la primera etapa de defoliación (imagen I1, **Figura 55**), más racimos fueron expuestos, las condiciones de iluminación sobre las uvas mejoraron, y en consecuencia, su tasa de detección aumentó (**Figura 55** (a), (b)).

En la **Figura 54 b**), la clase de Uva aparece en amarillo, la clase tronco en color azul oscuro, y las cuatro clases de hojas (Hojas jóvenes de grado 1, Hojas jóvenes de grado 2, Hojas Viejas grado 1, Hojas Viejas grado 2), se presentan en azul, azul claro, cian y rojo, respectivamente. Cuando se compara con las cuatro clases de hojas mostradas en la **Figura 54**, las hojas en la **Figura 54** (clases e), f), g) y h)) son similares, mostrando que el algoritmo de clasificación es robusto en la segregación de varios estados de madurez foliares, frente a condiciones de iluminación cambiantes, mientras se llevó a cabo la defoliación.

Figura 55. Ejemplo de clasificación, realizado sobre una imagen de la vid después de la primera etapa de la defoliación (I1) para mostrar la capacidad del algoritmo para separar entre diferentes tipos de hojas (a) imagen original. (b) Imagen de grupos. (c) la clase de Uva, sin operaciones morfológicas. (d) Clase Tronco. (e) Clase Hojas viejas de grado 1. (f) Clase Hojas viejas grado 2. (g) Clase Hojas jóvenes grado 1. (h) Clase Hojas jóvenes grado 2.

En esta imagen (I1), la clase Uva también mostró un evento de error en la clasificación. Lo mismo sucede en la esquina superior izquierda, encerrado en un círculo rojo (**Figura 55 b), c)**) donde unas hojas con una coloración azul inducida por la aplicación de un fungicida (sulfato de cobre) fueron confundidas y consideradas como píxeles de uva.

En este caso, la clasificación errónea no fue causada por un mal desempeño de la metodología de clasificación, sino por la similitud de tono entre las hojas pulverizadas y la clase de Uva. En este escenario, y para mejorar el rendimiento global del algoritmo, los píxeles son clasificados como píxeles de uva, sólo si se encuentran en la mitad inferior de la imagen.

También en el caso de imágenes con baja variabilidad de tipos de hojas y condiciones adversas de luz solar, el algoritmo mostró una respuesta adecuada y capacidad de adaptación (**Figura 56**).

Figura 56. Ejemplo de clasificación realizada sobre una imagen de vid después de la tercera etapa de defoliación (imagen I5). a) Imagen original. b) Imagen de grupos. c) Clase de Uva. d) Fondo brillante. e) Fondo oscuro. f) Clase Tronco. g) Clase Hojas jóvenes. h) Clase Hojas viejas.

Después de la etapa de defoliación final, sólo los sarmientos y racimos restantes quedan visibles en el cubierta (imagen I5) (**Figura 56 a**)). Bajo estas condiciones, y debido a que el fondo no pudo cubrir toda el área, el cielo puede ser parcialmente distinguible, lo que hace que la luz del sol penetre en la escena y cree sombras y zonas brillantes sobre el fondo. Para superar esta situación, se seleccionaron dos clases diferentes de fondo, y se las identificó como clases Fondo brillante y Fondo Oscuro (**Figura 56 d**), e)), y las cuatro clases de hojas descritas e identificadas en los pasos anteriores fueron reducidas a sólo dos grupos: Hojas viejas y Hojas jóvenes (**Figura 56 g**), h)). El desempeño de la clasificación se muestra en la **Figura 56 b**).

En la **Figura 56 f)**, se muestra la clase Tronco que incluye tronco, sarmientos y alambres del enrejado. Por otra parte, la **Figura 56 c)** muestra la clase de Uva, incluyendo los errores preliminares de clasificación, encerrados en rojo en la parte superior de la imagen, producidos por algunos píxeles de coloración azul correspondientes al cielo. Esta clasificación errónea también fue resuelta al considerar como píxeles de uva solamente los situados en la mitad inferior de la imagen.

3.4.3. Estimación de rendimiento a cosecha y área foliar

Las curvas de correlación y validación para la estimación del área foliar y rendimiento de uva utilizando la metodología de clasificación y análisis de imágenes propuesta, se muestran en las Figuras 57 y 58, respectivamente.

Se encontró una fuerte correlación entre el área foliar real en las imágenes vides y el número de píxeles correspondientes a la clase Hojas. Dicha correlación es lineal ($y = 0,1712x + 0,1863$) con un coeficiente de determinación $R^2 = 0,78$ $p < 0,001$ (**Figura 57 a)**). Cuando se utiliza esta función para predecir el área de la hojas de otro conjunto de imágenes de vid (conjunto de validación), la correlación entre las superficies de hojas observadas y previstas está muy cerca de la línea 1:1 ($y = 1.0598x + 0,0117$) y los valores de $R^2 = 0,81$ $p < 0,001$ y $RMSE = 0,745 \text{ m}^2$ (**Figura 57 b)**).

Figura 57. a) Correlación entre el área foliar real ($\text{m}^2 \text{vid}^{-1}$) de la cubierta en las imágenes del grupo de entrenamiento, y el número de píxeles correspondientes a la clases Hoja calculado por el método de clasificación y análisis de imágenes. b) Comparación entre los valores reales de área foliar ($\text{m}^2 \text{vid}^{-1}$) provenientes de las imágenes del grupo de validación, y los valores estimados de área de hojas calculados con la ecuación de correlación en (a). La línea punteada es la línea 1:1.

Las diferencias en la densidad foliar de las vides fotografiadas, interpretadas como mayor o menor número de capas de hojas, pueden haber afectado la estimación de área de hojas en la vid. De este modo, en cubiertas muy densas (con varias capas superpuestas de hojas) los pasos iniciales de defoliación (Imágenes I1 e I3) podrían no haber causado una "desaparición" significativa de hojas en la imagen como se esperaba, de modo que la vid se mantuvo completamente cubierta de hojas, y sólo hasta que se alcanzó la etapa 3 de defoliación, el área de la cubierta observada se redujo drásticamente. En otras palabras, para cubiertas muy densas, la eliminación de hojas en las primeras etapas no siempre significó bajas estimaciones de área foliar.

Las imágenes utilizadas en el presente trabajo se correspondían con vides de cubiertas medias a muy densas, en general, lo que parece ser el escenario menos favorable para la estimación de área de hojas mediante análisis de imágenes. Sin embargo, la predicción del área de hojas, a través del modelo basado en análisis de imágenes fue muy satisfactorio, y se prevé tenga un mejor desempeño en vides con cubiertas menos densas, como en el caso de viñedos de vigor bajo a moderado.

Por otra parte, una estimación fiable y precisa del área de hojas de la vid en diversos momentos durante la temporada de crecimiento puede ser de gran utilidad para los viticultores de cara a controlar, de una manera dinámica, el crecimiento vegetativo de la planta, e identificar los síntomas de varios tipos de estrés abiótico y biótico, tales como estrés hídrico y presión de enfermedades. Del mismo modo, esta información también puede ayudar al viticultor en la toma de decisiones sobre el manejo de la cubierta de modo de mejorar el equilibrio entre el crecimiento vegetativo y reproductivo.

En cuanto a la estimación del rendimiento de uva, la correlación entre el rendimiento real y el número de píxeles correspondientes a la clase de uva siguió una relación lineal ($y = 0.1787x + 0,611$) con un coeficiente de determinación $R^2 = 0,78$ $p < 0,001$ (**Figura 58 a**)).

Figura 58. a) Correlación entre el rendimiento de las vides (kg) obtenidos a partir de las imágenes del grupo de entrenamiento, y el número de píxeles correspondientes a la clase de uva, calculada por el método de clasificación y análisis de imágenes. b) Comparación entre los valores reales de rendimiento (kg) de las vides de las imágenes del conjunto de validación, y los valores de rendimiento estimados, calculados con la ecuación de correlación en a). La línea punteada representa la línea 1:1.

Cuando se emplea esta curva para predecir el rendimiento de las imágenes del grupo de validación, la correlación entre los valores de rendimiento observados y pronosticados también están cerca de la línea 1:1 ($y = 0.8907x + 0,253$), con un $R^2 = 0,73$ en $p = 0,002$, y $RMSE = 0,749$ kg (**Figura 58 b**)).

En follajes densos y muy densos, los racimos de uva están típicamente cubiertos con hojas, evitando su exposición a los rayos solares, pero también reducen su grado de visibilidad frente al ojo humano o una máquina. Este hecho, que se produjo en las imágenes iniciales de las vides no deshojadas, no raleadas (I0) y también en las imágenes correspondientes a la primera etapa de la defoliación (I2), parece haber afectado el desempeño de la predicción del rendimiento mediante la metodología de análisis de imágenes, ya que el coeficiente de determinación de los valores de la estimación de rendimiento fueron más pequeñas que los de la predicción de área foliar.

Al igual que para el área foliar, la predicción del rendimiento de uva, a partir del modelo establecido mediante análisis de imágenes, fue satisfactoria, y cubre una amplia gama de condiciones de exposición y visibilidad de uva, generada por sucesivas defoliaciones

y etapas de extracción de uva. En este sentido, la defoliación basal es una práctica de manejo de la cubierta, ampliamente realizado en todo el mundo –entre cuajado y envero– en uno o dos lados de la cubierta. Esta práctica está dirigida a mejorar la exposición de la fruta para mejorar la calidad de la uva (Crippen y Morrison, 1986; Reynolds et al., 1986; Percival et al., 1994; Reynolds et al., 1996; Bergqvist et al., 2001; Kliewer y Dokoozlian, 2005) y su salud (Bledsoe et al., 1988). Dado que la visibilidad de los racimos es ciertamente incrementada después de la defoliación basal, la exactitud de la estimación del rendimiento –mediante la metodología de clasificación y análisis de imágenes presentada en este trabajo – aumentaría significativamente, por lo tanto, permitiría una predicción de rendimiento muy precisa.

Una estimación veraz del rendimiento potencial de la uva cercano a envero provee una información muy valiosa no sólo para fines logísticos en la cosecha (e.g., necesidades de mano de obra, capacidad de elaboración del vino en la bodega...), sino también por razones económicas, especialmente cuando un productor de vino tiene que comprar uvas de otros viticultores y proveedores. De igual forma la producción total de uva de una región determinada o zona de denominación es una importante fuerza motriz del precio final de la uva en una cosecha determinada.

VIII. CONCLUSIONES

1. Conclusiones Predicción de rendimiento

La metodología, desarrollada para extraer características de la cubierta y análisis de imágenes, descrita en el presente trabajo, ha demostrado ser una herramienta útil y fiable para la evaluación del área foliar y del rendimiento en la viña. La configuración propuesta para la adquisición de imágenes es simple, de bajo costo y no destructiva ya que sólo se necesita una cámara RGB comercial. La metodología de procesamiento ha demostrado ser altamente adaptable y robusta frente a los cambios en iluminación y distancia a la vid objetivo, dos factores que resultan críticos en aplicaciones de visión artificial en condiciones de campo.

La metodología de clasificación permite discriminar siete diferentes clases, correspondientes a siete tipos de características de la cubierta presentes en imágenes de vides, aunque sólo las clases Hojas y Uva fueron calibradas y validadas con éxito contra las mediciones reales de la planta. El rendimiento del clasificador para la identificación de las hojas y las uvas fue muy alto y su efectividad sobrepasa el 90% en ambos casos.

Una estimación precisa del área de la hojas de la vid y del rendimiento durante la temporada de cultivo, hecha a través de un método rápido y no destructivo, como el que se describe en este trabajo, puede proporcionar información muy valiosa para la industria de la uva y del vino en la toma de decisiones de manejo del follaje, así como para fines logísticos y económicos.

2. Conclusiones Métodos de agrupamiento

En una primera etapa los resultados mostrados por Fuzzy C-Means resultaron prometedores, pero dicha técnica no puede ser aplicada en tiempo real (menos de 60 ms) por la limitación del tiempo de procesado, por lo que la técnica de K-Means resulta más adecuada para dicho objetivo.

Al evaluar otros clasificadores difusos, como PCM.FPCM, RFPCM, FCMSC, KRFCM, o FCM-GK se concluye que FCM-GK tiene el mejor desempeño como clasificador, pero con costo computacional elevado. La norma inducida con la que mide las distancias permite adaptarse localmente a los datos, generando grupos internamente homogéneos.

Ocho grupos es el mínimo número de clases que se debe emplear para separar adecuadamente uvas y hojas del resto de los componentes de la escena.

3. Conclusiones Estrategia de Aceleración

Se ha presentado un nuevo marco metodológico dedicado a caracterizar viñedos en tiempo real, lo que permite la cuantificación e identificación de características de la vid tales como presencia de uvas, tronco, fondo oscuro, fondo brillante y cuatro clases de hojas de distintas edades o grados de vigor.

Se logró una mejora significativa en los tiempos de ejecución y también en los rendimientos de K-means al inicializarlo con los centros de los grupos generados por FCM-GK. Esta estrategia permite la adquisición y procesamiento de imágenes en tiempo real a $13,5 \text{ km h}^{-1}$, velocidad suficiente para la mayoría de las tareas agrícolas. Por otra parte, se alcanzó una reducción en el sesgo de clasificación a través de la actualización, cada cuatro imágenes, de los valores de inicialización de K-means. Esta estrategia aporta la capacidad de adaptarse a las nuevas condiciones que puedan aparecer en terreno.

SVM permite la identificación del 100% de los grupos de hojas y uvas en menos de 3 ms. Es una mejora significativa en comparación con la ANN que necesita 30 ms para la identificación. Por lo tanto SVM es muy adecuado para esta tarea, y además el entrenamiento del algoritmo puede ser realizado en el terreno, ya que dicho proceso sólo toma 5 s.

En comparación con los resultados de otras investigaciones, este marco metodológico mejora el estado de la técnica, al realizar de varias tareas al mismo tiempo y la mayoría de ellas en tiempo real, con elevados índices de acierto.

Por último, futuros esfuerzos se concentraran en la aceleración del proceso de unión de imágenes, para de este modo, llevar a cabo todas las tareas en tiempo real. También se explorará la incorporación de un sistema de estabilización de la cámara, para evitar vibraciones y deformaciones de la imagen.

4. Conclusiones Finales

Se mostró que es posible desarrollar un sistema autónomo de caracterización de la cubierta vegetal en viñedos. Dicho sistema está formado por una plataforma móvil autónoma y un sistema de adquisición de imágenes visibles (RGB), de profundidad (D) e infrarrojas IR instalado.

- El sistema a través de información tridimensional del cultivo ha mostrado que permite estimar el área foliar en tiempo real, lo que permite, a partir de una rápida calibración, determinar el área fotosintéticamente activa de la vid.
- El sistema ha mostrado además que permite determinar el estado de vigor de vides en tiempo real, caracterizando las hojas entre distintos grados de vigor, desde verde intenso hasta amarillo o tonos marrones en caso de hojas secas.
- El sistema ha mostrado además que es posible estimar el rendimiento en terreno, en base al área de uva visible, es decir, sin considerar las uvas ocultas bajo las hojas.

IX. TRABAJOS FUTUROS

Se visualizan dos grande áreas de desarrollo a futuro.

La primera tiene relación con la construcción y evaluación de la plataforma robotizada, que no sólo serviría para la caracterización del viñedo, sino además para la aplicación inteligente de pesticidas.

La segunda área de desarrollo es la implementación de los algoritmos propuestos, en una aplicación (App) para teléfonos móviles, la que permitiría generar rápidamente y en terreno una aproximación de la cantidad de uva presente. Esto aprovechando las capacidades de procesamiento, cámaras y sensores de orientación de los que disponen los *Smartphone* actuales.

X. REFERENCIAS

- Al-Otum, H.M., Morphological operators for color image processing based on Mahalanobis distance measure. *Opt. Eng.* 2003, 42, 2595-2606.
- Amatya S., Karkee M., Gongal A., Zhang Q., Whiting M.D., Detection of cherry tree branches with full foliage in planar architecture for automated sweet-cherry harvesting, *Biosystems Engineering*, Volume 146, June 2016, Pages 3-15, ISSN 1537-5110, <http://dx.doi.org/10.1016/j.biosystemseng.2015.10.003>.
- Amorim R.C., Boris Mirkin. 2011. Minkowski metric, feature weighting and anomalous cluster initializing in K-Means clustering, *Pattern Recognition*, 45(3)(2012), pp 1061-1075.
- Arnó, J.; Martínez-Casasnovas, J.A.; Ribes-Dasi, M.; Rosell, J.R. Review. Precision viticulture. Research topics, challenges and opportunities in site-specific vineyard management. *Span. J. Agric. Res.* 2009, 7, 779–790.
- Arnó, J., Escolà, A., Vallès, J., Llorens, J., Sanz, R., Masip, J., Palacín, J., and Rosell-Polo, J. (2013). Leaf area index estimation in vineyards using a ground-based LiDAR scanner. *Precision Agriculture* 14, 290-306.
- Azzari, G., Goulden, M., and Rusu, R. (2013). Rapid Characterization of Vegetation Structure with a Microsoft Kinect Sensor. *Sensors* 13, 2384-2398.
- Babuka R., P. van der Veen, and U. Kaymak, "Improved covariance estimation for Gustafson-Kessel clustering," in *Fuzzy Systems*, 2002. FUZZ-IEEE'02. Proceedings of the 2002 IEEE International Conference on, vol. 2, 2002, pp. 1081 -1085.
- Berenstein R., O. Shahar, A. Shapiro, and Y. Edan, "Grape clusters and foliage detection algorithms for autonomous selective vineyard sprayer, "in *Intelligent Service Robotics*, september 2010, pp. 1-11.
- Bergqvist, J.; Dokoozlian, N.; Ebisuda, N., Sunlight exposure and temperature effects on berry growth and composition of cabernet sauvignon and grenache in the central San Joaquin valley of California. *Am. J. Enol. Vitic.* 2001, 52, 1-7.

- Berkhin P., "Survey of clustering data mining techniques". Technical Report, 2002.
[Online]. Available:
<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.18.3739>
- Bertamini, M.; Tardaguila, J.; Iacono, F., Valutazione dell'equilibrio vegeto-produttivo e microclimatico del vigneto per l'ottimizzazione delle tecniche colturali a verde: Aspetti teorici e pratici. Rivista dell'Istituto Agrario San Michele all'Adige (Bolletino ISMA) 1994, 2, 24-40.
- Bjurström, H.; Svensson, J. Assessment of grapevine vigour using image processing. Master, Linköping University, Sweden, 2002.
- Bledsoe, A.M.; Kliewer, W.M.; Marois, J.J., Effects of timing and severity of leaf removal on yield and fruit composition of sauvignon blanc grapevines. Am. J. Enol. Vitic. 1988, 39, 49-54.
- Blom, P.E.; Tarara, J.M., Trellis tension monitoring improves yield estimation in vineyards. HortScience 2009, 44, 678-685.
- Braun, T.; Koch, H.; Strub, O.; Zolynski, G.; Berns, K., Improving pesticide spray application in vineyards by automated analysis of the foliage distribution pattern in the leaf wall. In 1st Commercial Vehicle Technology Symposium Kaiserslautern, Germany, 2010.
- Bréda N.J.J. Ground-based measurements of leaf area index: a review of methods, instruments and current controversies Journal of Experimental Botany, 54 (2003), pp. 2403-2417.
- Brown, D. L., Giles, D. K., Oliver, M. N., and Klassen, P. (2008). Targeted spray technology to reduce pesticide in runoff from dormant orchards. Crop Protection 27, 545-552.
- Brox Thomas and J. Weickert. A TV Flow Based Local Scale Measure for Texture Discrimination, ECCV, 2004.
- Burgos, S., et al. "Use of Very High-Resolution Airborne Images to Analyse 3d Canopy Architecture of a Vineyard." The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 40.3 (2015): 399.
- Busin L., J. Shi, N. Vandenbroucke, and L. Macaire. Color space selection for color image segmentation by spectral clustering. In Signal and Image Processing

- Applications (ICSIPA), 2009 IEEE International Conference on, pages 262-267, 2009.
- Casser, V. Using Feedforward Neural Networks for Color Based Grape Detection in Field Images. 3rd Computer Science Conference for University of Bonn Students (CSCUBS).pp. 23-33, 2016
- Carbonneau, A., La surface foliaire exposee potentielle. Guide pour sa mesure. Prog. Agric. Vitic 1995, 112, 204-212.
- Chamelat R., E. Rosso, A. Choksuriwong, C. Rosenberger, H. Laurent, and P. Bro. Use of Zernike moments for grape detection with image processing. In Conference of the IEEE Industrial Electronics Society (IECON), pages 3697 -3702, 2006.
- Cheddad A., D. Mohamad and A. Abd Manaf, "Exploiting Voronoi diagram properties in face segmentation and features extraction," Pattern Recognition, 41 (12)(2008)3842-3859, Elsevier Science.
- Cheng, H. D.; Jiang, X. H.; Sun, Y. & Wang, J. L. 2001. Color image segmentation: Advances and prospects Pattern Recognition, 32(12)(2001), pp. 2259-2281.
- Chiang, C., Chiang, P., Fei, J.-C. & Liu, J. (2007). A comparative study of implementing Fast Marching Method and A* SEARCH for mobile robot path planning in grid environment: Effect of map resolution Advanced Robotics and Its Social Impacts, 2007. ARSO 2007. IEEE Workshop on, (pp. 1-6).
- Chikamasa, T. (2012). Simulink Support for Kinect. pp. is a set of Simulink blocks which enables to make an interface between Simulink and a natural interaction device such as Microsoft Kinect. MathWorks, Natick, Massachusetts, Estados Unidos.
- Chuang K., H. Tzeng, S. Chen, J. Wu, T. Chen. 2006. Fuzzy c-means clustering with spatial information for image segmentation, Computerized Medical Imaging and Graphics, pp 9-15.
- Chuvieco, E., Martín, M.P. y Palacios, A. 2002."Assessment of different spectral indices in the red-near-infrared spectral domain for burned land discrimination" Int. J. of Remote Sensing, vol. 23, pp. 5103-5110.
- Clingeffer, P.R. Final report for project csh 96/1. Crop development, crop estimation and crop control to secure quality and production of major wine grape varieties: A national approach; CSIRO and NRE: Australia, 2001; p 159.

- Corbalan, M.; Millan, M.; Gilabert, D. Computer vision system for color discrimination. A: "Proceedings of 4th Iberoamerican Meeting on Optics and 7th Latin American Meeting on Optics, Lasers, and their Applications". Washington: The International Society for Photo-Optical Instrumentation Engineers (SPIE), 2001, pp. 592-595.
- Correa C., A. Moya, E. Baguena, A. Herrero, J. Diago, M. and Baluja, J. Tardaguilla, C. Valero, and P. Barreiro. "Feature extraction of the vineyard canopies, using images acquired on-the-go (RGB and RGIR)/ extracción de características de la cubierta vegetal del viñedo mediante imágenes RGB y RGIR obtenidas de forma dinámica," in VI Congreso Ibérico de AgroIngeniería, 2011.
- Correa C., C. Valero, P. Barreiro J. Tardaguilla and M. P. Diago. 2012. Feature Extraction on Vineyard by Gustafson Kessel FCM and K-means. 2011. 16th IEEE Mediterranean Electrotechnical Conference, Medina-Yasmine Hammamet, Tunisia, 25-28 Mar, (pp. 481- 484).
- Correa C., C. Valero, P. Barreiro, and M. P. D. Javier Tardaguilla, "A comparison of fuzzy clustering algorithms applied to feature extraction on vineyard," in Avances en inteligencia artificial, J. Lozano, J. Gomez, and J. Moreno, Eds., vol. 1, no. 1, November 2011.
- Correa, C.; Valero Ubierna, C.; Barreiro Elorza, P., Characterization of vineyard's canopy through fuzzy clustering and svm over color images. In International Conference of Agricultural Engineering, Valencia, Spain, 2012.
- Correa C., C Valero, P Barreiro, J Ortiz-Cañavate, J Gil. (2013). Usando Kinect como Sensor para Pulverización Inteligente. Congreso Ibérico de Agroingeniería. 26 -29 Agosto. Madrid, España.
- Correa C., M. Garrido, A. Moya, C. Valero Comparación de desempeño de una cámara de profundidad comercial y un. LIDAR en explotaciones agrícolas (2013). Congreso Ibérico de Agroingeniería, 26 -29 Agosto. Madrid, España.
- Crippen, D.D.; Morrison, J.C., The effects of sun exposure on the compositional development of cabernet sauvignon berries. Am. J. Enol. Vitic. 1986, 37, 235-242.

- Crippen, D.D.; Morrison, J.C. The effects of sun exposure on the phenolic content of cabernet sauvignon berries during development. *Am. J. Enol. Vitic* 1986, 37, 243–247.
- Daughtry, C. S. (1990). Direct measurements of canopy structure. *Remote Sensing Reviews*, 5(1), 45-60.
- Dey, D.; Mummert, L.; Sukthankar, R., 2012. Classification of plant structures from uncalibrated image sequences. In *IEEE Workshop on Applications of Computer Vision (WACV)*, Breckenridge, USA, 9-11 Jan (pp. 329 -336).
- Diago, M.-P., Correa, C., Millán, B., Barreiro, P., Valero, C., and Tardaguila, J. (2012). Grapevine Yield and Leaf Area Estimation Using Supervised Classification Methodology on RGB Images Taken under Field Conditions. *Sensors* 12, 16988-17006.
- Duan, K., Keerthi, S. S. Which is the best multiclass SVM method? An empirical study. 2005 *Proceedings of the Sixth International Workshop on Multiple Classifier Systems*, Seaside USA, 13-15 Jun, (pp. 278-285).
- Dunn, G.M.; Martin, S.R., Yield prediction from digital image analysis: A technique with potential for vineyard assessments prior to harvest. *Aus. J. Grape Wine R.* 2004, 10, 196-198.
- Dunn, G.M.; S.R., M., The current status of crop forecasting in the Australian wine industry. In *ASVO Seminar Series: Grapegrowing at the Edge*, Tanunda, Barossa Valley, South Australia, 2003; pp 4-8.
- Dunstone, R.J. Final report for project dnr 02/02. Winegrape crop forecasting module; CWRDC: Australia, 2002; p 39.
- Fairlie, K.; Whitty, M.; Leach, M.; Norzahari, F.; White, A.; Cossell, S.; Guivant, J.; Katupitiya, J. Spatially Smart Wine-Testing Geospatial Technologies for Sustainable Wine Production. *Coordinates* 2011, VII, 14-16.
- Fathy M., Mohsen Sharifi and M. Tayefeh Mahmoudi, "An Experimental Comparison of Commonly used Edge Detectors", The 2002 International Conference on Imaging Science, Systems, and Technology (CISST'02), The 2002 International Multi-conference on Computer Science, CSREA Press, Las Vegas, USA, June 24-27, 2002.

- Fernández L. A.F., M. Oliveira. 2008. Real-time line detection through an improved Hough transform voting scheme. *Pattern Recognition*, Volume 41, Issue 1, Pages 299-314.
- Fernández, R.; Montes, H.; Salinas, C.; Sarria, J.; Armada, M. Combination of RGB and Multispectral Imagery for Discrimination of Cabernet Sauvignon Grapevine Elements. *Sensors* 2013, 13, 7838-7859.
- Fleck, S.; Van Der Zande, D.; Schmidt, M.; Coppin, P., Reconstruction of tree structure from laser-scans and their use to predict physiological properties and processes in canopies. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* M. Thies, B.K., H. Spiecker, H. Weinacker Ed. ISPRS: Freiburg, Germany, 2004; Vol. 36, pp 119-123.
- Ganesan P. and V. Rajini, "A method to segment color images based on modified fuzzy-possibilistic-c-means clustering algorithm," in *Recent Advances in Space Technology Services and Climate Change (RSTSCC)*, 2010, nov. 2010, pp. 157-163.
- Garrido, M.; Méndez, V.; Valero, C.; Correa, C.; Torre, A.; Barreiro, P., 2012. Online dose optimization applied on tree volume through a laser device. In *First International Conference on Robotics and Associated High-Technologies and Equipment for Agriculture*, Pisa, Italy, 19-21 Sept. pp. 325-330.
- Geng, C., Zhang, K., Zhang, E., Zhang, J., & Li, W. (2012). Assessment on spraying effect of intelligent spraying robot by experiment. *Transactions of the Chinese Society of Agricultural Engineering*, 28(SUPPL 2), 114-117.
- Gil, E.; Arnó, J.; Llorens, J.; Sanz, R.; Llop, J.; Rosell-Polo, J.R.; Gallart, M.; Escolà, A. Advanced Technologies for the Improvement of Spray Application Techniques in Spanish Viticulture: An Overview. *Sensors* 2014, 14, 691-708.
- Giles D.K., D. Downey. Quality control verification and mapping for chemical application Precision Agriculture, 4 (2003), pp. 103–124.
- González D. P., "Algoritmos de agrupamiento basados en densidad y validación de clusters," Ph.D. dissertation, Universitat Jaume I, 2010.

- Goutouly, J., R. Drissi, D. Forget, and J. Gaudillère (2006) Characterization of vine vigour by ground based NDVI measurements. Proceedings of the VI International Terroir Congress; 3-5 July 2006; Bordeaux, France. pp. 237-241
- Graves D., W. Pedrycz. 2007. Fuzzy c-means, Gustafson-Kessel FCM, and kernel-based FCM: A comparative study. *Advances in Soft Computing*, 41(2007), pp. 140-149.
- Gray, J.D.; Gibson, R.J.; Coombe, B.G.; Giles, L.C.; Hancock, T.W., Assessment of winegrape value in the vineyard-a preliminary, commercial survey. *Aus. NZ. Wine Ins. J.* 1994, 9, 253-261.
- Grocholsky, B.; Nuske, S.; Aasted, M.; Achar, S., Bates, T. 2011. A Camera and Laser System for Automatic Vine Balance Assessment. 2011. ASABE Annual International Meeting, Louisville, USA. 7-10. Aug (pp. 1-10).
- Guijarro M., G. Pajares, I. Riomoros, P.J. Herrera, X.P. Burgos-Artizzu, and A. Ribeiro. Automatic segmentation of relevant textures in agricultural images. *Computers and Electronics in Agriculture*, 75(1):75 -83, 2011.
- Hagebeuker, Dipl-Ing Bianca, and Product Marketing. "A 3d time of flight camera for object detection." (2007).
- Heath, M.; Sarkar, S.; Sanocki, T. & Bowyer, K. Comparison of Edge Detectors: A Methodology and Initial Study *Computer Vision and Image Understanding*, 1998, 69, 38 - 54.
- Herrero-Langreo, A.; Barreiro, P.; Diago, M.P.; Baluja, J.; Ochagavia, H.; Tardaguila, J., Pixel classification through Mahalanobis distance for identification of grapevine canopy elements on RGB images. In *The International Association for Spectral Imaging Conference*, Dublin, Ireland, 2010; pp 40-41.
- Igawa, H.; Tanaka, T.; Kaneko, S.; Tada, T.; Suzuki, S., "Visual and tactual recognition of trunk of grape for weeding robot in vineyards," *Industrial Electronics*, 2009. IECON '09. 35th Annual Conference of IEEE , vol., no., pp.4274,4279, 3-5 Nov. 2009.
- Kartikeyan,B and Sarkar,A, "An identification approach for 2-D autoregressive models in describing textures" *Graphical Models and Image Processing*, vol.53, pp.121-131, 1991.

- Khoshelham, K., and Elberink, S. O. (2012). Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications. *Sensors* 12, 1437-1454.
- Kicherer A, Herzog K, Pflanz M, et al. An Automated Field Phenotyping Pipeline for Application in Grapevine Research. Passaro VMN, ed. *Sensors* (Basel, Switzerland). 2015; 15(3):4823-4836.
- Killing, J., B. W. Surgenor, and C. K. Mechefske. "A machine vision system for the detection of missing fasteners on steel stampings." *The International Journal Of Advanced Manufacturing Technology* 41.7-8 (2009): 808-819.
- Kliewer, W.M., Effect of day temperature and light intensity on coloration of vitis vinifera L. grapes. *J. Am. Soc. Hort. Sci.* 1970, 95, 693-697.
- Kliewer, W.M.; Dokoozlian, N.K., Leaf area/crop weight ratios of grapevines: Influence on fruit composition and wine quality. *Am. J. Enol. Vitic.* 2005, 56, 170-181.
- Kliewer, W.M.; Lider, L.A., Influence of cluster exposure to sun on composition of Thompson seedless fruit. *Am. J. Enol. Vitic.* 1968, 19, 175-&.
- Kotsiantis S.B.. Supervised machine learning: A review of classification techniques. *Informatics*, 31(3):249-268, 2007.
- Krishnapuram R. and J.M. Keller. A possibilistic approach to clustering. *Fuzzy Systems, IEEE Transactions on*, 1(2):98-110, may 1993.
- Leemans V. and M.-F. Destain. Application of the Hough transform for seed row localization using machine vision. *Biosystems Engineering*, 94(3):325-336, 2006.
- Lei F. Tian & D. Slaughter. Environmentally adaptive segmentation algorithm for outdoor image segmentation. *Computers and Electronics in Agriculture*, 21(3):153-168, 1998.
- Li B.N., C.K. Chui, S. Chang, S.H. Ong (2011) Integrating spatial fuzzy clustering with level set methods for automated medical image segmentation. *Computers in Biology and Medicine* 41(1) 1-10.
- Li, Y.& Y.Shen, "Robust Image Segmentation Algorithm Using Fuzzy Clustering Based on Kernel-Induced Distance Measure," *Computer Science and Software*

- Engineering, 2008 International Conference on*, vol.1, no., pp.1065,1068, 12-14 Dec. 2008. doi: 10.1109/CSSE.2008.694
- Tan H., P Li, S. Lee, H Hsu, PEEA 2011 Study on citrus fruit image data separability by segmentation methods, *Procedia Engineering*, Volume 23, 2011, Pages 408-416, ISSN 1877-7058, <http://dx.doi.org/10.1016/j.proeng.2011.11.2522>.
- Liming Xu, Zhao Yanchao, Automated strawberry grading system based on image processing, *Computers and Electronics in Agriculture*, Volume 71, Supplement 1, April 2010, Pages S32-S39, ISSN 0168-1699.
- Liu, S., Marden, S., & Whitty, M. Towards Automated Yield Estimation in Viticulture. *Proceedings of Australasian Conference on Robotics and Automation*, 2-4 Dec 2013, University of New South Wales, Sydney Australia.
- Livingston, M. A., Sebastian, J., Zhuming, A., and Decker, J. W. (2012). Performance measurements for the Microsoft Kinect skeleton. In "Virtual Reality Short Papers and Posters (VRW), 2012 IEEE", pp. 119-120.
- Liyang Zheng, Daming Shi, and Jingtao Zhang. Segmentation of green vegetation of crop canopy images based on mean shift and Fisher linear discriminant. *Pattern Recognition Letters*, 31(9):920 -925, 2010.
- Loy, G. Zelinsky, A. Fast radial symmetry for detecting points of interest. *IEEE PAMI*, Vol. 25, No. 8, August 2003. pp 959-973.
- Lowe D., "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, 60, 2 (2004), pp. 91-110.
- Mabrouk, H.; Sinoquet, H., Indices of light microclimate and canopy structure of grapevines determined by 3d digitising and image analysis, and their relationship to grape quality. *Aus. J. Grape Wine R.* 1998, 4, 2-13..
- Mark Girolami. Mercer kernel-based clustering in feature space. *IEEE Transactions on Neural Networks*, 13:780-784, 2002.
- Martin, S.; Dunstone, R.; Dunn, G. How to forecast wine grape deliveries using grape forecaster excel workbook version 7; GWRDC: Australia, 2003; p 100.

- Mathews, A.J.; Jensen, J.L.R. Visualizing and Quantifying Vineyard Canopy LAI Using an Unmanned Aerial Vehicle (UAV) Collected High Density Structure from Motion Point Cloud. *Remote Sens.* 2013, 5, 2164-2183.
- McLachlan, G.J., Mahalanobis distance. *Resonance, journal of science education* 1999, 4, 20-26.
- Merzlyak M.N., A.A. Gitelson, O.B. Chivkunova, V.Y. Rakitin Nondestructive optical detection of leaf senescence and fruit ripening *Physiol Plant*, 106 (1999), pp. 135–141.
- Meunkaewjinda, A.; Kumsawat, P.; Attakitmongcol, K.; Srikaew, A. 2008. Grape leaf disease detection from color imagery using hybrid intelligent system. *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, ECTI-CON 2008, 5th International Conference on*. Krabi, Thailand, 14-17 May (pp. 513-516).
- Meyer G., Joao Camargo Neto, David D. Jones, and Timothy W. Hindman. Intensified fuzzy clusters for classifying plant, soil, and residue regions of interest from color images. *Computers and Electronics in Agriculture*, 42(3):161 -180,2004.
- Miin-Shen Yang and Hsu-Shen Tsai. A gaussian kernel-based fuzzy c-means algorithm with a spatial bias correction. *Pattern Recognition Letters*, 29(12):1713 -1725, 2008.
- Ming Gao; Tien-Fu Lu; 2006. "Image Processing and Analysis for Autonomous Grapevine Pruning," *Mechatronics and Automation, Proceedings of the 2006 IEEE International Conference*, vol., no., pp.922-927.
- Moreda G.P., J. Ortiz-Cañavate, F.J. García-Ramos, M. Ruiz-Altisent, Non-destructive technologies for fruit and vegetable size determination – A review, *Journal of Food Engineering*, Volume 92, Issue 2, May 2009, Pages 119-136, ISSN 0260-8774.
- Murillo-Bracamontes E., Martinez-Rosas M., Miranda-Velasco M., Martinez-Reyes H., Martinez-Sandoval J., Cervantes-de-Avila H., Implementation of Hough transform for fruit image segmentation, *Procedia Engineering*, Volume 35, 2012, Pages 230-239, ISSN 1877-7058

- Muthukannan K. and M.M. Moses. Color image segmentation using k-means clustering and optimal fuzzy c-means clustering. In *Communication and Computational Intelligence, 2010 International Conference on*, pages 229-234, dec. 2010.
- Nguyen T.T., K. Vandevoorde, N. Wouters, E. Kayacan, J.G. De Baerdemaeker, W. Saeys, Detection of red and bicoloured apples on tree with an RGB-D camera, *Biosystems Engineering*, Volume 146, June 2016, Pages 33-44, ISSN 1537-5110, <http://dx.doi.org/10.1016/j.biosystemseng.2016.01.007>.
- Nuske S., S. Achar, T. Bates, S. Narasimhan, and S. Singh, "Yield estimation in vineyards by visual grape detection," in *Intelligent Robots and Systems (IROS)*, 2011 IEEE/RSJ International Conference on, sept. 2011, pp. 2352 -2358.
- Nuske, S., Achar, S., Gupta, K., Narasimhan, S. G., & Singh, S. (2011). Visual Yield Estimation in Vineyards: Experiments with Different Varietals and Calibration Procedures. Tech. Report CMU-RI-TR-11-39. Carnegie Mellon University.
- Nuske, S., Wilshusen, K., Achar, S., Yoder, L., Narasimhan, S., & Singh, S. (2014). Automated visual yield estimation in vineyards. *Journal of Field Robotics*, 31(5), 837-860.
- Ohta Yu-Ichi, Tokeo Kamode.1980. Color information for region segmentation. [J]. *Computer Graphics and Image Processing* 13.Pages 222-241.
- Omasa K., F. Hosoi, and A. Konishi. 3D lidar imaging for detecting and understanding plant responses and canopy structure. *Journal of Experimental Botany*, 58(4):881-898, 2007. 330.
- Omid M., M. Khojastehnazhand, A. Tabatabaeefar, Estimating volume and mass of citrus fruits by image processing technique, *Journal of Food Engineering*, Volume 100, Issue 2, September 2010, Pages 315-321, ISSN 0260-8774, <http://dx.doi.org/10.1016/j.jfoodeng.2010.04.015>.
- Omran, M. G., Engelbrecht, A. P., & Salman, A. (2005, September). Differential evolution methods for unsupervised image classification. In *Evolutionary Computation*, 2005. The 2005 IEEE Congress on (Vol. 2, pp. 966-973). IEEE.
- Padmavathi G.and Muthukumar. Image segmentation using fuzzy c means clustering method with thresholding for underwater images. *Int. J. Advanced Networking and Applications*, 02:514-518, 2010.

- Pal N.R., K. Pal, and J.C. Bezdek. A mixed c-means clustering model. In *Fuzzy Systems, 1997., Proceedings of the Sixth IEEE International Conference on*, volume 1, pages 11-21 vol.1, jul 1997.
- Palliotti, A.; Cartechini, A.; Ferranti, F., Morpho-anatomical and physiological characteristics of primary and lateral shoot leaves of cabernet franc and trebbiano toscano grapevines under two irradiance regimes. *Am. J. Enol. Vitic.* 2000, 51, 122-130.
- Pandey S. K. & S. Hema. A Simple, Cost-Effective Method for Leaf Area Estimation, *Journal of Botany*, vol. 2011, Article ID 658240, 6 pages, 2011. doi:10.1155/2011/658240
- Percival, D.C.; Fisher, K.H.; Sullivan, J.A., Use of fruit zone leaf removal with vitis vinifera I. Cv. Riesling grapevines. II. Effect on fruit composition, yield, and occurrence of bunch rot (*botrytis cinerea* pers.:Fr.). *Am. J. Enol. Vitic.* 1994, 45, 133-140.
- Proffitt, T. (2015). Do unmanned aerial vehicles (drones) offer new opportunities to viticulture?. *Australian and New Zealand Grapegrower and Winemaker*, (615), 38.
- Rabatel, G. & Guizard, C. 2007. Grape berry calibration by computer vision using elliptical model fitting, *ECPA 2007, 6th European Conference on Precision Agriculture*, Skiathos, Greece, 3-6 Jun, (pp. 581-587).
- Ray S. and R H Turi: 1999. Determination of number of clusters in K-means clustering and application in color image segmentation, *Proceedings of ICAPRDT'99*, pp 137-143.
- Reis M.J.C.S., R. Morais, E. Peres, C. Pereira, O. Contente, S. Soares, A. Valente, J. Baptista, P.J.S.G. Ferreira, J. Bulas Cruz, Automatic detection of bunches of grapes in natural environment from color images, *Journal of Applied Logic*, Volume 10, Issue 4, December 2012, Pages 285-290, ISSN 1570-8683
- Reynolds, A.G.; Pool, R.M.; Mattick, L.R., Influence of cluster exposure on fruit composition and wine quality of seyval blanc grapes. *Vitis* 1986, 25, 85-95.
- Robledo, A.; Cossell, S. & Guivant, J. (2011). Outdoor Ride: Data Fusion of a 3D Kinect Camera installed in a Bicycle. *Proceedings of Australasian Conference on Robotics and Automation*. Monash University, Melbourne Australia.

- Roscher R., K. Herzog, A. Kunkel, A. Kicherer, R. Töpfer, W. Förstner, Automated image analysis framework for high-throughput determination of grapevine berry sizes using conditional random fields, *Computers and Electronics in Agriculture*, Volume 100, January 2014, Pages 148-158, ISSN 0168-1699
- Rodríguez-Pérez, J.R.; Riaño, D.; Carlisle, E.; Ustin, S.; Smart, D.R. Evaluation of hyperspectral reflectance indexes to detect grapevine water status in vineyards. *Am. J. Enol. Vitic.* 2007, 58, 302-317.
- Rosell J., Ricardo Sanz, Jordi Llorens, Jaume Arnó, Alexandre Escolà, Manel Ribes-Dasi, Joan Masip, Ferran Camp, Felip Gràcia, Francesc Solanelles, Tomàs Pallejà, Luis Val, Santiago Planas, Emilio Gil, Jordi Palacín. 2009. A tractor-mounted scanning LIDAR for the non-destructive measurement of vegetative volume and surface area of tree-row plantations: A comparison with conventional destructive measurements, *Biosystems Engineering*, 102(2)(2009), pp 128-134.
- Sáiz-Rubio, V., and F. Rovira-Más. 2012. Dynamic segmentation to estimate vine vigor from ground images. *Spanish Journal of Agricultural Research*, 10 (3): 596-604.
- Shafarenko L., H. Petrou, and J. Kittler. Histogram-based segmentation in a perceptually uniform color space. *Image Processing, IEEE Transactions on*, 7(9):1354{1358, sep 1998.
- Skrabanek, P., & Majerík, F. (2016). Simplified Version of White Wine Grape Berries Detector Based on SVM and HOG Features. In *Artificial Intelligence Perspectives in Intelligent Systems* (pp. 35-45). Springer International Publishing.
- Smart, R.; Robinson, M., *Sunlight into the wine. A handbook for winegrape canopy management*. Winetitles: Adelaide, 1991; p 88.
- Smart, R.E., Principles of grapevine canopy microclimate manipulation with implications for yield and quality. A review. *Am. J. Enol. Vitic.* 1985, 36, 230-239.
- Son, J.; Inoue, N.; Yamashtia, Y., Geometrically local isotropic independence and numerical analysis of the Mahalanobis metric in vector space. *Pattern Recog. Lett.* 2010, 31, 709-716.

- Storchi, P., Perria, R., Sarri, D., Rimediotti, M. and Vieri, M. (2013). Comparative Assessment of Different Sensing Technologies for Mapping The Vineyard. *Acta Hortic.* 978, 71-76
- Tardaguila, J.; de Toda, F.M.; Poni, S.; Diago, M.P., Impact of early leaf removal on yield and fruit and wine composition of vitis vinifera l. Graciano and carignan. *Am. J. Enol. Vitic.* 2010, 61, 372-381.
- Tardaguila, J.; Herrero-Langreo, A.; Barreiro, P.; C, V.; S, P.; Diago, M.P., Using rgb image analysis to assess the impact of early defoliation on the fruit zone. In 17th International Symposium GiESCO (Group of International Experts of Vitivinicultural Systems for CoOperation), Novello, V., Ed. Asti-Alba, Italy, 2011.
- Tardaguila, J.; Martines de Toda, F., Assessment of tempranillo grapes quality in the vineyard by vitur score-sheet. *J. Int. Sci. Vigne Vin* 2008, 42, 59-65.
- Tian L., David C. Slaughter. 1998. Environmentally adaptive segmentation algorithm for outdoor image segmentation, *Computers and Electronics in Agriculture*, Volume 21, Pages 153-168.
- Tian Y., Li T., Li C., Piao Z., Sun G., Wang B., 2007. Method for recognition of grape disease based on support vector machine. *Transactions of the CSAE*, 23(6)(2007) pp. 175-180.
- Tokeo Kamode Yu-Ichi Ohta. Color information for region segmentation. *Computer Graphics and Image Processing*, 13:222-24, 1980.
- Vibhute, Anup, and S. K. Bodhe. "Color Image Processing Approach for Nitrogen Estimation of Vineyard." *International Journal of Agricultural Science and Research (IJASR)* 3.3 (2013): 189-195.
- Vitrabin E., Y. Edan, Adaptive thresholding with fusion using a RGBD sensor for red sweet-pepper detection, *Biosystems Engineering*, Volume 146, June 2016, Pages 45-56, ISSN 1537-5110, <http://dx.doi.org/10.1016/j.biosystemseng.2015.12.002>.
- von Barga K. Mortensen D. A. Woebbecke D. M., Meyer G. E. Color indices for weed identification under various soil, residue, and lighting conditions. *Transactions of the ASAE*, 38:259-269, 1995.

- Walklate, P. J., Cross, J. V., Richardson, G. M., Murray, R. A., and Baker, D. E. (2002). IT-Information Technology and the Human Interface: Comparison of Different Spray Volume Deposition Models Using LIDAR Measurements of Apple Orchards. *Biosystems Engineering* 82, 253-267.
- Wang, H.; Li, G.; Ma, Z. & Li, X. 2012. Image recognition of plant diseases based on principal component analysis and neural networks *Natural. Computation (ICNC)*, 2012 Eighth International Conference on, Chongqing, China, 29-31 May (pp 246 -251).
- Watanabe, Y., Kondo, N., Ogawa, Y., Iida, M., Shiigi, T., Nishiwaki, K., Omori, H., Kurosaki, H. Machine vision system for detecting tomato fruits (2010) *American Society of Agricultural and Biological Engineers Annual International Meeting* 2010, 7, pp. 5714-5724.
- Williams L.E., J.E. Ayars, Grapevine water use and the crop coefficient are linear functions of the shaded area measured beneath the canopy, *Agricultural and Forest Meteorology*, Volume 132, Issues 3-4, 3 October 2005, Pages 201-211, ISSN 0168-1923.
- Wolpert, J.A.; Vilas, E.P., Estimating vineyard yields: Introduction to a simple, two-step method. *Am. J. Enol. Vitic.* 1992, 43, 384-388.
- Xiao-Hong Wu and Jian-Jiang Zhou. An improved possibilistic C-means algorithm based on kernel methods. In *SSPR/SPR*, pages 783-791, 2006.
- Yan Y., Y. Shen, and S. Li, "Unsupervised color-texture image segmentation based on a new clustering method," in *New Trends in Information and Service Science*, 2009. NISS '09. International Conference on, June 30 2009-july 2 2009, pp. 784 -787.
- Zaman Q.-U. and A. W. Schumann. 2005. Performance of an Ultrasonic Tree Volume Measurement System in Commercial Citrus Groves. *Precision Agriculture*, 6(5)(2005) pp. 467-480.
- Zhang, D. Q., & Chen, S. C. (2003, June). Kernel-based fuzzy and possibilistic c-means clustering. In *Proceedings of the International Conference Artificial Neural Network* (pp. 122-125).

Zheng L, D. Shi, and J. Zhang. Segmentation of green vegetation of crop canopy images based on mean shift and Fisher linear discriminant. *Pattern Recognition Letters*, 31(9):920 -925, 2010.

Zhou Yong, Li Yue'e, and Xia Shixiong. Robust fuzzy-possibilistic c-means algorithm. In *Intelligent Information Technology Application*, 2008. IITA '08. Second International Symposium on, volume 1, pages 669-673, dec. 2008.