Recuperador de imágenes en base a contenido: una aplicación con esporas

Meléndez Acosta Nayeli Joaquinita, Rios Figueroa Homero Vladimir y Marín Hernández Antonio

Departamento de Física e Inteligencia Artificial
Sebastián Camacho No.5 Xalapa, Veracruz,
nayelimelendez@gmail.com , hrios@uv.mx, anmarin@uv.mx

Resumen La identificación taxonómica por género y especie de los hongos micorrízicos arbusculares (HMA) tiene su fundamento en los detalles de cada estructura, como son color, textura, dimensiones y forma. En este documento, se explica el desarrollo de un sistema de recuperación de imágenes basada en contenido para ser utilizado en el área de la microbiología empleando esporas. Para el color se utilizaron los Histogramas, usando el espacio HSV, en textura se aplicaron 3 características de Tamura y en forma 4 Momentos Invariantes Afines. La base de datos contiene 225 imágenes con 60 clases, cada clase tiene de 1 a 6 esporas. 3 experimentos se realizaron uno por característica. En promedio nuestro sistema realiza un diagnostico correcto del 90 % y una predicción de clasificación del 87 %. Esta propuesta se enfoca a un área donde aun no se cuenta con una herramienta similar según expertos del área, lo que nos llevo a realizar dicho sistema.

Keywords: Hongos Micorrízicos Arbusculares (HMA), Momentos Invariantes Afines, Características de Tamura

1. Introducción

Debido al aumento excesivo en la colección de imágenes digitales, son requeridas búsquedas eficientes de imágenes, navegación y herramientas de recuperación por los usuarios de dominios diferentes [1].

Este trabajo se enfocara en la recuperación de imágenes basada en contenido (Content-based Image Retrieval o CBIR), es decir la búsqueda analizará los contenidos reales de la imagen. El término contenido se refiere a colores, formas, texturas o cualquier información que se derive de la imagen.

Los sistemas CBIR se basan en características visuales de las imágenes como el color, el brillo, las texturas y su distribución espacial (que describen el contenido de la imagen) para reconocer patrones similares en las distintas imágenes.

Los hongos micorrízicos arbusculares (HMA) son microorganismos del suelo que son simbiontes obligados de las plantas. Su estudio siempre está ligado a la microscopía, sobre todo en el campo de la taxonomía. La taxonomía clásica se basa en el estudio microscópico de las esporas y estructuras de estos hongos,

cuyas dimensiones varían entre las 300 y 20 μm. La identificación taxonómica por género y especie tiene su fundamento en los detalles de cada estructura, como son color, textura, dimensiones, forma y complejidad de la estructura [2]. Las técnicas de identificación taxonómica de los HMA requieren de un conocimiento especializado, sobretodo de la experiencia y entrenamiento en la identificación de estructuras; este proceso suele ser tedioso y requiere tiempo, algunas veces hasta días solo para separar una especie de una muestra de varios individuos. La biología molecular es una técnica que, aunque precisa, es costosa y requiere de infraestructura y un entrenamiento muy especializado del personal.

Los procesos de análisis de imágenes para clasificar objetos han demostrado ser una herramienta útil en el ahorro de tiempo dentro de los procesos de investigación científica. Las ventajas de tener software especializado que nos permita separar o clasificar no solo representan un ahorro de tiempo, sino que también permite facilitar el proceso de enseñanza a las generaciones nuevas, que pretendan estudiar los HMA y especializarse en la rama de taxonomía.

La utilidad de un software que permita clasificar e incluso identificar especies de HMA mediante el análisis de imágenes, abriría área de oportunidad dentro del área de estudio.

2. Fundamentos

Los sistemas CBIR realizan las consultas mediante una imagen de referencia, la cual contiene las características visuales que deseamos buscar, con esta imagen, el sistema realiza el procesamiento necesario para extraer el vector de características el cual no es más que un conjunto de descriptores que forman el índice de la imagen. Este vector es comparado con los vectores de características de cada una de las imágenes que conforman la base de datos. La comparación se realiza mediante una métrica o función de similitud que permite recuperar una lista de imágenes que se aproximen mejor a la consulta realizada. La Figura 1 ilustra el diagrama del proceso de consulta, el cual representa la estructura básica de un sistema CBIR [4].

Los principales componentes que constituyen un sistema CBIR son:

Interfaz de usuario: debe permitir realizar las consultas y visualizar los resultados de las mismas. Las consultas pueden ser realizadas mediante imágenes de referencia dadas al sistema. Esta permite la interacción del usuario para refinar el proceso de búsqueda a partir de la realimentación de valores para obtener el resultado deseado.

Análisis de imágenes: se encarga de realizar el procesamiento de las imágenes, su análisis y extracción de las características mediante técnicas de visión. Aquí se calcula y cuantifica la información relativa al color, a texturas o formas presentes en la imagen.

Comparación y funciones de similitud: este es un aspecto fundamental y se relaciona con la definición de los vectores de características. Las métricas o funciones de similitud permiten comparar el vector de características de la imagen de referencia con los vectores de características de las imágenes almacenados en

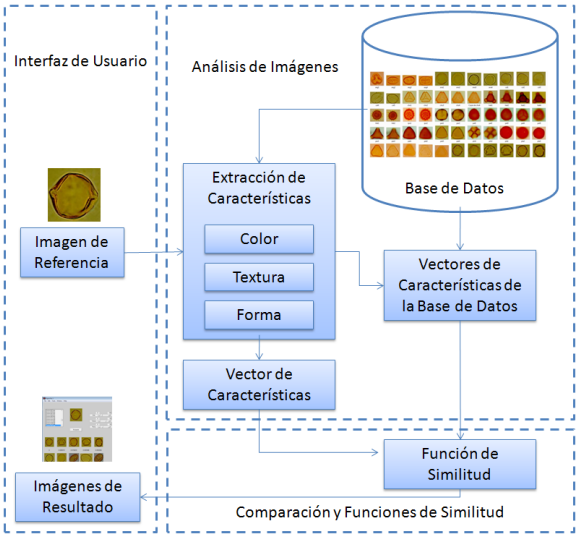

Figura 1. Esquema de una Consulta en un Sistema CBIR

la base de datos del sistema, y establecer cuáles son más cercanas o similares a la consulta dada.

3. Arquitectura del sistema

Ahora se describen los métodos utilizados en el proceso de extracción de características primitivas de la imagen, que es el aspecto clave en los sistemas CBIR.

3.1. Color

Esta características es uno de los aspectos más representativos de la imagen. Aunque la mayoría de las imágenes están en el espacio de color RGB, no es éste el más adecuado para la obtención de su caracterización cromática, ya que es el espacio de color más ajeno a la percepción del ser humano.

Histogramas de Color: Una forma de extraer esta característica de color es mediante el uso de histogramas. Los histogramas de color son una forma que representa la distribución del color en imágenes, donde cada bin del histograma representa un color en el espacio de color [5].

Espacio de Color: El espacio de color determina la manera en como representamos los diferentes colores. Existe una gran variedad de espacios de color, en este trabajo se utilizó el espacio HSV (Hue-Satuation-Value) que define los colores en base a tono (hue), cantidad de color presente (saturation) y brillo (value), esta elección se hizo en base a los resultados en [1] donde HSV es mejor que RGB en el contexto de cambios de iluminación, además en este espacio las

diferencias de color son más cercanas a las diferencias halladas en la percepción de los humanos.

Métrica de similitud en Color: Una de tantas métricas utilizadas para medir la similitud entre histogramas es la chi cuadrada [5] la cual es utilizada en este trabajo, esta métrica da un valor pequeño si las imágenes son similares y un valor grande en caso contrario. En dos imágenes idénticas da un valor de 0.

$$d_{chi\ cuadrada}(H_1, H_2) = \sum \frac{(H_1(i) - H_2(i))^2}{H_1(i) + H_2(i)} \ . \tag{1}$$

donde, H1 y H2 son los histogramas a comparar.

3.2. Textura

La textura es otra característica importante de los objetos que nos permite medir la regularidad de una imagen. La textura es definida por un conjunto de pixeles, es decir un solo pixel no define algún tipo de textura [5].

Características de Tamura: Las características de Tamura son en total 6, y sus autores encontraron que hay tres de ellas que correlacionan fuertemente con la percepción humana: coarseness (se entiende como rugosidad), contraste y direccionalidad. En su trabajo original en [6] presenta la forma de calcular las 6 características para obtener un valor escalar por cada imagen procesada.

1. *Tamura Coarseness:* El coarseness proporciona información acerca del tamaño de las texturas en la imagen. Cuando las texturas tienen un micro patrón y un macro patrón el procedimiento de Tamura considera el patrón más grande, y la forma de hacerlo es aplicando operadores de distinto tamaño de la siguiente manera:
   Sea I una imagen; I(x, y) el valor del píxel en la posición (x, y).
   a) Para cada píxel (x, y) se calcula el promedio sobre una vecindad definida que debe ser de un tamaño potencia de dos, por ejemplo $2 \times 2, 4 \times 4, ..., 32 \times 32$:

$$A(x, y) = \frac{1}{2^k} \sum_{i=1}^{2^k} \sum_{j=1}^{2^k} I(x - 2^{k-1} + i, y - 2^{k-1} + j) \tag{2}$$

   b) Para cada punto (x, y) se calcula la diferencia entre las vecindades que no se sobre solapan a lados opuestos del punto en las direcciones horizontal y vertical.

$$D_h^k(x, y) = \left| A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y) \right| \tag{3}$$

   o

$$D_h^v(x, y) = \left| A_k(x, y + 2^{k-1}) - A_k(x, y - 2^{k-1}) \right| \tag{4}$$

   c) Para cada punto (x, y) seleccionar el valor con la diferencia más grande:

$$S(x, y) = Max_{d=h,v} \left\{ E_k^d(x, y) \right\}, k = \{1 \cdots 5\} \tag{5}$$

d) Finalmente tomar el promedio 2^S como una medida de rugosidad para la imagen:

$$F_{crs} = \frac{1}{nm} \sum_{x=1}^{n} \sum_{y=1}^{m} 2^{S(x,y)} \tag{6}$$

2. *Contraste de Tamura:* El contraste está influenciado por los niveles de grises, la polarización en los niveles de blanco y negro, definición de los bordes y período de repetición de los patrones. Este se calcula utilizando la media y la varianza de los valores de intensidad en una vecindad dada de la siguiente manera:

$$F_{con} = \frac{\sigma}{\alpha_4^z} \; con \; \alpha_4^z = \frac{\mu_4}{\sigma 4} \tag{7}$$

en donde

$$\sigma 4 = \frac{1}{XY} \sum_{x=1}^{X} \sum_{y=1}^{Y} (I(x,y) - \mu)^4 \tag{8}$$

$z = 0{,}25$ es recomendado como el mejor discriminante de texturas

3. *Direccionalidad de Tamura*: La presencia de dirección en la imagen es más interesante, que la orientación. Para calcular la direccionalidad, las derivadas horizontal y vertical son calculadas al convolucionar la imagen I utilizando los siguientes operadores de 3×3:

$$\Delta_H = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \; y \; \Delta_V = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \tag{9}$$

Luego, para cada posición (x, y) se calcula el ángulo de orientación utilizando:

$$\theta = \frac{\pi}{2} + tan^{-1} \frac{\Delta_V(x,y)}{\Delta_H(x,y)} \tag{10}$$

Con la definición de las características de Tamura presentadas anteriormente, se obtienen tres medidas de textura por imagen, con los cuales puede construirse el vector de características que contiene información de las texturas, para construir un vector de Tamura.

Métrica de Similitud en Textura: La medida de similitud de textura entre dos imágenes está representada de forma cuadrática entre el vector característica de la imagen de entrada y el vector característica de una imagen de la base de datos (Distancia Euclidiana), la cual es definida como:

$$d = \sqrt{(f_{11} - f_{12})^2 + (f_{12} - f_{22})^2 + \cdots + (f_{1n} - f_{1n})^2} \tag{11}$$

donde, fij esta definida como una característica de Tamura

3.3. Forma

Para el análisis de forma existe el uso de momentos invariantes afines los cuales son invariantes a traslación, rotación, escala y transformadas afines.

Momentos Invariantes Afines: El interés de los momentos generales está en que los contornos pueden modelarse como un tipo especial de funciones acotadas y calcular los momentos de la misma, es por esto que se aplica en el reconocimiento de formas, los momentos invariantes clásicos fueron presentados por [8]. En [9] derivaron un nuevo conjunto de momentos invariantes que son invariantes bajo transformadas afines, este trabajo hace uso de 4. *Momentos Geométricos y Momentos Centrales:* La definición de los momentos geométricos 2-D de orden $(p+q)$ de una funciones de distribución de densidad $f(x,y)$ como la siguiente:

$$m_{pq} = \sum\sum x^p y^q f(x,y) \tag{12}$$

Los momentos que tiene la propiedad de traslación invariante son llamados momentos centrales y son definidos como:

$$\mu_{pq} = \sum\sum (x-\overline{x})^p (y-\overline{y})^q f(x,y) \tag{13}$$

donde $\overline{x}$ y $\overline{y}$ son las coordenadas del centroide de la función de la imagen $f(x,y)$

$$\overline{x} = \frac{m_{10}}{m_{00}}, \overline{y} = \frac{m_{01}}{m_{00}} \tag{14}$$

Momentos invariantes afines implementados:

$$Inv_1 = \frac{1}{\mu_{00}^4}(\mu_{20}\mu_{02} - \mu_{11}^2) \tag{15}$$

$$\tag{16}$$

$$Inv_2 = \frac{1}{\mu_{00}^6}(\mu_{40}\mu_{04} - 4\mu_{13}\mu_{31} + 3\mu_{22}^2) \tag{17}$$

$$Inv_3 = \frac{1}{\mu_{00}^7}(\mu_{20}\mu_{21}\mu_{30} - 4\mu_{20}\mu_{12}^2 - \mu_{11}\mu_{03}\mu_{30} + \mu_{11}\mu_{21}\mu_{12} + \mu_{02}\mu_{30}\mu_{12} - \mu_{02}\mu_{21}^2 \tag{18}$$

$$Inv_4 = \frac{1}{\mu_{00}^{10}}(\mu_{30}^2\mu_{02}^2 - 6\mu_{21}\mu_{12}\mu_{30}\mu_{03} - 4\mu_{30}\mu_{12}^3 + 4\mu_{03}\mu_{21}^3 - 3\mu_{21}^2\mu_{12}^2) \tag{19}$$

Métrica de Similitud en Forma: La métrica utilizada para comprar dos vectores de forma es la Distancia Euclidiana. Si la variación se acerca a 0 significa que son similares en caso contrario no se parecen.

3.4. Recuperación de imágenes usando color, textura y forma

El resultado de una recuperación usando más de una características puede ser mucho más eficiente, por ello uso la combinación de tres características color, textura y forma.

La similitud entre la imagen de referencia y una de la base de datos es medida usado tres características: color, textura y forma. Estas características representan diferentes aspectos por lo que es conveniente agregar un peso que diferencie su valor. La distancia entre dos imágenes fue tomada de [10] y es medida como:

$$d = w1 * d1 + w2 * d2 + w3 * d3 \quad , \quad \sum_{i=1}^{3} w_i = 1 \tag{20}$$

donde, $w1$ es el peso de la característica de color, $w2$ es el peso de la textura, $w3$ es el peso de la forma y $d1, d2$ y $d3$ son las distancias calculadas de color, textura y forma respectivamente entre las imágenes. La suma de los pesos es un promedio ponderado.

4. Resultados experimentales

Este sistema persigue permitir al usuario clasificar e incluso identificar especies de HMA mediante el análisis de imágenes. Se realizaron los experimentos haciendo uso de la base de datos de imágenes en [11]. La Figura 2 muestra algunas imágenes de la base de datos utilizada.

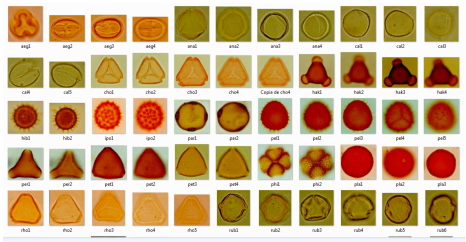

Figura 2. Base de Datos de Esporas

Las imágenes utilizadas en la experimentación contienen un solo objeto esto permite que la extracción de características sea más fácil, principalmente en forma. La base de datos tiene 225 imágenes, reparidas en 60 clases, cada clase tiene de 1 a 6 imágenes. El usuario del sistema tiene la elección de elegir el peso de cada característica.

La Figura 3 muestra los resultados de 3 experimentos, un experimento por característica: La Figura 3 (a) ejemplifica el color con un porcentaje de uso mas alto que el de textura y forma, los resultados son invariantes a color mostrando así que la característica mas importante para esta consulta es el color. En la Figura 3 (b) se puede observar claramente el buen desempeño de las Momentos Invariantes Afines, la consulta se realiza usando forma al 100 %, el resto de las características no son tomadas en cuenta, en los resultados se puede ver en la

primera fila que las esporas uno, tres, cuatro y cinco pertenecen a la misma clase y que ni color, ni textura influyen en la recuperación de la espora dos. En el último experimento mostrado en la Figura 3 (c) el porcentaje de uso de textura es el mas alto y se puede observar como que los resultados son buenos ya que predomina la textura, incluso el color se preservó.

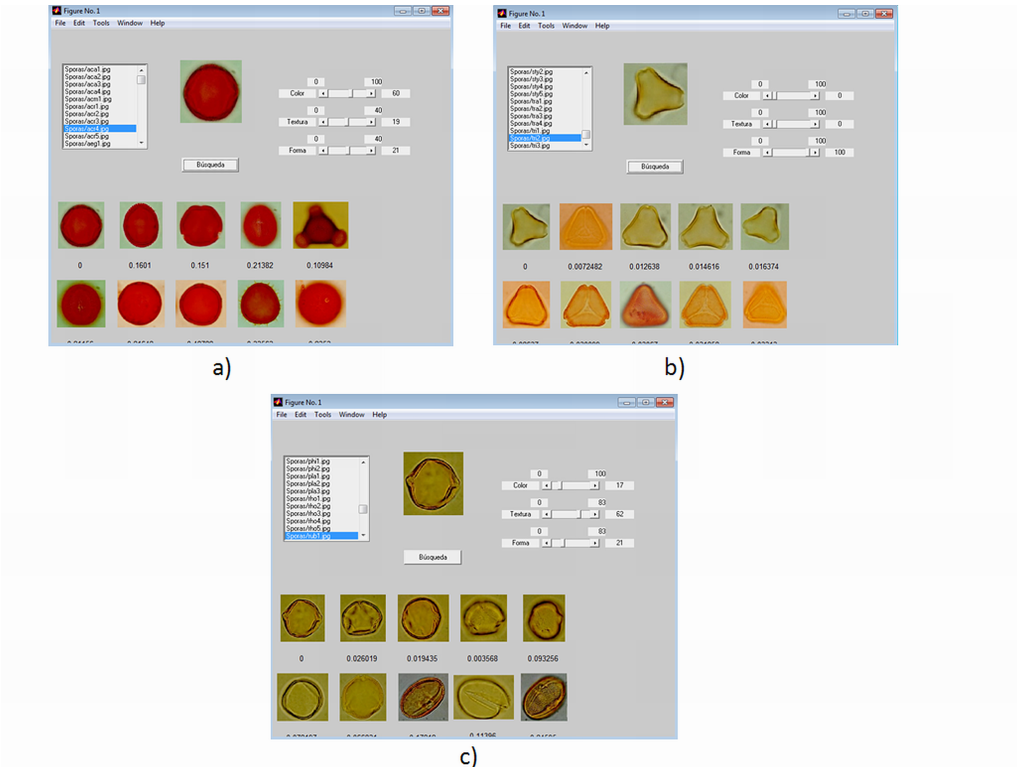

Figura 3. a)Espora=acronychia_oblongifolia, Color=60 %, Textura=19 % y Forma=21 % b) Espora=tristaniopsis_laurina, Color=0 %, Textura=0 % y Forma=100 % c) Espora= rubus_fruticosus, Color=17 %, Textura=62 % y Forma=21 %

4.1. Evaluación y comparación

Generalmente los sistemas CBIR son evaluados con relación a su rendimiento usando la observación humana. Nuestro sistema de CBIR primero es evaluado en términos de efectividad de recuperación. Para evaluar la efectividad de los sistemas de recuperación, son usadas dos métricas bien conocidas:

$$Precisión = \frac{Número\ de\ imágenes\ relevantes\ recuperadas}{Número\ total\ de\ imágenes\ recuperadas} \tag{21}$$

$$Recuperación = \frac{Número\ de\ imágenes\ relevantes\ recuperadas}{Número\ total\ de\ imágenes\ relevantes\ en\ la\ base\ de\ datos} \tag{22}$$

En la evaluación 5 expertos con conocimiento en esporas fueron jueces, los cuales valoraron los resultados y expusieron si las imágenes resultantes son similares a la imagen de referencia, la decisión de similitud es dominada por la mayoría de votos. Se tomaron en cuenta 3 de varios experimentos realizados, los cuales ya han sido mostrados en la Figura 3. El cuadro 1 muestra los resultados obtenidos en los experimentos, en promedio nuestro sistema realiza un diagnostico correcto (precisión) del 90 % y una predicción de clasificación (recuperación) del 87 %, proporcionando buenos resultados.

Cuadro 1. Evaluación de Efectividad

Experimento	**Espora**	**Precisión**	**Recuperación**
1	acronychia_oblongifolia	0.9	0.81
2	tristaniopsis_laurina	1	0.94
3	rubus_fruticosus	0.8	0.88
Promedio		0.9	0.87

En el cuadro 2 se observa una comparativa entre sistemas tomando en cuenta las características que usan: Color, Textura, Forma y Palabras Clave, y que tipo de sistema es: especifico o general. CBIR-Esporas usa tres características y es específico aunque puede ser general si la base de datos cumpla con las especificaciones requeridas, estas características hablan bien de nuestro sistema.

Cuadro 2. Comparativa de Sistemas

Sistema	**Color**	**Textura**	**Forma**	**Palabras clave**	**General**	**Especifico**
QBIC	*	*	*	*	*	
PicToSeek	*		*		*	
Photobook	*	*	*	*	*	
CIRES	*	*				*
NETRA	*	*	*		*	
Simplicity	*	*	*			*
Cortina	*	*		*	*	
SIMBA		*			*	
Alta Vista	*	*	*	*	*	
CBIR-Esporas	*	*	*		*	*

5. Conclusiones y trabajo futuro

El método propuesto usa color, textura y forma para crear un sistema de recuperación de imágenes. La recuperación de imágenes de esporas puede usarse para la educación e investigación ecológica. Usando las tres características el sistema proporciona mejores resultados. Además esta propuesta se enfoca a un área donde aun no se cuenta con una herramienta similar según expertos del área, lo que nos llevo a realizar una adaptación innovadora del conocimiento. Los resultados obtenidos en los experimentos muestran que en promedio el sistema realiza un diagnostico correcto del 90 % y una predicción de clasificación del 87 %, proporcionando buenos resultados. En un futuro, probaremos con una base de datos de imágenes de HMA que será proporcionada por el Laboratorio de Organismos Benéficos del Departamento de Agronomía de la Universidad Veracruzana. También consideraremos colocarlo en la web y será un Recuperador de Imágenes Basado en Contenido para cualquier base de datos siempre y cuando dicha base de datos cumpla con las especificaciones requeridas, cualquier usuario podrá subir su base de datos y utilizar el sistema, lo que significa que este sistema podrá ser aplicado a una gran variedad problemas similares.

Referencias

1. Ying Liu, Dengsheng Zhang, Guojun Lu,Wei-Ying Ma, .A survey of content-based image retrieval with high-level semantics"Pattern Recognition Society, accepted 28 April 2006.
2. Sieverding Ewald, Fritz Oehl, Javier Palenzuela, Kurt Ineichen, and Gladstone Alves da Silva. "Advances in Glomeromycota taxonomy and classification". IMA Fungus. 2011 December; 2(2): 191-199.
3. Yong Rui and Thomas S. Huang."Image retrieval: Current techniques, promising directions, and open issues". Journal of Visual Communication and Image Representation, 10:39-62, 1999
4. Maldonado Bautista José Orlando y Graña Romay Manuel, .Estudio de métodos de indexación y recuperación en bases de datos de imágenes", Memoria para obtener al grado de Doctor en Informática.
5. Maldonado Méndez Carolina Gabriela y Ríos Figueroa Homero, Recuperación de Información Visual". Tesis de Maestría Septiembre 2001
6. Sawhney Harpreet S. and Hafner James L."Efficient Color Histogram Indexing". IEEE 1994.
7. Tamura, H., S. Mopi, and T. Yamawaki, "Textural features corresponding to visual perception". IEEE Transactions on Systems, Man, and Cybernetics, 1978. SMC-8(6): p. 460-473.
8. Hu, M.K., 1962. "Visual pattern recognition by moments invariants". IRE Trans. Information Theory, 8: 179- 87.
9. Flusser J. and Suk T., "Pattern recognition by affine moment invariants,"Pattern Recogn., vol.26, pp.167-174, 1993.
10. Zhang Dengsheng, "Improving Image Retrieval Performance by Using Both Color and Texture Features"
11. Data Base by Peter Shimeld `ggpws@cc.newcastle.edu.au` for the Pollen Laboratory at the The University of Newcastle, Callaghan, NSW, Australia, NSW, Australia