

MORFOMETRÍA EN IMÁGENES DE CÉRVIX PARA LA DETERMINACIÓN DE LESIONES ESCAMOSAS

Ing. Eduard Leonardo Sierra

**UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA DE SISTEMAS E INDUSTRIAL
BOGOTÁ D.C.
2006**

MORFOMETRÍA EN IMÁGENES DE CÉRVIX PARA LA DETERMINACIÓN DE LESIONES ESCAMOSAS

Ing. Eduard Leonardo Sierra

Tesis presentada como requisito parcial para obtener el título de:
**MAGISTER OF SCIENCE EN INGENIERÍA DE SISTEMAS Y
COMPUTACIÓN**

Director
Eduardo Romero PhD.

**UNIVERSIDAD NACIONAL DE COLOMBIA
FACULTAD DE INGENIERÍA
DEPARTAMENTO DE INGENIERÍA DE SISTEMAS E INDUSTRIAL
BOGOTÁ D.C.
2006**

Nota de Aceptación

Director

Jurado

Jurado

Comentarios

Fecha: _____

Muchas gracias

CONTENIDO

	pág.
LISTA DE TABLAS	11
LISTA DE FIGURAS	12
1 INTRODUCCIÓN	15
2 MARCO TEÓRICO	17
2.1 Cancer de cuello uterino	17
2.1.1 Diagnóstico automático de cáncer	17
2.2 Procesamiento de imágenes histológicas	17
2.3 Procesamiento de imágenes en color	19
2.4 Procesamiento de texturas	19
2.5 Aprendizaje de máquina	19
3 MÉTODOS	21
3.1 Preprocesamiento	21
3.2 Estudio de segmentación	21
3.3 Construcción de la Morfometría	21
4 RESULTADOS	23
5 ANÁLISIS DE RESULTADOS	25
6 DESARROLLO DE SOFTWARE	27
6.1 Prototipo Inicial	27
6.1.1 Análisis y diseño	27
6.1.2 Algoritmos	27
6.1.3 Verificación	27
6.2 Prototipo segmentación	27
6.2.1 Análisis y diseño	27
6.2.2 Algoritmos	27
6.2.3 Verificación	27
6.3 Prototipo Morfometría	27
6.3.1 Análisis y diseño	27
6.3.2 Algoritmos	27
6.3.3 Verificación	27

7 CONCLUSIONES	29
BIBLIOGRAFÍA	31

LISTA DE TABLAS

pág.

LISTA DE FIGURAS

pág.

1. INTRODUCCIÓN

En Colombia el cáncer cuello uterino es uno de los de más alta incidencia, un estudio realizado en el Instituto Nacional de Cancerología (INC) durante 2002 [9] revisó los casos nuevos de cáncer en dicho instituto y encontró que: "los tipos de cáncer más frecuentes fueron cuello uterino (14,4%), piel (13,3%), mama (12,2%), estómago (6,5%) y sistema hematopoyético (5,5%). El cáncer de pulmón tuvo una baja participación (2,2%). El 79,8% de los casos proviene de Bogotá y los departamentos vecinos" [9]. Otro estudio del mismo instituto hecho sobre causas básicas de defunción en el año 2002; muestra que el cáncer de cuello uterino es la segunda causa de defunción con 10.8% (el primero es el cáncer de estómago con 10.9%) y la primer causa de defunción en mujeres con un porcentaje del 18.8% [8].

La biopsia de cervix es uno de los estudios que se realizan para la determinación de alteraciones histopatológicas y permite ayudar en el diagnóstico de lesiones escamosas intraepiteliales (en [11] se encuentra una descripción detallada de estas lesiones). A las muestras se les aplica un método de tinción (Hematoxilina-Eosina) y se observan en el microscopio. El patólogo determina si hay alteraciones basándose en los siguientes criterios: pérdida de la polaridad de las células, mitosis anormales, alteración de la relación núcleo-citoplasma, hiperchromasia nuclear y alteración de la maduración. En general los patólogos tienen un alto grado de seguridad cuando la muestra es normal o cuando definitivamente tiene displasia, pero es muy difícil determinar el estado en algunos casos en los cuales se presenta un desorden celular temprano. Esto es muy importante, puesto que la efectividad de los tratamientos depende directamente de un diagnóstico temprano y de la posibilidad de hacer seguimiento y/o tratamiento a las pacientes que verdaderamente presentan patología pre-neoplásica.

Con este proyecto se busca encontrar métodos de cuantificación, que puedan ser usados como herramientas que permitan apoyar al patólogo en sus diagnósticos y hacerlos más objetivos. Se desarrollan métodos de cuantificación que soporten el diagnóstico de lesiones escamosas basados de imágenes de cervix. Estos métodos se seleccionaron de acuerdo criterios de los empleados por los patólogos apropiados para construir una morfometría usando procesamiento digital de imágenes. Para el desarrollo de la morfometría se implementaron algoritmos que permiten segmentar las células en imágenes de cervix, estas segmentaciones se validaron con los resultados de las segmentaciones manuales de células en imágenes de cervix hechas con el apoyo de patólogos expertos. Finalmente se desarrolló una herramienta que puede ser usada como apoyo por patólogos en el diagnóstico de lesiones escamosas en imágenes de cervix.

2. MARCO TEÓRICO

2.1 Cancer de cuello uterino

El cáncer cuello uterino (cáncer cervical) es hoy el segundo más común en mujeres en todo el mundo [2], el diagnóstico temprano de esta enfermedad en una etapa precancerosa es muy importante para la efectividad de los tratamientos. Existen diversos exámenes para su diagnóstico como la citología vaginal, la colposcopia, la inspección visual directa, las espectroscopías fluorescente y reflectiva, entre otras; la mayoría de países del tercer mundo usan estudios de citología vaginal o inspección visual directa ya que son de muy bajo costo. En los sitios donde se ha implementado la citología vaginal se ha logrado bajar la incidencia de cáncer en un 70% pero esta prueba tiene inconvenientes ya que la sensibilidad de la misma es sólo del 51%, y la interpretación para dar un diagnóstico es altamente subjetiva [2].

Se han desarrollado diferentes proyectos que permitan ayudar a controlar esta enfermedad. Como por ejemplo el desarrollo de redes de investigación para estudiar la incidencia del virus del papiloma humano en el desarrollo de este cancer como se ve en [2].

2.1.1 Diagnóstico automático de cáncer

Recientemente se han reportado diferentes proyectos que buscan introducir el diagnóstico automático de cáncer basado en imágenes histopatológicas. En el reporte técnico de Demir y Yener [1] se encuentran documentados algunos de los pasos del diagnóstico de cáncer, ilustrados en la figura 1, estos pasos incluyen: preprocesamiento para eliminar ruido y segmentar regiones de interés en imágenes microscópicas, selección y extracción de características. Para el caso del cáncer de origen cervical entre las características mencionadas se encuentra la textura, basada en fractales y la topología. Finalmente en el diagnóstico se emplean algoritmos de aprendizaje con clasificadores, métodos de evaluación y pruebas estadísticas.

Otro ejemplo es el de Price et. al. quién presenta el desarrollo de un sistema de decision basado en redes de decisión bayesianas (bayesian belief networks) para dar soporte más objetivo al diagnóstico asociado a los exámenes de patología; que permitan a patólogos dar mejor sus diagnósticos, reproducirlos y que puedan ayudar a patólogos jóvenes en el aprendizaje [10].

2.2 Procesamiento de imágenes histológicas

Otro resumen importante es el de Loukas y Linney [6], el cual presenta el análisis de imágenes histológicas basado en tres factores biológicos que influyen en la respuesta de

los tumores a la radiación: la proliferación, la hipoxia y la vasculatura. Al igual que la revisión de Demir y Yener [1] presentan diferentes pasos donde se destaca el problema de segmentar las células. Aquí hay muchos enfoques que incluyen técnicas tradicionales de imágenes usando umbralización, detección de bordes con máscaras de gradiente, hasta técnicas donde se aprovecha la información de color incluyendo métodos usando modelos donde se separa la componente de iluminación como el HSI (Hue Saturation Intensity) y el HSL (Hue Saturation Luminance). También uso de algoritmos de clasificación de píxeles en el espacio de color RGB, también usando análisis de componentes principales (ACP) fusión de regiones, emparejamiento de plantillas (contornos activos, transformada de Hough), análisis de texturas y clasificación (Redes neuronales, lógica difusa) [6].

Por otro lado, Lezoray [5] en su tesis doctoral hace segmentación de células en imágenes en color con morfología matemática aplicada a imágenes de redes neuronales. Un trabajo anterior hecho por Walker et al. en 1994 [12], aplicado a células cervicales usa también morfología y extracción de características de textura.

Entre las técnicas de segmentación de células con aplicaciones en cancerología hay una comparación de métodos supervisados, basada en píxeles hecho por Meurie et al. [7], con imágenes de cáncer bronquial, en donde usa un esquema de segmentación en cinco etapas: simplificación de imágenes usando pre-tratamiento con suavizado, eliminación de ruido; seguido de clasificación de píxeles usando Bayes, kNN, SVM; extracción de marcadores basados en morfología matemática; luego watersheds para hacer crecimiento usando la información de color de la imagen obteniendo las regiones correspondientes al núcleo y al citoplasma. En este trabajo se concluyó que es de importancia la escogencia del binomio compuesto por el modelo de color y del clasificador para obtener mejores resultados, hay diferentes combinaciones que dan mejores resultados ellos encontraron que para Bayes los mejores espacios de color son YCh1Ch2, RGB y XYZ, para SVM son YCh1Ch2, HSL y RGB, y para kNN los mejores son HSL, XYZ y RGB [7].

Otros enfoques tratan de solucionar problemas que se presentan con segmentación de células usando representaciones paramétricas para hacer emparejamiento de plantillas como contornos activos o transformada Hough [6][3]. Entre estas también hay técnicas que plantean el problema de segmentar las células como un problema de optimización de una representación paramétrica de la forma celular. Segmentación satisfactoria de células aisladas, múltiples células y células traslapadas se logra con la técnica de Wu, Barba y Gil [13]. Este método hace minimización sobre una representación paramétrica pero el costo computacional es prohibitivo para imágenes grandes donde hay muchas células. Otro esquema basado también en optimización usa un método evolutivo basado en búsqueda tabú de Jiang y Yang [4]. Este método usa memoria de corto plazo y tiene una lista para diversificar el espacio de búsqueda pero el problema esta precisamente en la determinación del tamaño de esta lista. El número de familias es importante, si es pequeño se acerca a la búsqueda tabú, de otra manera la velocidad de convergencia empeora. Además el determinar al número de hijos también afecta la velocidad de convergencia para ese caso usaron valores entre 2 y 8. Además los autores encontraron que este método tenía casos donde no se obtenían segmentaciones óptimas.

2.3 Procesamiento de imágenes en color

2.4 Procesamiento de texturas

2.5 Aprendizaje de máquina

El procesamiento.. [3]

3. MÉTODOS

3.1 Preprocesamiento

3.2 Estudio de segmentación

3.3 Construcción de la Morfometría

4. RESULTADOS

5. ANÁLISIS DE RESULTADOS

6. DESARROLLO DE SOFTWARE

6.1 Prototipo Inicial

6.1.1 Análisis y diseño

6.1.2 Algoritmos

6.1.3 Verificación

6.2 Prototipo segmentación

6.2.1 Análisis y diseño

6.2.2 Algoritmos

6.2.3 Verificación

6.3 Prototipo Morfometría

6.3.1 Análisis y diseño

6.3.2 Algoritmos

6.3.3 Verificación

7. CONCLUSIONES

.....

BIBLIOGRAFÍA

- [1] Cigdem Demir and Bülent Yener. Automated cancer diagnosis based on histopathological images: a systematic survey. Technical Report TR-05-09, Rensselaer Polytechnic Institute, Department of Computer Science,, Troy, NY 12180 , US, March 2005. 2.1.1, 2.2
- [2] Daron G. Ferris. Analysis of digitized cervical images detect cervical neoplasia. In J. Michael Fitzpatrick and Milan Sonka, editors, *Proceedings of SPIE*, volume 5370 of *Medical Imaging 2004: Image Processing*, pages 181–194, 2004. 2.1
- [3] Rafael González and Richard E. Woods. *Digital Image Processing*. Prentice-Hall Inc., Upper Saddle River, New Jersey., 2 edition, 2001. 2.5
- [4] Tianzi Jiang and Faguo Yang. An evolutionary tabu search for cell image segmentation. *IEEE Transactions on Systems, Man, and Cybernetics*, 32(05):675–678, October 2002. 2.2
- [5] Olivier M. Lezoray. *Segmentation d’images couleur par morphologie mathématique et classification de données par réseaux de neurones : application à la classification de cellules en cytologie des séreuses*. Ph. d. these, Ecole Doctorale Structure, Information, Matière et Matériaux, U.F.R de Sciences, Université de Caen/Basse-Normandie, 2000. 2.2
- [6] Constantinos G. Loukas and Alf Linney. A survey on histological image analysis-based assessment of three major biological factors influencing radiotherapy: proliferation, hypoxia and vasculature. *Computer Methods and Programs in Biomedicine*, 74(03):183–199, June 2004. 2.2
- [7] C. Meurie, G. Lebrun, O. Lezoray, and A. Elmoataz. A comparison of supervised pixel-based color image segmentation methods. application in cancerology. *WEAS Transactions on Computers, special issue on ICOSIP’03*, 02(03):739–744, July 2003. 2.2
- [8] Miguel Ángel Castro, Marion Piñeros, and Teresa Martínez. Mortalidad en el Instituto Nacional de Cancerología, Colombia, 2002. *Revista Colombiana de Cancerología*, 07(03):20–13, 2003. 1
- [9] Constanza Pardo, Raúl Murillo, Marion Piñeros, and Miguel Ángel Castro. Casos nuevos de cáncer en el Instituto Nacional de Cancerología, Colombia, 2002. *Revista Colombiana de Cancerología*, 07(03):4–19, 2003. 1

- [10] G. J. Price, W. G. McCluggage, M. L. Morrison, G. McClean, L. Venkatraman, J. Diamond, H. Bharucha, R. Montironi, P. H. Bartels, D. Thompson, and P. W. Hamilton. Computarized diagnostic decision support system for the classification of preinvasive cervical squamous lesions. *Human Pathology*, 34(11):1193–1203, 2003. 2.1.1
- [11] José Saulo Torres. Lesiones escamosas intraepiteliales cervicales. l.e.i.c. *Revista Colombiana de Obstetricia y Ginecología*, 49(04), 1998. 1
- [12] Ross Walker, Paul Jackwa, Brian Lovell, and Dennis Longstaff. Classification of cervical cell nuclei using morphological segmentation and textural feature extraction. In *2nd ANZIS Australian and New Zealand Conference on Intelligent Information Systems*, 1994. 2.2
- [13] Hai-Shan Wu, Joseph Barba, and Joan Gil. A parametric fitting algorithm for segmentation of cell images. *IEEE Transactions on Biomedical Engineering*, 45(03):400–407, March 1998. 2.2