

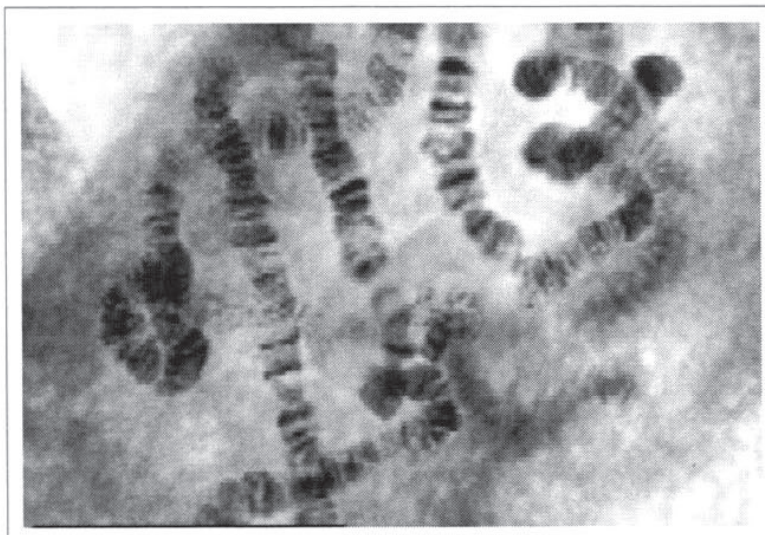
.....

Procesamiento Digital de Imágenes

OBTENCIÓN DE CARIOTIPOS

Aplicaciones al Área de la

Citogenética



RESUMEN

El presente artículo reseña un método para la obtención de cariotipos a partir de fotografías de metafases cromosómicas humanas [1] [2], procedimiento de gran importancia en el área de la citogenética. Para la obtención de los cariotipos se utilizan técnicas de procesamiento digital de imágenes que permiten obtener las características de cada cromosoma, como son su tamaño, la ubicación de su centrómero y la organización de sus bandas. El desempeño es analizado en términos de eficiencia en la organización de los diferentes pares de cromosomas.

PALABRAS CLAVES

Cariotipo, metafase, citogenética, procesamiento digital de imágenes, extracción de características.

ABSTRACT

This paper presents the design of a karyotyping method for spread metaphase images, one of the key procedures in the cytogenetics area. By means of digital image processing techniques, chromosome feature extraction (like centromere index, size computation as well as banding patterns evaluation) is possible. Methods for classifying are also

OSCAR GUALDRÓN GONZÁLEZ, Ph.D.
Profesor Asociado. Escuela de Ingenierías Eléctrica, Electrónica y Telecomunicaciones. Universidad Industrial de Santander.
Bucaramanga, Colombia.

CLARA INÉS VARGAS CASTELLANOS, MD.
Profesor Asistente. Jefe Laboratorio de Genética. Escuela de Medicina. Universidad Industrial de Santander.
Bucaramanga, Colombia.

MIGUEL ÁNGEL SOTAQUIRÁ GUTIÉRREZ.
Ingeniero Electrónico. Estudiante Maestría en Potencia Eléctrica. Universidad Industrial de Santander.
Bucaramanga, Colombia.

JORGE MARIO SÁNCHEZ NOGUERA.
Ingeniero Electrónico. Universidad Industrial de Santander.
Bucaramanga, Colombia.

*Grupo de Óptica y Tratamiento de Señales (GOTS)
Universidad Industrial de Santander
Bucaramanga-Colombia.*

22

.....

* Recibido : Diciembre 2002
* Aceptado: Septiembre 2003

described. The obtained results were measured in terms of classification efficiency of different types of chromosomes. Finally, research and educational applications are discussed.

KEYWORDS

Karyotyping, citogenetics, segmentation, metaphase, chromosome.

1. INTRODUCCIÓN

Para las personas que trabajan en el dominio de la genética, biología, medicina y afines, el análisis de los cromosomas es de gran utilidad, pues permite diagnosticar malformaciones y enfermedades en los individuos tratados, al igual que llevar a cabo investigaciones tendientes a determinar sus posibles orígenes y métodos de prevención o cura. Dicho análisis requiere la organización de los cromosomas de una manera reconocible en una imagen (cariotipo), de tal forma que sus características sean fácilmente visibles[1]. La técnica desarrollada comprende tanto la obtención de las características de la imagen (cromosomas), como su posterior organización en un cariotipo.

A pesar de que en la actualidad se cuenta con sistemas comerciales que realizan este trabajo, el objetivo principal es desarrollar un método que sea completamente abierto a futuras mejoras, con un menor costo y más asequible a grupos de investigación o como herramienta educativa.

En este artículo se resumen las etapas seguidas en el trabajo.

2. DESCRIPCIÓN DE CADA UNA DE LAS ETAPAS

2.1 Segmentación

Permite la diferenciación entre los cromosomas y el fondo dentro de la metafase celular (imagen en escala de grises adquirida previamente). La segmentación incluye etapas como la binarización, la limpieza de objetos o impurezas, la separación de cromosomas unidos y traslapados y la corrección de algunos cromosomas recortados o incompletos, producto de la binarización inicial.

2.1.1 Binarización. Busca una diferenciación preliminar entre los cromosomas y el fondo. Como resultado de este procedimiento, se obtiene una imagen binaria, donde los objetos de interés (cromosomas) están etiquetados con 1 y el fondo con 0. Para llegar a ello se

realiza un procesamiento global de la imagen en escala de grises. El núcleo del procedimiento se basa en el cálculo del histograma de niveles de gris de dicha imagen, con el cual se determina el umbral (o nivel de gris) adecuado para la binarización. Este tipo de procesamiento global permite conocer de manera directa los píxeles que pertenecen al fondo y los que pertenecen a un cromosoma.

2.1.2 Limpieza. La binarización identifica en ocasiones, como válidos, objetos que no corresponden a cromosomas. La limpieza permite eliminar dichos objetos de la imagen binaria.

2.1.2.1 Limpieza de zonas. Requiere la selección manual, por parte del usuario, de la zona que se desea eliminar (zona que contiene impurezas).

2.1.2.2 Limpieza de objetos. Uno o varios puntos pertenecientes al objeto que se desea eliminar, son seleccionados por el usuario. El algoritmo se encarga de la eliminación del objeto.

2.1.3 Separación. Algunos cromosomas aparecen unidos en la metafase binarizada. Es necesario separarlos para que el procesamiento posterior sea válido. Para ello, se crea inicialmente una matriz conformada únicamente por los cromosomas a separar, y el resultado se muestra al usuario para que delimite los puntos por donde desea realizar la separación (x,y). A través de cada par consecutivo de puntos x,y se traza un segmento de línea recta y los píxeles pertenecientes a él se hacen cero en la imagen que contiene el grupo de cromosomas a separar. El resultado es erosionado con una pasada, y luego dilatado dos veces [4], evitando en este caso conectar puntos que anteriormente no estaban unidos [3], [4], [5].

2.1.4 Corrección de cromosomas recortados. El procesamiento global de la imagen en grises, para obtener una imagen binaria, también da origen a cromosomas recortados o incompletos en algunas de sus zonas.

Inicialmente se hace la selección manual de la zona que se desea corregir. Con la zona en grises completamente aislada, se procede a determinar su umbral para su binarización, procedimiento exactamente igual al usado en el procesamiento global descrito con anterioridad. Finalmente se reemplaza el área correspondiente de la imagen binaria por la nueva zona corregida.

2.1.5 Separación de cromosomas traslapados. Debido a la preparación de la metafase, algunos cromosomas resultan traslapados y se hace entonces

necesaria su separación para obtener resultados adecuados en etapas posteriores. Se requiere en principio la selección manual de un punto que pertenezca a los dos cromosomas que se desean separar, para a continuación extraer sus componentes conexas y de esta forma determinar todos los píxeles correspondientes a los mismos. Después de ello el usuario debe encerrar el primer cromosoma a separar, de tal forma que la zona marcada se ajuste a la frontera del cromosoma en mención; con ello se crea una máscara del mismo tamaño de la imagen que contiene aquellos traslapados. Teniendo esta máscara se realiza un AND lógico entre ella y la matriz que contiene los cromosomas a separar, obteniendo así el primer de ellos aislado. Se determinan las dimensiones de éste (ancho y alto) así como las dimensiones de la imagen binaria, y de esta forma se inicia una búsqueda iterativa de una zona vacía dentro de la imagen binaria (con todos sus píxeles en 0), que se ajuste perfectamente al tamaño de dicho cromosoma. Se sitúan en ella el cromosoma separado tanto en escala de grises como el binario. Después de ubicar el primer cromosoma, el usuario debe encerrar el cromosoma restante, para realizar exactamente el mismo procedimiento descrito anteriormente.

2.2 Obtención de características

Características relacionadas con la forma y tamaño de los cromosomas (índices de centrómero y áreas), así como la distribución de las bandas[1],[2] a través de cada uno de ellos, permiten su diferenciación y posterior clasificación. El cálculo del índice de centrómero y el área de cada cromosoma hacen uso de la imagen binaria. La obtención del patrón de bandeo requiere, además de la imagen binaria, la imagen en escala de grises adquirida inicialmente.

2.2.1 Cálculo del índice de centrómero. Se determina para cada uno de los cromosomas su ángulo de inclinación con respecto a la vertical[4], se rota en dicha medida y se obtiene un cromosoma orientado verticalmente. Posteriormente se efectúa un barrido horizontal a través del objeto rotado, almacenando en cada caso el ancho de cada una de las filas, correspondiente al número de píxeles con valor 1. La posición del centrómero será entonces aquella fila que presente el menor ancho (menor cantidad de unos) y que a la vez divida al cromosoma en dos; de esta manera se elimina la posibilidad de tomar uno de los extremos del objeto como la posición del centrómero. Teniendo los dos componentes del cromosoma, se calcula el área de cada uno de ellos. El índice de centrómero es entonces la relación área menor sobre área mayor, y por tanto menor que 1. Ver figura 1.



Figura 1. Procedimiento para calcular el índice del centrómero

2.2.2 Cálculo de áreas. Equivale a evaluar la cantidad total de unos que se encuentran en el objeto (binario) analizado. Debido a que en ocasiones el poder de resolución usado en la adquisición de las diferentes imágenes no es el mismo (esto producto del uso de diferentes objetivos en el microscopio), es necesario llevar a cabo una normalización de cada una de las áreas, con respecto a aquella de mayor tamaño. De esta manera, el rango de valores estará entre 0 y 1, donde 1 corresponde al cromosoma de mayor tamaño.

2.2.3 Obtención de los patrones de bandeo. Inicialmente, se requiere calcular el ángulo de inclinación del cromosoma (tanto binario como en escala de grises), para así rotarlo en dicha medida. Con el cromosoma orientado verticalmente se procede a hacer un barrido por filas de la matriz que contiene al cromosoma binario, evaluando en cada caso el punto intermedio entre un extremo y otro del objeto ver figura 2. El conocimiento del punto intermedio permite evaluar el nivel de gris presente en dicha ubicación. Con ello se obtiene el perfil de gris a través del eje central del cromosoma, que es almacenado como un vector de longitud variable, dependiente de la imagen adquirida y del tamaño del cromosoma. Ver figura 3.



Figura 2. Obtención del eje central del cromosoma (cromosoma en grises, binario y con su eje central)

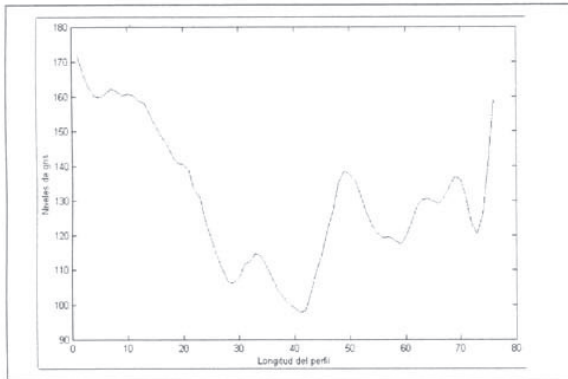


Figura 3. Perfil de gris típico obtenido al evaluar los niveles de gris del cromosoma a través de su eje central

2.3 Clasificación

Se entrenó una red neuronal probabilística, donde los datos suministrados para dicho entrenamiento consistían únicamente de los índices de centrómero y áreas (normalizadas con respecto a la de mayor tamaño) de una metafase humana normal, ya organizada en su cariotipo. La red posee 24 neuronas de salida (correspondientes a cada uno de los 22 posibles cromosomas homólogos, y a los cromosomas sexuales: X, Y), de las cuales solo una de ellas se activa (es decir, toma el valor 1), indicando la clase a la cual corresponde el cromosoma. Las demás neuronas permanecen inactivas (con valor cero). En ocasiones, no clasifica correctamente los cromosomas y es posible que, por ejemplo, identifique un cromosoma del par 4 como perteneciente al par 5, esto debido principalmente a la semejanza en la forma de estos dos tipos de cromosomas.

Como paso siguiente se calcula el coeficiente de correlación entre los perfiles de gris (patrones de bandeo) normalizados (con valores entre 0 y 1) de los cromosomas de una base de datos con las parejas ya encontradas mediante la red. Así por ejemplo, es posible que la red neuronal clasifique como correspondientes al par 13, un cromosoma del par 14, uno del par 15 y los dos del par 13 (por las razones descritas anteriormente) y por tanto, se tienen cuatro posibles cromosomas pertenecientes al par 13. El problema es resuelto al evaluar el coeficiente de correlación entre dichos cromosomas candidatos y el perfil correspondiente de la base de datos correspondiente al par 13. Se toman como correspondientes a dicho par, aquellos cromosomas que poseen mayor coeficiente de correlación.

2.4. Resultados

Se tomaron treinta imágenes de metafases cromosómicas que fueron organizadas usando el algoritmo explicado, evaluando entonces la eficiencia en la correcta clasificación de los diferentes tipos de cromosomas; como es mostrado en la tabla 1.

Tabla 1. Resultados obtenidos al usar el algoritmo planteado para la obtención del cariotipo

PARES	EFICIENCIA (%)
1	76.6
2	71.6
3	75.0
4	66.6
5	41.6
6	50.0
7	65.0
8	51.6
9	30.0
10	15.0
11	61.6
12	38.3
13	53.3
14	38.3
15	41.6
16	66.6
17	60.0
18	41.6
19	68.3
20	66.6
21	73.3
22	40.0
X	48.8
Y	0.0

Los resultados son bastante buenos en la clasificación de los pares 1,2,3,7,16,19 y 21, muy similares e incluso superiores a los obtenidos usando aplicaciones de tipo comercial. Pares como el 4,11,17 y 20 arrojan igualmente resultados satisfactorios. Sin embargo, aún existen algunas falencias en la clasificación de las parejas restantes. Específicamente, la eficiencia obtenida en la clasificación del cromosoma Y (0%) es debida a que no se contó con la suficiente cantidad de metafases con la presencia de cromosomas de este tipo.

2.5. Presentación

Las etapas descritas anteriormente (segmentación, obtención de características y clasificación) fueron

implementadas usando una interfaz gráfica desarrollada bajo la aplicación Matlab (versión 5.1). En ella, es posible la obtención del cariotipo partiendo de la imagen de una metafase celular, previamente adquirida. En la etapa final, clasificación, se permite al usuario la reubicación de los cromosomas clasificados de manera incorrecta, además del almacenamiento de los resultados obtenidos.

2.6. Perspectivas

En el desarrollo de futuros trabajos relativos al área de la citogenética, se puede pensar en el diseño de algoritmos que permitan elevar el grado de automatización del proceso, como son las rutinas para la separación automática de cromosomas unidos o traslapados[6].

De igual manera, procedimientos como los involucrados en la etapa de obtención de características y la clasificación, pueden ser mejorados para así elevar el grado de eficiencia en la organización de los cromosomas. Se puede pensar en refinar el procedimiento de obtención de los patrones de bandeo, de tal manera que las características de cada cromosoma no sean tan puntuales, y en lugar de ello recolecten la mayor cantidad posible de información, lo que redundará en mejores resultados en la etapa de clasificación. Se puede implementar el sistema desarrollado como una herramienta independiente del entorno Matlab posiblemente programando los diferentes algoritmos en lenguajes de alto nivel, como C o C++.

3. CONCLUSIONES

Se diseñaron y comprobaron procedimientos para la binarización de las metafases, separación de cromosomas unidos y traslapados, eliminación de impurezas y corrección de cromosomas recortados. La correcta segmentación de la metafase celular permitió la obtención de diferentes parámetros y características necesarias en la etapa final, el reconocimiento. Así, se desarrollaron rutinas para la obtención de índices de centrómero, áreas y patrones de bandeo a través del eje central, para cada cromosoma.

Se analizaron las diferentes posibilidades existentes para abordar el problema de reconocimiento y clasificación de los cromosomas. Se hizo uso de herramientas como las redes neuronales y otras técnicas del procesamiento digital de imágenes, logrando en última instancia la elaboración de un método asistido que permite la obtención del cariotipo.

Los resultados obtenidos son en muchos aspectos, comparables e incluso superiores a los proporcionados por las aplicaciones comerciales. Además, se tiene una herramienta lo suficientemente robusta que permitirá el futuro desarrollo de algoritmos que automaticen aún

más el proceso de organización de los cromosomas. El sistema desarrollado permite su uso como herramienta educativa y de investigación, de fácil uso y un costo mucho menor comparado con el de herramientas de carácter comercial.

4. BIBLIOGRAFÍA

- [1.] SALAMANCA G. Fabio. Citogenética humana, fundamentos y aplicaciones clínicas. Médica Panamericana. México. 1990. 440p.
- [2.] FORD E.H.R. Human chromosomes. Academic Press. Inglaterra. 1973. 381p.
- [3.] GONZÁLEZ Rafael - WOODS Richard. Tratamiento digital de imágenes. Addison-Wesley Iberoamericana S.A. Estados Unidos. 1996. 773p.
- [4.] JAIN Anil. Fundamentals of digital image processing. Prentice Hall. Estados Unidos. 1989. 569p.
- [5.] SERRA Jean P. Image análisis and mathematical morphology. Academic Press. Inglaterra. 1982. 610p.
- [6.] AGAM Gady - DINSTEIN Its'hak. Geometric separation of partially overlapping nonrigid objects applied to automatic chromosome classification. En: IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 19, No. 11. Noviembre. 1997. 16p.

AUTORES



OSCAR GUALDRÓN GONZÁLEZ. Físico, Ingeniero de Sistemas. Universidad Industrial de Santander. Bucaramanga, Colombia. Ph.D. Óptica. Universidad Laval. Québec, Canadá. Profesor Asistente. Escuela de Ingenierías Eléctrica, Electrónica y Telecomunicaciones. Universidad Industrial de Santander. e-mail: gualdron@uis.edu.co.



CLARA INÉS VARGAS CASTELLANOS. Médico. Universidad Industrial de Santander. Profesor Asociado, Jefe Laboratorio de Genética. Escuela de Medicina. Universidad Industrial de Santander. e-mail: cvargas@uis.edu.co.



MIGUEL ÁNGEL SOTAQUIRÁ GUTIÉRREZ. Ingeniero Electrónico. Estudiante Maestría en Potencia Eléctrica. Universidad Industrial de Santander. e-mail: miguel_sotaquirá@hotmail.com.



JORGE MARIO SÁNCHEZ NOGUERA. Ingeniero Electrónico. Universidad Industrial de Santander. e-mail: jeinzu1@hotmail.com