

Segmentación de Imágenes en Viñedos para la Medición Autónoma de Variables Vitícolas.

D. S. Pérez, F. Bromberg

Laboratorio de Inteligencia Artificial DHARMA
Departamento de Sistemas de la Información, UTN Regional Mendoza
Rodríguez 273, CP 5500, Mendoza, Argentina.
{sebastian.perez,fbromberg}@frm.utn.edu.ar
<http://dharma.frm.utn.edu.ar>

Resumen En el presente trabajo se ofrece un estudio experimental para segmentación de *madera* sobre imágenes de viñedos en condiciones realistas de clima e iluminación. Los resultados son muy alentadores y se presenta la *medición del diámetro de tronco* en las vides fotografiadas como un caso de uso real. Se consideran tres tecnologías de segmentación: *K-means*, *Gaussian Mixture* y *Support Vector Machines*. Además de *diámetro de tronco* se exponen otros casos en donde nuestro enfoque es potencialmente útil dentro del contexto de *Viticultura de Precisión*.

Palabras claves: *Image Segmentation, K-means, Gaussian Mixtures, Support Vector Machines, Precision Viticulture.*

1. Introducción

En las últimas décadas, y gracias a la disponibilidad de nuevas tecnologías de la información y sensado, se han comenzado a administrar las explotaciones agrarias de forma más precisa mejorando la gestión de la producción. Este conjunto de tecnologías reciben el nombre de *Agricultura de Precisión (AP)* [3], y cuando se aplican a la administración de viñedos toma el nombre de *Viticultura de Precisión (VP)* [4]. Algunos aportes de la VP que se pueden citar son: predicción de calidad de la uva y rendimiento del cultivo para una cosecha; detección de zonas de productividad alta, media y baja; detección y predicción de enfermedades; aplicación precisa de fertilizantes y pesticidas; maximización de la eficiencia de los insumos; minimización del impacto ambiental negativo, y más [3]. En la práctica, sin embargo, la medición de muchas variables de interés conlleva un elevado costo tanto en recursos humanos, como en equipamiento especializado para la medición. Por esto, uno de los objetivos de nuestra investigación es aplicar técnicas de *Segmentación de Imágenes* para la medición autónoma de variables vitícolas, siempre que sean medibles visualmente. Es común encontrar entre las tecnologías empleadas en AP técnicas de *Procesamiento y Segmentación de Imágenes*, como ser *Markov Random Fields*, *clustering*, *thresholding*, *edge detection*, *fuzzy sets*, y otros [1; 12], para resolver diversos problemas tales como fumigación diferencial, localización de

frutos en árboles, detección de defectos en un fruto, identificación de malezas, obtención de índices de color de vegetación, medición del diámetro de frutos, y más [11; 6; 13].

Segmentar una imagen consiste en particionar la misma en un conjunto de regiones no superpuestas que tienen una semántica coherente con respecto a una aplicación en particular. Formalmente, se define *segmentación* como un método para particionar una imagen I en R_k regiones, con $k = 1, \dots, K$, tal que cada *región* R_k es un objeto candidato [10]. Por lo tanto, *segmentación* es el agrupamiento de píxeles en regiones tal que

- Una región R_k es un subconjunto de píxeles de una imagen I .
- Las regiones R_k forman una partición.
- Cada *región* R_k satisface un *criterio de uniformidad*.
- Píxeles de regiones adyacentes no satisfacen el *criterio de uniformidad*.

Cada píxel de una imagen I es un punto compuesto por tres componentes que comprende las intensidades de los canales rojo, verde y azul (RGB) [10]. RGB es usado comúnmente para visualización de color, pero no es adecuado para análisis y segmentación de imágenes debido a la alta correlación que existe entre las intensidades de los tres componentes. Por lo tanto, las imágenes RGB originales son preprocesadas y convertidas al espacio de color $L^*u^*v^*$ (L =*luminescence*, u =*saturation*, v =*hue angle*) de la CIE (Comisión Internationale de l'Éclairage), escrito generalmente como $CIELUV$ o solamente LUV [10]. $CIELUV$ tiene varias ventajas sobre RGB en el ámbito de la segmentación de imágenes (consultar [7]).

A continuación presentamos un estudio de segmentación para imágenes de viñedos en escenas auténticas y condiciones realistas de clima y luminosidad. El objetivo es agrupar todos aquellos píxeles que pertenecen al material *madera*, discriminándolos de forma precisa del resto de los píxeles de la escena (llamados *background*). Los algoritmos a evaluar son K -means, *Gaussian Mixtures (GM)* y *Support Vector Machines (SVM)* [2]. Además se presentan resultados para la medición del *diámetro de tronco* de vides partiendo de las imágenes segmentadas y comparándolas con segmentaciones realizadas manualmente. En la Sección 2 se describen los elementos empleados en la experimentación, se especifica la metodología de evaluación y se dan detalles de los algoritmos implementados. Los resultados obtenidos se exponen en la Sección 3, y se presentan los resultados de la *medición del diámetro de tronco* en la Sección 4. Finalmente en la Sección 5 se ofrecen las conclusiones obtenidas en este trabajo.

2. Elementos de Trabajo y Metodología

Las imágenes necesarias para llevar a cabo la experimentación propuesta corresponden a viñedos de tipo *espaldero*, pertenecientes a la *Facultad de Ciencias Agrarias, UNCuyo* (Lujan de Cuyo, Mendoza). Las fotografías fueron realizadas con una cámara NIKON COOLPIX L16, en formato JPEG con una resolución de 2034×3072 . Sin embargo las imágenes fueron recortadas (no escaladas) a una resolución de 1024×1420 para reducir los requerimientos de memoria y procesamiento de los algoritmos. Dos de las imágenes empleadas en este trabajo se pueden apreciar en la Figura 1(a) y 1(c). Las mismas fueron

tomadas los primeros días del mes de Mayo de 2012, entre las 14:00 y las 15:00 horas, sin alterar la escena natural y bajo condiciones de clima y luminosidad realistas. Los únicos elementos “artificiales” agregados en la escena corresponden a la cinta azul que envuelve el tronco de la vid a unos 20cm de la cruz (lugar en donde se bifurca el tronco) y a la marca azul clavada sobre el tronco a unos 15cm de la cruz. Estas marcas fueron colocadas con propósitos de calibrar las distancias y medidas obtenidas por los algoritmos (Sección 4).

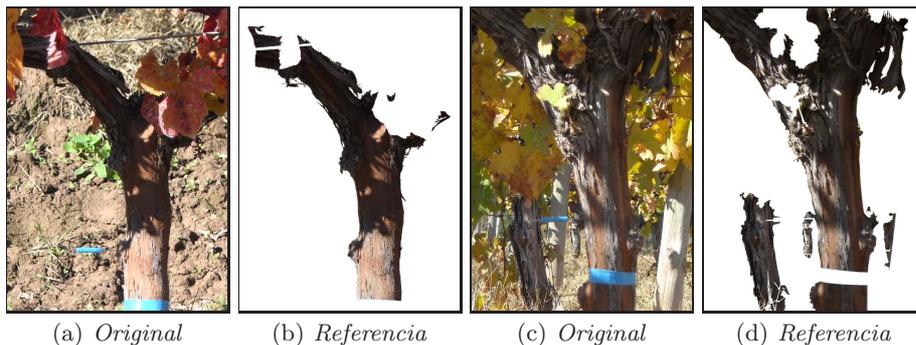


Figura 1. Imágenes originales e imágenes de referencia.

Se utilizaron tres algoritmos en este trabajo, cada uno con características y parámetros que los hacen muy diferentes entre sí: *K-means* agrupa datos no etiquetados en K clusters; *Gaussian Mixtures* aprende, a partir de regiones pre-etiquetadas manualmente, K distribuciones de probabilidad sobre los píxeles de la escena que representan la probabilidad del pixel de pertenecer a la k -ésima región; *SVM* aprende un clasificador, también a partir de regiones pre-etiquetadas, que determina a cual de las K clases pertenece cada pixel de la escena. Por lo tanto, existen dos enfoques: *supervisado* (*GM* y *SVM*) y *no-supervisado* (*K-means*). La cantidad de regiones se establece en $K = 3$, correspondientes a *madera*, *background*, y *marcas azules*, lo que define la cantidad de clusters en *K-means*, distribuciones en *GM* y clases en *SVM*. La región para las *marcas azules* se incluyó para evitar confundir al *background*. El total de píxeles disponibles para entrenamiento de los algoritmos es de 5,675,199, extraídos de un conjunto de *imágenes de entrenamiento* diferentes a las imágenes usadas para evaluación, de los cuales 1,646,961 pertenecen a *madera*, 3,981,693 pertenecen al *background* y 46,545 pertenecen a las *marcas* de referencia. Estos píxeles corresponden a zonas parciales de las distintas imágenes de entrenamiento que fueron demarcadas de forma aproximada en una o más zonas de *madera*, *background* y *marcas*, dejando sin etiquetar aquellos píxeles cercanos a la frontera de diferentes materiales ya que, debido al comportamiento del algoritmo JPEG, los píxeles en las fronteras se presentan de forma difusa. De esta manera se descartan píxeles dudosos en el entrenamiento de los algoritmos. Esta sección prosigue con la explicación de los métodos empleados en la evaluación de los algoritmos y a continuación se dan detalles de los algoritmos implementados.

Metodología de Evaluación. La evaluación de los algoritmos se realiza mediante *métodos de discrepancia* que comparan la imagen segmentada con una *imagen de referencia*, usando la diferencia para valorar el rendimiento [16]. La segmentación se considera un proceso de clasificación de píxeles con lo que se propone como *medida de discrepancia* diferentes índices sobre el número de píxeles correctamente/erróneamente clasificados. Las *imágenes de referencia* fueron segmentadas manualmente por los autores (Figura 1(b) y 1(d)).

Tabla 1. Ecuaciones y explicación de cada medida de calidad implementada.

$precision = \frac{tp_k}{tp_k + fp_k}$ (1)	Fracción de píxeles correctamente clasificados como <i>positive</i> para una clase, i.e., evalúa la capacidad de predicción del algoritmo para la <i>k-ésima</i> clase.
$recall = \frac{tp_k}{tp_k + fn_k}$ (2)	Proporción de píxeles <i>positive</i> reales que son clasificados correctamente como tal, i.e., evalúa la eficacia del algoritmo para los <i>positive</i> de la <i>k-ésima</i> clase.
$F-measure = 2 \times \frac{precision \times recall}{precision + recall}$ (3)	Proporción compuesta que relaciona <i>precision</i> y <i>recall</i> , calculado como la <i>media armónica</i> entre ellos, i.e, la inversa de la media aritmética.

En la Tabla 1 se presentan las medidas de calidad para la comparación de los algoritmos utilizadas típicamente en problemas de clasificación [14; 8]. Los términos *true positive* (*tp*), *true negative* (*tn*), *false positive* (*fp*) y *false negative* (*fn*) corresponden a contadores que se utilizan para contrastar los resultados de la segmentación con la *imagen de referencia*. Las expresiones *positive* y *negative* se refieren a la clasificación producida por el algoritmo, mientras que *true* y *false* se refieren a si esa clasificación concuerda con el valor de la etiqueta (*positive* o *negative*) del pixel en la *imagen de referencia*. En el contexto de nuestro problema un valor de *precision* = 1.0 significa que todos los píxeles clasificados como *positive* son *positive* en la *imagen de referencia*, pero no dice nada acerca de si todos los píxeles *positive* de la *imagen de referencia* fueron “encontrados” y etiquetados como tal, mientras que un valor de *recall* = 1.0 significa que todo pixel *positive* de la *imagen de referencia* fue clasificado como tal, pero no dice nada acerca de los píxeles incorrectamente clasificados como *positive*. Por lo tanto, un algoritmo que clasifique a todos los elementos como parte de una sola clase obtendría un valor de *recall* = 1.0, pero valores de *precision* muy bajos para esa clase, mientras que un algoritmo que clasifique correctamente solo un elemento de la clase lograría un valor de *precision* = 1.0 y un valor de *recall* cercano a cero. Esta relación inversa comienza a balancearse a medida que la clasificación se acerca a la clasificación real haciendo que *precision* y *recall* tiendan a 1.0. Por lo tanto estas magnitudes no se tratan de forma aislada, combinándose en una sola medida tal como *F-measure*. A continuación pasamos a describir los algoritmos de segmentación utilizados.

K-means. *K-means* es un método de *Clustering* iterativo descendiente usado para agrupar datos no etiquetados en un espacio multidimensional. Aquí se omite el desarrollo del algoritmo, el cual se encuentra explicado en detalle en [2]. Es necesario aclarar que la decisión de incluir *K-means* en la experimentación está fundamentada en la utilidad que tiene como línea base para la comparación del resto de los algoritmos presentados, ya que éste no es un enfoque particularmente sofisticado para el problema de segmentación de imágenes y en general es extremadamente difícil enfrentarlo con esta técnica, siendo propenso a errores cuando la escena esta compuesta por objetos que presentan superficies no homogéneas, superficies texturadas y/u objetos que se ven afectados por cambios de luminosidad o sombras. Estos problemas se hacen evidentes en la Figura 1(a) y 1(c). Las mismas exhiben una variedad de materiales, tonalidades y texturas diversas, pertenecientes a madera, follage, uva, tierra, malezas, alambre y otros elementos, mientras que la cantidad de clusters $K = 3$ que deseamos como salida es bastante menor. Sumado a esto, la *medida de disimilaridad* basada en información del color hace a la imagen extremadamente “ruidosa”, aún cuando el algoritmo opera en el espacio de color *CIELUV* sobre los componentes *UV*.

Gaussian Mixture. Una *GM* es un modelo compuesto por una superposición lineal simple de distribuciones Gaussianas que tiene por objeto proveer una clase multi-modal de modelos de densidad, frente a la clase uni-modal de una única Gaussianas. En [2] se explica de forma detallada el desarrollo del algoritmo. El enfoque propuesto aquí consiste en encontrar K distribuciones mediante *GMs*, una por cada material a segmentar, que modelen la distribución de probabilidad de los píxeles sobre los componentes *UV* del espacio de color *CIELUV*. Una vez entrenados los modelos se evalúa la probabilidad de cada pixel de la imagen a segmentar en cada distribución de probabilidad correspondiente a cada material, clasificando el pixel con la etiqueta de la distribución que mayor probabilidad entregó. En el proceso de entrenamiento fueron empleados la totalidad de los píxeles disponibles, i.e., 5,675,199 divididos en 1,646,961 de *madera*, 3,981,693 de *background* y 46,545 de *marcas*. Por último, el número C de componentes de la *GM* no es el mismo en cada modelo y es un parámetro provisto externamente al algoritmo: la distribución de píxeles pertenecientes a *madera* fue entrenada con $C = 4$; la distribución de píxeles que pertenecen al *background* (conformado por un conjunto de materiales) con $C = 6$ y la distribución de los píxeles que pertenecen a las *marcas* de referencia con $C = 1$ (lo que es igual a una distribución Gaussianas simple). Para establecer el número de componentes C de cada modelo se entrenaron distintas *GMs* para cada material y luego se optó reportar aquella con el mejor rendimiento sobre las *imágenes de evaluación*. El rango de valores explorado fue de $C = 1, \dots, 8$ para cada material.

Support Vector Machines para clasificación. Aquí se dan detalles sobre la implementación del algoritmo *SVM* para clasificación y se explican sus parámetros, para una revisión completa del mismo invitamos al lector a seguir la referencias [2; 5]. En la experimentación hemos utilizado *SVM* como un clasificador multiclase mediante el enfoque *one-versus-the-rest* [2], con un *kernel* de tipo *radial basis functions* y los parámetros $\gamma = 1/M$ y $C = 1$, siendo γ

un parámetro de la *radial basis function*, M la dimensionalidad del conjunto de datos, y C una constante de penalización por cada punto que viola los *límites de decisión*. A diferencia de GM , durante el proceso de entrenamiento del modelo no se empleó la totalidad de píxeles disponibles debido a los altos requerimientos de computo y memoria que hubiese demandado tal operación. Por lo tanto se tomó una muestra aleatoria de píxeles de cada material a segmentar, a saber, 10% de los 1,646,961 píxeles de *madera*, 5% de los 3,981,693 píxeles de *background* y 100% de los 46,545 píxeles de *marcas*, lo que hace un total de 410,325 píxeles disponibles para entrenamiento. Los porcentajes se fijaron en estos valores por los requerimientos de memoria y costo computacional asociados al algoritmo. Por último, con el objetivo de entregarle más información a SVM y mejorar su clasificación, el dataset de entrenamiento se aumentó con información de los componentes UV de los vecinos más próximos de cada pixel. Definimos el *vecindario* de un pixel p_{ij} como la región encerrada por una elipse vertical con centro en p_{ij} , semieje mayor $a = 4$ (sobre las abscisas) y semieje menor $b = 1$ (sobre las ordenadas). De esta manera el *vecindario* de p_{ij} queda conformado por diez vecinos, lo que aporta veinte nuevas variables al dataset de entrenamiento (cada pixel está definido por sus componentes UV) elevando la dimensionalidad del mismo a $M = 22$. El diseño del aspecto del *vecindario* no es azaroso y está fundamentado en las proporciones que mantiene el tronco de la vid en relación al resto de los objetos.

Wood Area. Si bien GM y SVM mejoran notablemente los resultados frente a la segmentación realizada por K -means (Sección 3), ambos algoritmos confunden ciertos píxeles de las escenas que pertenecen a *madera* con píxeles que pertenecen al *background* y viceversa. Sin embargo, en general, este fenómeno se produce en pequeñas zonas de la imagen y de forma diseminada, generando una segmentación “ruidosa” con grupos de píxeles dispersos, como muestran las imágenes de la Figura 2(c) y 2(d) sobre la segmentación obtenida para *madera*. Con este problema en mente elaboramos un mecanismo que realiza una reclasificación de los píxeles segmentados por GM y SVM . Este post-procesamiento, que llamamos *Wood Area (WA)*, mejora los resultados de la segmentación para GM y SVM , denominados GM -WA y SVM -WA. Este enfoque ha sido originalmente propuesto por [15] usando GM para segmentación de piel humana en imágenes y video a color. El mecanismo consiste en definir un área circular A_{ij} de radio r con centro en el píxel p_{ij} a reclasificar y un porcentaje P que determina el umbral de reclasificación. Si la proporción de píxeles *madera* en A_{ij} con $r = 30$ es mayor a $P = 0.7$, el píxel p_{ij} es reclasificado como *madera*, sino el píxel es reclasificado como *background*. Para establecer el valor de r se evaluó GM -WA y SVM -WA en el rango $r = 5, 10, 15, 20, 30, 45, 60, 90$, siendo $r = 30$ el caso en el que mejor rendimiento se obtuvo. Por otro lado, el valor de P se tomó directamente de [15].

3. Resultados

En la Figura 2(b) se puede observar el resultado de la segmentación obtenida con K -means para $K = 3$ clusters, mientras que en la primer fila de la Tabla

2(a) se puede consultar los resultados de *K-means* para *madera*, donde vemos que si bien obtiene uno de los valores más altos de *recall*, el puntaje de *precision* es demasiado bajo, lo que perjudica notablemente el valor de *F-measure*. Por lo tanto, *K-means* es bueno clasificando píxeles de *madera* (o *true positives*) pero incluye incorrectamente demasiados píxeles del *background* (o *false positives*). Los resultados que se observan en la Tabla 2(b) para *background* muestran el comportamiento inverso, es decir, valores altos de *precision* y bajos de *recall*.

El mismo análisis presentado para *K-means* se realiza para *GM* y *SVM*. Las segmentaciones obtenidas por ambos algoritmos se encuentra en la Figura 2(c) y 2(d) respectivamente. Ambos algoritmos mejoran la marca global de *K-means*, obteniendo para *madera* valores aproximados de *recall* y muy superiores de *precision*, lo que eleva el valor de *F-measure*. Comparando *GM* y *SVM* a través de *F-measure*, el segundo es superior por un margen pequeño. Mirando en la Tabla 2(b) correspondiente a *background*, vemos que *SVM* vuelve a superar a *GM* según *F-measure*, esta vez por una diferencia más ajustada que las hace prácticamente equivalente.

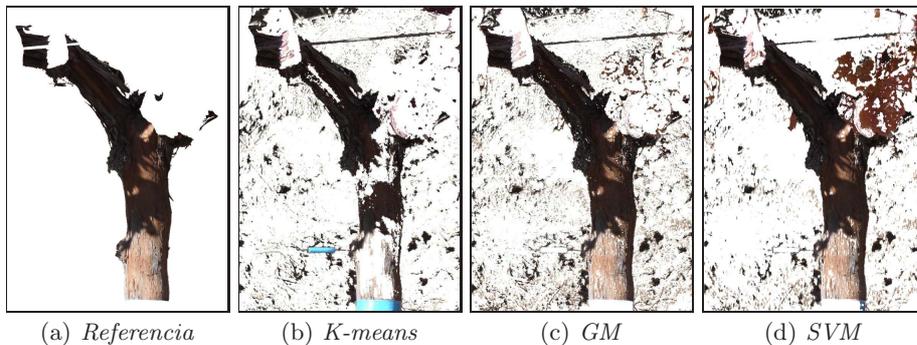


Figura 2. Imagen de referencia y segmentación lograda por cada algoritmo.

El análisis sobre *GM-WA* y *SVM-WA* nos revela mejoras considerables sobre las técnicas anteriores para *madera* y *background*. En el caso de *madera*, si bien los valores de *recall* caen apreciablemente, el aumento en *precision* es tan importante que *F-measure* supera a sus competidores anteriores. Para el caso de *background*, las mejoras son más notables, con valores de *recall* cercanos a 0.98 y valores de *precision* aproximados a 0.87, llegando *F-measure* hasta valores próximos a 0.92. Estas mejoras se pueden ver en la Figura 3(b) y 3(d).

Por último, comparando los resultados de la Tabla 2(a) con la Tabla 2(b) encontramos que *GM-WA* es el algoritmo que mejores resultados de segmentación obtiene para las métricas planteadas en el presente trabajo. Además, en general, *GM* necesita menos capacidad de cómputo y requerimientos de memoria que *SVM*, por lo que lo hace preferible frente a éste. Sin embargo *GM* necesita un primer procesamiento empírico o analítico de los datos para definir el número de componentes *C* de cada material a segmentar, mientras que *SVM* no lo requiere.

Tabla 2. Valores promedio para cada métrica propuesta sobre cinco imágenes.

(a) Resultados para <i>madera</i>				(b) Resultados para <i>background</i>			
	<i>precision</i>	<i>recall</i>	<i>F-measure</i>		<i>precision</i>	<i>recall</i>	<i>F-measure</i>
<i>K-means</i>	0.4461	0.8138	0.5672	<i>K-means</i>	0.9090	0.5652	0.6707
<i>GM</i>	0.6662	0.8187	0.7275	<i>GM</i>	0.9143	0.8379	0.8731
<i>SVM</i>	0.7185	0.8065	0.7511	<i>SVM</i>	0.9042	0.8760	0.8882
<i>GM-WA</i>	0.9596	0.6809	0.7915	<i>GM-WA</i>	0.8799	0.9860	0.9291
<i>SVM-WA</i>	0.9229	0.6607	0.7602	<i>SVM-WA</i>	0.8638	0.9784	0.9161

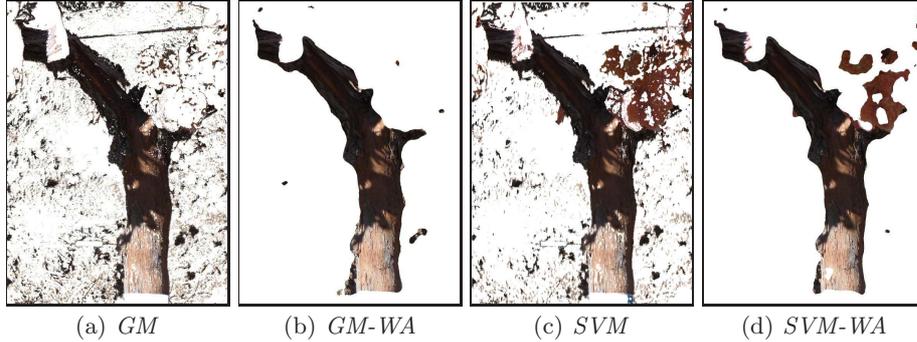


Figura 3. WA aplicado a *GM* en (b) y a *SVM* en (d) con $r = 30$ y $P = 0.7$.

4. Aplicaciones: medición del diámetro de tronco

Como aplicación de los resultados obtenidos en este trabajo se presenta la medición del *diámetro de tronco*, una variable de interés agrícola que se correlaciona con la capacidad de la planta de sostener el crecimiento y maduración de los brotes, hojas y frutos, ya que el tronco tiene la función de transportar agua y nutrientes, así como acumular sustancias reservantes [9]. En la práctica, la variable diámetro de tronco es medida de forma manual usando calibres o cintas métricas. De acuerdo a comunicaciones privadas con Ingenieros Agrónomos especializados en viticultura, en campo generalmente se trabaja con una escala de $5mm$ debido a las condiciones de trabajo, logrando una precisión de $\pm 2.5mm$, redondeando valores tales como $3.7cm$ a $3.5cm$ y $3.8cm$ a $4cm$. En este trabajo presentamos resultados medidos en píxeles y realizamos una calibración manual para estimar el error cometido en milímetros, dejando el problema de calibrado automático, complejo e interesante, para trabajos futuros. Estos resultados se presentan como medidas del diámetro en diferentes posiciones horizontales (filas de la imagen), correspondientes a las diferentes posiciones donde puede medirse el diámetro.

La Figura 4(a) corresponde la imagen segmentada por *GM-WA*, y requiere de una breve explicación: las regiones de color rojo representan píxeles correctamente clasificados como *madera* (*true positives*); las de color blanco píxeles clasificados correctamente como *no madera* (*true negatives*); las regiones de color azul representan píxeles incorrectamente clasificados como *madera* (*false positives*); y finalmente las regiones de color verde representan píxeles que pertenecen a *madera* y no han sido clasificados como tal (*false negatives*). Por

otro lado, las flechas grises en la Figura 4(a) señalan la zona del tronco totalmente vertical que comienza aproximadamente a 5cm de la cruz del tronco y finaliza a 20cm de la misma, y que aquí llamamos *región de medición*. Dado que en cada fila de la imagen los errores cometidos son de dos tipos (*false positives* y *false negatives*), los resultados se reportan en términos de *F-measure*, *precision* y *recall*. Analizando las posiciones horizontales dentro de la *región de medición*, en la Figura 4(c) notamos que la mayoría de los valores de *precision* y *recall* son mayores a 0.9, mientras que en la Figura 4(b) *F-measure* alcanza valores mayores a 0.95. Estos valores representan un error menor a 3.4mm (15 píxeles en esta imagen) en el diámetro del tronco con respecto al diámetro calculado desde la imagen segmentada manualmente. Comparado con la precisión de una medición manual vemos una diferencia de 0.9mm, sin embargo estos resultados nos motivan a continuar investigando técnicas que reduzcan estos errores.

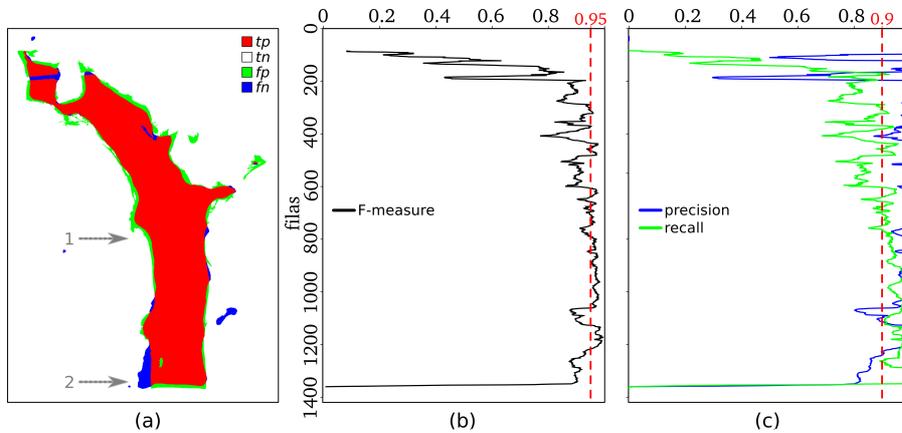


Figura 4. Valor de *F-measure* (b), *precision* y *recall* (c) para cada fila de píxeles de la imagen segmentada por *GM-WA* (a).

5. Conclusiones y Trabajo Futuros

En primer lugar queda claro que *K-means* no es un enfoque robusto para alcanzar resultados de segmentación que compitan con *GM* y *SVM*. Tanto *GM* como *SVM* obtienen resultados de segmentación equivalentes, mientras que el mecanismo *WA* aumenta la calidad de la segmentación para estos algoritmos en forma apreciable. Además, la información de color de un píxel aumentada con información de color de píxeles vecinos mejora notablemente los resultados de segmentación. Los resultados obtenidos en este trabajo son prometedores, e incentivan la investigación de otras posibles aplicaciones. De acuerdo a comunicaciones privadas con Ingenieros Agrónomos, consideramos investigar en un futuro otras aplicaciones en *VP* donde *Segmentación de imágenes* es potencialmente aplicable, como ser área foliar, índice verde, fertilidad de brote, entre otros [9]. Algunos enfoques que pueden mejorar los resultados logrados en este trabajo consisten en aumentar la información entregada a los algoritmos a través de la generación de mapas de profundidad de la escena con técnicas de

visión estereoscópica o sensores de profundidad, así como la inclusión de técnicas de modelos dinámicos ajustables (dynamic templates).

Agradecimientos. Este trabajo fue financiado por la *Universidad Tecnológica Nacional* a través de la beca de postgrado del Ing. D.S. Pérez y por *CONICET* mediante el cargo perteneciente al Dr. F. Bromberg. Agradecemos a la *Cátedra de Fisiología Vegetal, Facultad de Ciencias Agrarias, UNCuyo*, al Ing. Agr. Francisco Gonzalez, Ing. Agr. Roberto Borgo e Ing. Agr. Carina Gonzalez, por el tiempo y toda la información que nos brindaron de forma desinteresada.

Referencias

1. S. Bhattacharyya, *A brief survey of color image preprocessing and segmentation techniques*, Journal of Pattern Recognition Research **6** (2011), no. 1, 120–129.
2. C.M. Bishop, *Pattern recognition and machine learning*, Springer, 2009.
3. R.G.V. Bramley, *Lessons from nearly 20 years of precision agriculture research, development, and adoption as a guide to its appropriate application*, Crop and Pasture Science **60** (2009), no. 3, 197–217.
4. R.G.V. Bramley and D.W. Lamb, *Making sense of vineyard variability in australia*, Proc. Internat. Symp. on Precision Viticulture, Ninth Latin American Congr. on Viticulture and Oenology, 2003, pp. 35–54.
5. C. Campbell, *An introduction to kernel methods*, Studies in Fuzziness and Soft Computing **66** (2001), 155–192.
6. Y.R. Chen, K. Chao, and M.S. Kim, *Machine vision technology for agricultural applications*, Computers and Electronics in Agriculture **36** (2002), no. 2, 173–191.
7. H.D. Cheng, XH Jiang, Y. Sun, and J. Wang, *Color image segmentation: advances and prospects*, Pattern recognition **34** (2001), no. 12, 2259–2281.
8. T. Fawcett, *Roc graphs: Notes and practical considerations for researchers*, Machine Learning **31** (2004), 1–38.
9. L. Hidalgo Fernandez-Cano and J. Higoalgo Togores, *Tratado de viticultura*, Mundi prensa, 2011.
10. R. Jain, R. Kasturi, and B.G. Schunck, *Machine vision*, vol. 5, McGraw-Hill New York, 1995.
11. AR Jimenez, R. Ceres, JL Pons, et al., *A survey of computer vision methods for locating fruit on trees*, Transactions of the ASAE-American Society of Agricultural Engineers **43** (2000), no. 6, 1911–1920.
12. L. Lucchese and S.K. Mitra, *Color image segmentation: A state-of-the-art survey*, Proceedings of the Indian National Science Academy (INSA-A). Delhi, Indian: Natl Sci Acad **67** (2001), no. 2, 207–221.
13. CL McCarthy, NH Hancock, and S.R. Raine, *Applied machine vision of plants: a review with implications for field deployment in automated farming operations*, Intelligent Service Robotics **3** (2010), no. 4, 209–217.
14. M. Sokolova, N. Japkowicz, and S. Szpakowicz, *Beyond accuracy, f-score and roc: a family of discriminant measures for performance evaluation*, AI 2006: Advances in Artificial Intelligence (2006), 1015–1021.
15. M.H. Yang and N. Ahuja, *Gaussian mixture model for human skin color and its application in image and video databases*, Proc. SPIE: Storage and Retrieval for Image and Video Databases VII, vol. 3656, Citeseer, 1999, pp. 458–466.
16. Y.J. Zhang, *A summary of recent progresses for segmentation evaluation*, Advances in image and video segmentation (2006), 423.