

Estudio comparativo de algoritmos de segmentación de piel usando atributos de color

Sheila Gonzalez-Reyna¹, Marlene Elizabeth López-Jiménez¹,
Emmanuel Zavala-Mateo¹, Israel Yañez-Vargas¹,
Jesús Guerrero-Turrubiates²

¹ Universidad Politécnica de Juventino Rosas,
Salamanca, Gto., México

² Universidad de Guanajuato, División de Ingenierías Campus Irapuato-Salamanca,
Salamanca, Gto., Mexico

{sgonzalez_ptc,marlene.lopez,emmanuel.zavala,jyanez_pa}@upjr.edu.mx
jdj.guerreroturrubiates@ugto.mx

Resumen. El presente artículo analiza una serie de métodos para el procesamiento de imágenes de personas, con el fin de segmentar su piel. La base de datos Jochen Triesch Static Hand Posture Database [9] proporciona una amplia gama de tonalidades de piel, obtenida de personas de razas y edades distintas. Las tareas de detección de piel en imágenes se enfrentan a diversos retos, debido a factores diversos como: variaciones en la iluminación, factores ambientales y fondo de la escena. En este trabajo se muestran los resultados obtenidos al desarrollar algoritmos de segmentación basados en distintas técnicas heurísticas y probabilísticas, a partir de la utilización de una base de datos en el espacio de color RGB.

Palabras clave: Segmentación de piel, umbralización, naive Bayes, perceptrón multicapa.

A Comparative Study of Skin Segmentation Algorithms Using Color Attributes

Abstract. This article analyzes a set of methods for processing images of people for skin segmentation. The Jochen Triesch Static Hand Posture Database [9] provides a wide range of skin tones obtained from people of different ages and breeds. The skin detection task on images give face to distinct challenges, due to various factors like: illumination variations, environment factors and background. In this paper we show the results obtained in developing segmentation algorithms based on different heuristic and probabilistic techniques using RGB color space.

Keywords: Skin segmentation, thresholding, naive Bayes, multilayer perceptron.

1. Introducción

La detección de piel tiene un papel muy importante en una amplia gama de aplicaciones del procesamiento de imágenes, que van desde la detección y seguimiento de rostros, detección de gestos manuales y sistemas varios de interacción humano-computadora. Recientemente, métodos de detección de piel basados en información de color han sido ampliamente utilizados, debido a que la segmentación del color de piel es computacionalmente eficaz, característica deseable para la implementación de funciones más complejas.

Los algoritmos de segmentación permiten separar los objetos de interés del fondo o del resto de la escena en una imagen digital. La segmentación de imágenes es considerada una etapa importante en la detección y seguimiento de objetos. Algunas aplicaciones importantes incluyen reconocimiento de rostros y gestos manuales, interacción humano-computadora, entre otros.

La información de color en las imágenes digitales está representada en el espacio de color RGB (Red, Green and Blue). Sin embargo, este espacio de color se ve altamente afectado por cambios en la iluminación, dificultando las tareas de segmentación por color. Para sobrellevar este problema, existen distintos espacios de color, capaces de separar la intensidad del color, del color en sí, generando de esta manera, una cierta invarianza a los cambios en iluminación.

El presente artículo realiza un análisis de distintos métodos de segmentación de piel utilizando el color en espacio RGB como fuente de información, y realizando la tarea de segmentar por tres métodos distintos: umbralización, modelado del espacio de color mediante probabilidades y redes neuronales artificiales.

El resto del artículo se estructura de la siguiente manera. En la Sección 2 se presenta un resumen de algunos algoritmos relacionados. La Sección 3 describe los conceptos teóricos en que se sustentan los experimentos. Los resultados experimentales se analizan en la Sección 4. Finalmente, la Sección 5 da las conclusiones finales y presenta las perspectivas de trabajo futuro.

2. Antecedentes

La segmentación de objetos en imágenes digitales puede realizarse mediante atributos de color, forma y textura. Las características de color pueden considerarse las más simples debido a que no suelen considerar datos de posición, rotación y escalamiento. Sin embargo, el color de los objetos se ve seriamente afectado debido a factores ambientales tales como sol, sombra, lluvia y niebla. Debido a lo anterior, la segmentación de imágenes utilizando atributos de color se implementa utilizando modelos de color alternativos, donde el color se separa de la iluminación. Algunos de los espacios de color más utilizados para llevar a cabo la tarea de segmentación de piel son RGB [1, 10], YCbCr [8, 9], HSV, HSI, CieLab [8, 12].

En [12] los autores realizan un estudio comparativo de nueve modelos de crominancia, incluyendo CIELab y CIELuv, utilizados en la segmentación de piel para la detección de rostros. Los autores de [8] realizan un experimento de

comparación en la segmentación de color de piel utilizando dos modelos de color: CIELab y YCbCr, dicho experimento lleva a la conclusión de que el modelo de color CIELab produce mejores resultados debido a su capacidad para representar una mayor cantidad de colores.

Además se han realizado trabajos en donde la segmentación de piel se realiza en espacio de color RGB. En [1], se usó la información de color RGB para clasificar piel mediante un árbol difuso. El porcentaje de reconocimiento obtenido en su base de datos es del 94.1%. Santos y Pedrini propusieron otro método para la segmentación de piel en espacio de color RGB, utilizando histogramas de color y apoyándose en mapas de salientes [10]. Sobieranski et al. [11] propuso la segmentación usando el mismo espacio de color, y clasificadores de distancia no lineal (Mahalanobis).

De acuerdo con el método de implementación, los algoritmos de segmentación por color pueden ser de tres tipos distintos [9]:

1. **Umbralización.** Este método implica un análisis de colores, con la finalidad de encontrar rangos de valores entre los cuales se encuentran los colores del objeto de interés,
2. **Modelos probabilísticos.** Se tiene una base de datos de colores, y se generará un modelo utilizando probabilidades que permitan segmentar el objeto de su fondo [10, 11], y
3. **Algoritmos de clasificación.** Los algoritmos clasificadores tienen su origen en distintas teorías, ellos pueden realizar una clasificación de objeto/fondo utilizando la información de color [1].

Los métodos de segmentación por color que involucran el uso de clasificadores, han tenido una especial atención en años recientes [7]. Esto debido a que los clasificadores se muestran robustos en la generalización de la información, es decir, se ven menos afectados por cambios de iluminación que los métodos de umbralización. Guerrero-Curries et al. [3] realiza una comparación del desempeño de distintos clasificadores en tareas de segmentación, usando también información de color en diferentes espacios (CIELab, RGB normalizado, YCbCr).

3. Métodos

En esta Sección se enumeran los conceptos necesarios para la reproducción de los experimentos.

3.1. Método de umbralización

El color de piel en los humanos se determina por la cantidad de pigmento “melanina” en la piel en función de la absorción de la radiación ultravioleta del sol. Las personas con grandes cantidades de melanina tienen piel oscura, y las personas con una cantidad pequeña de esta tienen piel blanca [7].

El primer paso para la clasificación de color de piel es la elección del espacio de color en que se trabajará. El color RGB es el espacio predeterminado

para la mayoría de formatos de imagen. Cualquier otro espacio de color es una transformación del RGB.

Los colores de piel humana difieren entre un grupo de individuos por muy pequeño que éste sea, incluso difiere en las zonas del cuerpo de un mismo individuo, y bajo cierto nivel de iluminación. Uno de los métodos más utilizados y más sencillos es la definición de límites o rangos (umbrales) de color, donde mediante una decisión se hará la agrupación de componentes espaciales. Los valores de los píxeles de una imagen que se encuentran dentro de los umbrales calculados son considerados píxeles de piel, de manera que

$$y(c) = \begin{cases} 1 & thr_{low} \leq c \leq thr_{high} \\ 0 & \text{cualquier otro caso,} \end{cases} \quad (1)$$

donde thr_{low} y thr_{high} son los umbrales inferior y superior, respectivamente para el canal de color que está siendo evaluado.

3.2. Clasificador Naive Bayes

El objetivo de las técnicas de clasificación es asignar a cada elemento de un conjunto la pertenencia a una clase determinada. Las clases identifican a un conjunto de elementos que comparten cierta similitud en una serie de características. Las distintas clases definidas dependen de la aplicación. Por ejemplo, en este caso el objetivo es distinguir entre dos clases: las que son “piel” y las que son “no piel”.

De acuerdo con [10], se puede construir un modelo Bayesiano para la clasificación de píxeles que representan piel/no piel. Considerando que c representa el color de un píxel, se tiene la siguiente relación:

$$P(piel|c) = \frac{P(c|piel)P(piel)}{P(c|piel)P(piel) + P(c|\neg piel)P(\neg piel)}. \quad (2)$$

La Eq. (2) representa la regla de Bayes para la toma de decisiones. Las probabilidades pueden calcularse mediante una distribución Gaussiana:

$$P(c|clase) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \frac{(c - \mu)^2}{2\sigma^2}, \quad (3)$$

donde μ es el valor del promedio de cada componente de color, y σ^2 es la varianza de la misma componente. La regla de Bayes puede ser simplificada suponiendo que cada característica (R, G, B) es independiente de las otras dos, generando de este modo el clasificador Naive Bayes [6].

3.3. Perceptrón Multicapa

El Perceptrón Multicapa (MLP) es el tipo más popular de redes neuronales [5]. Se ha aplicado con éxito en muchos problemas de reconocimiento de patrones

debido a su capacidad de aprender complejas relaciones entre entradas y salidas no lineales y la habilidad para generalizar los datos dados.

El perceptrón multicapa utilizado en este artículo tiene su estructura similar a la que se observa en la Fig. 1. A la entrada se encuentran los valores R, G y B del pixel a clasificar. A la entrada de la neurona 1, se calcula la suma ponderada de cada una de las entradas, más una señal de activación, como se muestra en 4.

$$N_{1,entrada} = [R \ G \ B \ 1] [W_{R,1} \ W_{G,1} \ W_{B,1} \ b_1]^T, \quad (4)$$

donde b_1 es la señal de activación, y los pesos $W_{c,1}, c \in \{R, G, B\}$ son las ponderaciones de los valores del pixel hacia la neurona 1. Una vez calculado el valor $N_{1,entrada}$, la salida de la neurona se obtiene aplicando la función sigmoideal a dicho valor:

$$N_1 = \frac{1}{1 + \exp(-N_{1,entrada})}. \quad (5)$$

Operaciones similares se realizan con las neuronas N_2 , N_3 y N_4 . Las salidas de las neuronas 3 y 4, etiquetadas como S_3 y S_4 se comparan para tomar la decisión final [2].

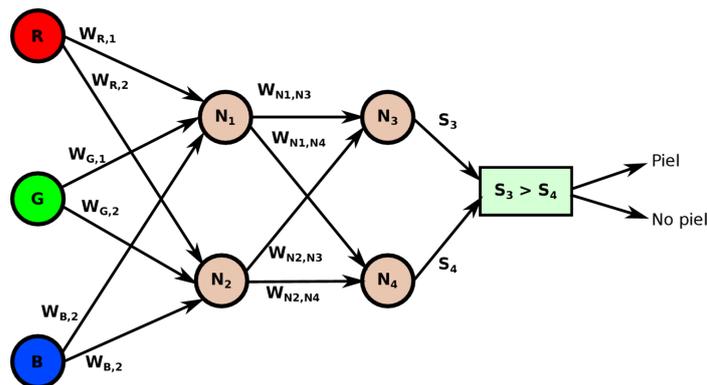


Fig. 1. Perceptrón multicapa para clasificación de color en espacio RGB

4. Resultados

Este artículo presenta un estudio comparativo de tres métodos para la segmentación de piel, que pueden ser utilizados para aplicaciones de interacción humano-computadora, detección de peatones, detección y reconocimiento de rostros, etc. Los métodos analizados son: 1) Umbralización, 2) clasificación usando Naive Bayes y 3) clasificación usando Redes Neuronales Multicapa (MLP). Estos métodos se aplican para lograr la segmentación de piel en imágenes en espacio de

color RGB ya que es la forma en la que se encuentra la base de datos; asimismo el tratamiento de imágenes en este espacio de color ahorra tiempo de procesamiento y muestra buenos resultados.

4.1. Base de datos de color de piel

Para la realización de los experimentos de este artículo, se utilizó la base de datos de Segmentación de Piel (Jochen Triesch Static Hand Posture Database) reportada en [1]. Esta base de datos está compuesta por un muestreo aleatorio de valores tomados de rostros de distintos grupos de edad (jóven, adulto, adulto mayor), grupos raciales (blanco, negro, asiático) y género. Está compuesta por un total de 200,045 muestras, de las cuales 50,859 pertenecen a muestras de tono de piel, y 149,186 son muestras de colores distintos a la piel. La Fig. 2 muestra las nubes de dispersión de la base de datos, graficada en dos canales de color de manera simultánea.

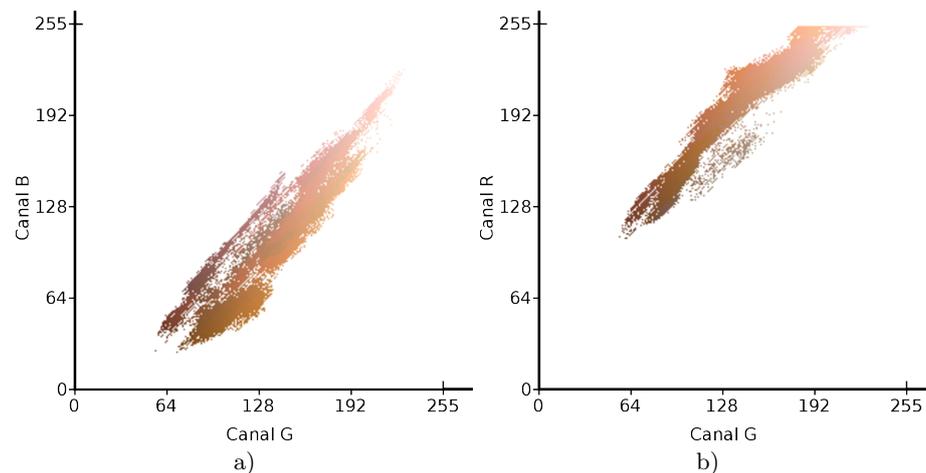


Fig. 2. Dispersión de la base de datos, a) canal G contra canal B, b) canal G contra canal R

4.2. Detección manual de umbrales para segmentación

La tarea de umbralización consiste en definir un rango de valores para cada canal de color, entre los cuales se encuentran los valores más comunes para el objeto de interés. La elección de los umbrales se realizó de forma heurística, es decir, acotando manualmente la nube de dispersión de la Fig. 2, en su región de mayor densidad, por inspección visual. Los valores de los umbrales escogidos se pueden observar en la Tabla 1.

Tabla 1. Umbrales propuestos para la segmentación, basados en la nube de dispersión de tonos de piel

Canal de color	Umbral menor	Umbral mayor
Azul	29	191
Verde	74	202
Rojo	130	250

Como se mencionó anteriormente, cada umbral es representado en el algoritmo de segmentación como una condición que controla el rango en cada canal(B,G,R) de la imagen sometida para cada pixel determinando así si este pertenece a la clase “Piel” en este caso con valor “1”, o a la clase “No Piel” con valor “0” para así obtener como resultado una nueva imagen binarizada.

La Tabla 2 muestra la matriz de confusión del proceso de umbralización, que se obtuvo para los umbrales de la Tabla 1. Los valores “Verdaderos Positivos” y “Falsos Positivos”, que se refieren a los datos que el algoritmo clasificó correctamente, son muy bajos en comparación con sus contrarios “Verdaderos Negativos” y “Falsos Negativos”, respectivamente. Debido a lo anterior, se obtuvo un porcentaje bajo de eficacia (16.2028 %) en este proceso de segmentación.

Tabla 2. Matriz de confusión utilizando método de umbralización

Clase “Piel”	Clase “No Piel”	
7,690	43,169	Clase “Piel”
124,463	24,723	Clase “No Piel”

4.3. Segmentación utilizando algoritmo Naive Bayes

La clasificación supervisada se basa en la disponibilidad de datos de entrenamiento, de los que se conoce a qué clase pertenecen; extrayendo características de estos elementos es posible diseñar un algoritmo clasificador que asigne a muestras futuras una clase determinada en función de sus características.

Uno de los algoritmos más populares de clasificación es el conocido como Naive Bayes, que para este caso, genera un modelo Gaussiano para cada canal de color, pues asume que son independientes. La Tabla 3 muestra la matriz de confusión del clasificador naive Bayes que segmenta piel a partir del color de un pixel. Puede observar una mejora consistente en la eficacia de la clasificación, que es de un 92.36 %, comparado con el 16.2 % que se obtuvo con la umbralización.

Tabla 3. Matriz de confusión utilizando el clasificador Naive Bayes

Clase "Piel"	Clase "No Piel"	
38,956	1,903	Clase "Piel"
3,378	145,808	Clase "No Piel"

4.4. Segmentación utilizando Redes Neuronales Multicapa

La tarea de segmentación se llevó a cabo utilizando un MLP. Esta red neuronal fue modelada utilizando Weka [4], y posteriormente se implementó en MATLAB[®].

La imagen que sea sometida al algoritmo será dividida en los tres canales de color para su posterior procesamiento. La red diseñada consta de cuatro nodos distribuidos en dos capas. Los valores de entrada a la red son los tres canales de color RGB de la imagen sometida y como salida se obtiene una etiqueta para el caso de cada pixel determinando si estos pertenecen a la clase "piel" o "no piel". El resultado es una nueva imagen binarizada, a la cual corresponden valores de "0" y "1" para las clases "No piel" y "Piel" respectivamente. Esta nueva imagen tiene las mismas medidas que la imagen original pero en blanco y negro.

Tabla 4. Matriz de confusión utilizando clasificador MLP

Clase "Piel"	Clase "No Piel"	
50,597	262	Clase "Piel"
768	148,418	Clase "No Piel"

En los resultados obtenidos de la matriz de confusión para el MLP (Tabla 4) se puede observar que los "Verdaderos Positivos" y los "Falsos Positivos" son cifras muy grandes, lo cual quiere decir que el porcentaje de eficacia es muy alto al utilizar este algoritmo de segmentación (99.48%).

4.5. Comparación de resultados obtenidos por los tres métodos

La Fig. 3.b) muestra el resultado de segmentar la imagen de la Fig. 3.a) mediante el proceso de umbralización. Observe que a pesar del bajo porcentaje de clasificación, la piel de las personas se segmenta correctamente. Por otro lado, hay información que debería formar parte del fondo (no es piel) y la umbralización no la descarta (falsos positivos).

La Fig. 3.c) despliega el resultado de la segmentación de la misma imagen, al utilizar el clasificador Naive Bayes para la toma de decisiones. Observe la mejora de los resultados, al existir una menor cantidad de falsos positivos (pixeles clasificados como piel, que realmente no lo eran).



Fig. 3. Resultados de la segmentación de una imagen. a) Imagen original. Segmentación utilizando b) umbralización, c) Naive Bayes y d) MLP

La imagen segmentada de la Fig. 3.d) se obtuvo mediante la red neuronal MLP. Observe una mayor limpieza en las zonas segmentadas, además de una cantidad considerablemente menor en el número de falsos positivos. Lo anterior se debe a los valores altos de sensibilidad (98.5%) y especificidad (99.82%) obtenidos tras el proceso de clasificación.

De acuerdo con los datos de la Tabla 5, es posible verificar la superioridad de un algoritmo de clasificación en una tarea de segmentación, cuando se compara contra el resultado obtenido por la umbralización. Considerando la misma tabla, los porcentajes globales tienen cierta variación, teniendo un 16.20% de eficacia en segmentación por “umbralización”, y una gran diferencia en el porcentaje global del método por “Naive Bayes” obteniendo un 92.36% y obteniendo un porcentaje de clasificación casi perfecta por el método “MLP” logrando un 99.48% de eficacia, teniendo menos del 1% como posible falla.

5. Conclusiones

En este proyecto se realizó el proceso de segmentación de color de piel, mediante tres algoritmos distintos para la base de datos Jochen Triesch Static

Tabla 5. Comparación en la eficiencia de los tres métodos de segmentación

Descripción	Umbralización	Naive Bayes	MLP
Porcentaje de pixeles clasificados correctamente	16.2028 %	92.3612 %	99.4851 %
Porcentaje de pixeles clasificados incorrectamente	83.7971 %	7.6388 %	0.5149 %
Sensitividad	5.8190 %	92.0205 %	98.5048 %
Especificidad	36.4151 %	92.456 %	99.8237 %

Hand Posture Database en espacio de color RGB. Se observó que el algoritmo basado en redes neuronales tiene mejores resultados para la detección de piel, sobre el algoritmo probabilístico de Naive Bayes y el método de umbralización, respectivamente. El espacio RGB crea muchos conflictos en el procesamiento dependiendo la iluminación del lugar donde sea tomada la imagen, sin embargo, los resultados mostrados aquí demuestran que dicho espacio es suficiente para llevar a cabo tareas de segmentación de piel, en ambientes con iluminación controlada (ambientes de interior, por ejemplo).

Este estudio se llevó a cabo para elegir el mejor proceso de segmentación y reducción de zonas de interés en imágenes que posteriormente serán sometidas a algoritmos de reconocimiento de gestos manuales. De la misma manera, las técnicas aquí exploradas pueden tener aplicaciones en otras tareas como detección de rostros o peatones.

Referencias

1. Bhatt, R.B., Sharma, G., Dhall, A., Chaudhury, S.: Efficient Skin Region Segmentation Using Low Complexity Fuzzy Decision Tree Model. In: 2009 Annual IEEE India Conference. pp. 1–4. Institute of Electrical & Electronics Engineers (IEEE) (2009), <http://dx.doi.org/10.1109/INDCON.2009.5409447>
2. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern classification. John Wiley & Sons (2012)
3. Guerrero-Curienes, A., Rojo-Álvarez, J.L., Conde-Pardo, P., Landesa-Vázquez, I., Ramos-López, J., Alba-Castro, J.L.: On the Performance of Kernel Methods for Skin Color Segmentation. EURASIP Journal on Advances in Signal Processing 2009(1), 1–13 (2009), <http://dx.doi.org/10.1155/2009/856039>
4. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA data mining software. SIGKDD Explor. Newsl. 11(1), 10–18 (nov 2009), <http://dx.doi.org/10.1145/1656274.1656278>
5. Haykin, S.: Neural Networks: A Comprehensive Foundation. Prentice Hall (1999), <https://books.google.com.mx/books?id=3-1HPwAACAAJ>
6. John, G.H., Langley, P.: Estimating continuous distributions in bayesian classifiers. In: Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence. pp. 338–345. UAI'95, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1995), <http://dl.acm.org/citation.cfm?id=2074158.2074196>

7. Kakumanu, P., Makrogiannis, S., Bourbakis, N.: A survey of skin-color modeling and detection methods. *Pattern Recognition* 40(3), 1106–1122 (mar 2007), <http://dx.doi.org/10.1016/j.patcog.2006.06.010>
8. Kaur, A., Kranthi, B.: Comparison between YCbCr color space and CIELab color space for skin color segmentation. *International Journal of Applied Information Systems, IJAIS* 3(4), 30–33 (2012)
9. Phung, S.L., Chai, D., Bouzerdoum, A.: A universal and robust human skin color model using neural networks. In: *IJCNN International Joint Conference on Neural Networks. Proceedings (Cat. No.01CH37222)*. vol. 4, pp. 2844–2849. Institute of Electrical & Electronics Engineers (IEEE) (2001), <http://dx.doi.org/10.1109/IJCNN.2001.938827>
10. Santos, A., Pedrini, H.: Human Skin Segmentation Improved by Saliency Detection. In: *Computer Analysis of Images and Patterns*, pp. 146–157. Springer Science & Business Media (2015)
11. Sobieranski, A.C., Chiarella, V.F., Barreto-Alexandre, E., Linhares, R.T.F., Comunello, E., von Wangenheim, A.: Color Skin Segmentation Based on Non-linear Distance Metrics. In: *Progress in Pattern Recognition Image Analysis, Computer Vision, and Applications*, pp. 143–150. Springer Science & Business Media (2014)
12. Terrillon, J.C., Shirazi, M., Fukamachi, H., Akamatsu, S.: Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images. In: *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*. pp. 54–61. Institute of Electrical & Electronics Engineers (IEEE) (2000), <http://dx.doi.org/10.1109/AFGR.2000.840612>