

Sistemas de recuperación de contenido (CBMR)

Los sistemas de recuperación de contenido multimedia (**Content Based Multimedia Retrieval, CBMR**) son sistemas de software compuestos normalmente por varios programas asociados que operan bien en un ordenador local, o bien a través de la web.

Su característica esencial es que son capaces de **indexar información multimedia** usando bien etiquetas de texto asociadas a la misma, bien elementos que sirvan para describir el contenido extraídos automáticamente del propio contenido.

Una vez indexados los elementos multimedia, a los que se puede considerear como incluídos en una base de datos, el usuario realiza una **consulta** ("query") usando bien etiquetas ("tags") de texto, bien otros elementos multimedia a modo de ejemplos, o ambas cosas. El sistema debe mostrarle el número de elementos de la base de datos, **ordenados por su proximidad a la consulta**.

Sistemas de recuperación de contenido (CBMR)

Nos centraremos en la recuperación de **información visual** (imágenes o secuencias)

Existe necesidad de estos sistemas: cada vez existe más información visual disponible, raramente bien etiquetada.

¿Cuáles serán las claves utilizadas para la búsqueda?

Elementos visuales como el **color**, la **textura**, la **forma** de los objetos contenidos en la propia imagen. etc.

Sistemas de recuperación de contenido (CBMR)



Pregunta ("query")

Colección de imágenes (o una pequeña parte de ella...)



¿Qué resultado se espera?

Tecnologías Multimedia en Web

Sistemas de recuperación de contenido (CBMR) (III)

Sistemas de recuperación de contenido (CBMR)

Pero ¿en qué podríamos basar la búsqueda?

En color: hay elementos marrones y verdes en otras imágenes...

En textura: la textura de la hierba se puede confundir con la de algunos fuegos artificiales...

En forma: ¿cómo describimos la forma de una caballo (o de una flor...)?

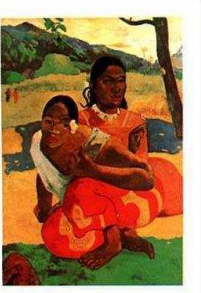





¿Y si el usuario lo que busca son "medios de transporte"?

Primera consideración:

Los sistemas CBMR están dirigidos a UN usuario concreto, y él es el único que puede decidir si el resultado de la búsqueda es satisfactorio.

Sistemas de recuperación de contenido (CBMR)

¿Qué información existe en una imagen?

				<input type="text" value="Autor"/> <input type="text" value="Título"/> <input type="text" value="Fecha"/> <input type="text" value="Museo"/>
		<input type="text" value="Autor"/> <input type="text" value="Fecha"/> <input type="text" value="Lugar"/>		

Tecnologías Multimedia en Web

Sistemas de recuperación de contenido (CBMR) (v)

Sistemas de recuperación de contenido (CBMR)

Veamos, antes de seguir, qué tipo de información nos puede proporcionar una imagen:

- Datos independientes del contenido: formato, nombre del autor, fecha, propietario,...
- Datos contenidos en la propia imagen:
 - Características de bajo nivel: textura, color, forma, estructura, movimiento,...
 - Datos dependientes del contenido semántico: datos descriptivos del contenido (objetos mostrados, emociones y significados asociados a la imagen,...).

Sistemas de recuperación de contenido (CBMR)

Primera generación de Sistemas de Recuperación Visual (SRV):

- Estos sistemas permiten acceder a la información visual a través de atributos de texto.
- Los datos independientes de la imagen se representan como cadenas de caracteres.
- Los datos dependientes del contenido se expresan a través de palabras claves o descripciones más extensas.
- Se utilizan los tradicionales lenguajes de búsqueda como SQL para realizar las búsquedas.

Problemas de esta representación:

- Deficiencia para expresar ciertas características perceptuales.
- No consigue representar bien la similitud perceptual.
- Depende directamente del 'anotador'.

La primera generación de sistemas se basaban en los tradicionales sistemas de BDs; por ejemplo, se utilizaba el modelo relacional para almacenar información relevante de cada imagen.

Sistemas de recuperación de contenido (CBMR)

Nueva generación de SRV:

Estos sistemas utilizan:

- Descriptores objetivos del contenido visual.
- Modelos apropiados de similitud.

¿Qué tipo de información podemos extraer de una imagen?

- Propiedades perceptuales de bajo nivel (color, textura, forma y relaciones espaciales).
- Primitivas correspondientes a la abstracción, como objetos, papeles y escenas.
- Impresiones, emociones y significado asociado a la combinación de características perceptuales.

Así, se necesitará un módulo de procesamiento de imágenes y reconocimiento de patrones como partes integrantes del sistema total.

La nueva generación de SRV intenta solucionar los problemas anteriores. Para ello cambiará las palabras claves que representan el contenido por información objetiva del contenido, y necesitará trabajar con modelos de similitud apropiados.

Sistemas de recuperación de contenido (CBMR)

¿Cómo funcionan los SRV?

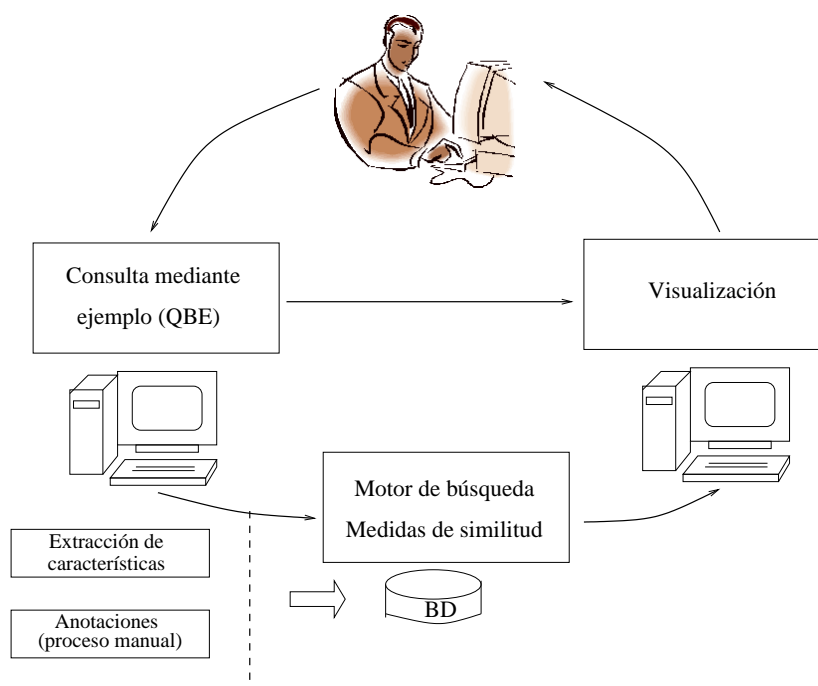
- Se calcula un conjunto de descriptores para cada imagen en la BD.
- Las consultas se expresan mediante ejemplos visuales. ("Query By Example", QBE)
- El sistema comprueba el parecido entre el contenido visual de la consulta y las imágenes en la BD (medida de similitud)

No se trata de un encaje. Las diferencias entre búsqueda por similitud y encaje ("matching"):

- El encaje decide si un objeto observado se corresponde o no con un modelo mientras que la búsqueda por similitud es una tarea de reordenación de las imágenes en la BD.
- En el encaje la imprecisión se maneja dentro del proceso mientras que en la búsqueda por similitud es el usuario quien hace este papel.

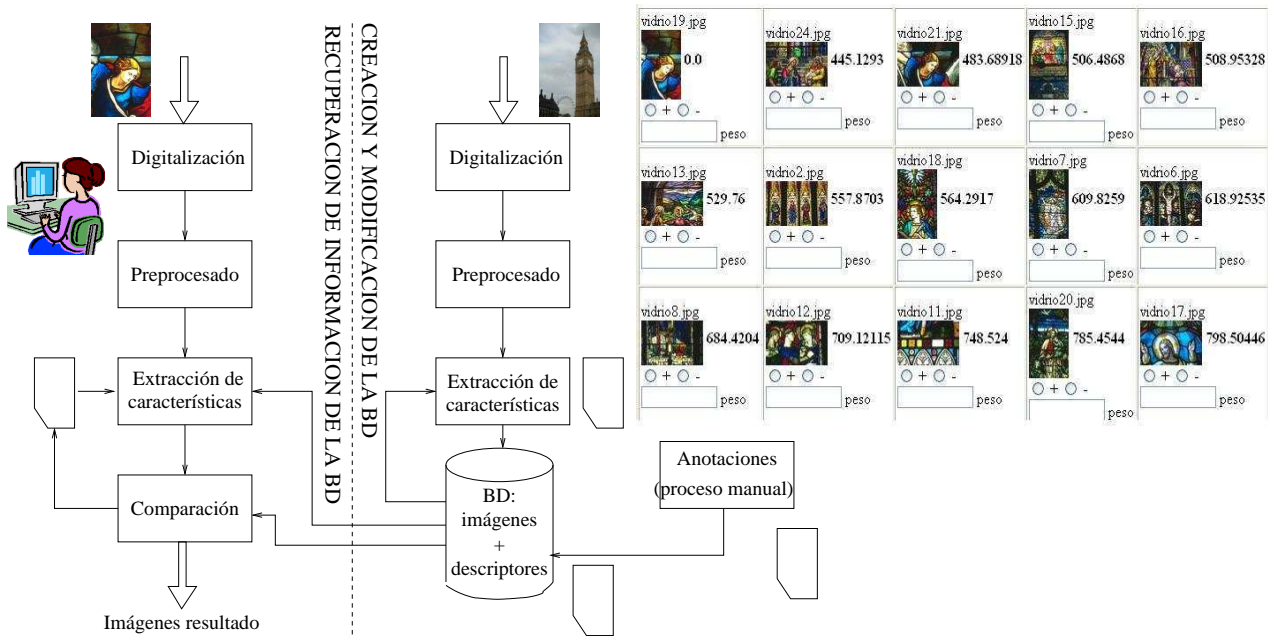
Sistemas de recuperación de contenido (CBMR)

Esquema de uso de un SRV



Sistemas de recuperación de contenido (CBMR)

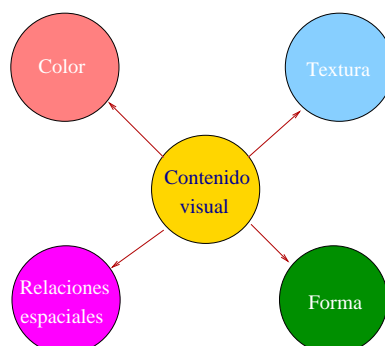
Esquema de construcción de un SRV



Sistemas de recuperación de contenido (CBMR)

Un problema clave en estos sistemas es la selección de características visuales de las imágenes en la BD. Se necesitan métodos para extraer descriptores numéricos adecuados que den una medida cuantitativa de estas características; estos métodos proceden del análisis de imágenes y reconocimiento de patrones.

Algunas primitivas semánticas como objetos, acciones, sucesos se pueden derivar por análisis de combinaciones de características de bajo nivel según un modelo determinado. En cualquier caso el problema clave es la selección de buenas características que describan el contenido visual.



Sistemas de recuperación de contenido (CBMR)

Descriptores de color: están relacionados con los atributos cromáticos de las imágenes. Desde un punto de vista físico, la percepción del color deriva de la distribución del espectro de la radiación electromagnética que llega a la retina, pero desde el punto de vista psicológico, la percepción de un color depende de varios factores: atributos del color, colores del entorno, organización espacial del color, la memoria/conocimiento del observador.

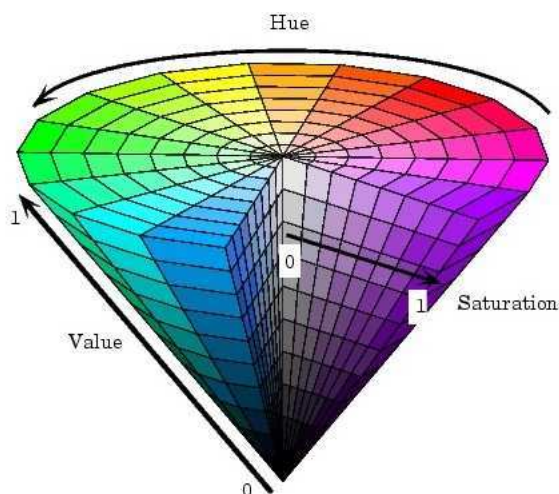
Las cualidades cromáticas del color se representan mediante modelos geométricos de color

Los espacios de color se pueden clasificar en

- Espacios de color no uniformes
 - RGB (orientado al dispositivo)
 - HSV (Hue, Saturation, Value). (muy intuitivo)
- Espacios de color perceptualmente uniformes
 - Lab
 - Luv: parece aproximar las diferencias de color como las percibe el ojo humano

Sistemas de recuperación de contenido (CBMR)

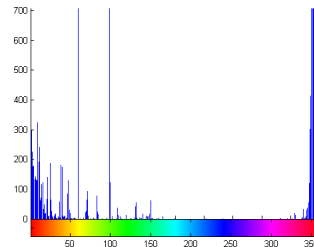
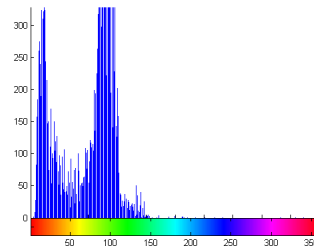
El espacio de color más usado en este contexto es el HSV:



La característica más relevante para distinguir un color de otro, independientemente de la iluminación, es la componente H (Hue-tinte).

Sistemas de recuperación de contenido (CBMR)

Una forma de resumir la información cromática sin tener en cuenta la intensidad o la saturación del color, que pueden variar entre unas imágenes similares y otras, es el histograma de la componente H:



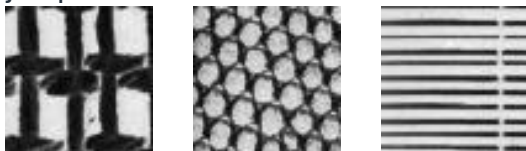
Sistemas de recuperación de contenido (CBMR)

Descriptores de texturas:

El término textura se utiliza en reconocimiento de patrones para identificar trozos de imágenes que se caracterizan por cambios de intensidad. Tiene que ver con altas frecuencias del espectro de la imagen (imágenes suaves no se consideran texturas). Se necesita una escala de referencia a la hora de analizar una textura.

Términos utilizados para su caracterización:

- Granularidad (fina o gruesa)
- Direcciones privilegiadas,
- Periodicidad (repetición).
- Ejemplos de texturas:



- Clases de descriptores de texturas.
 - Basados en modelos estadísticos
 - Basados en transformadas.

Sistemas de recuperación de contenido (CBMR)

Descriptores de forma

Para su extracción se puede emplear:

- Segmentación de los objetos. (Mayor dificultad)
- Representación de las formas de los objetos una vez segmentados (basados en el contorno y/o en las regiones).

Entre los descriptores de forma más utilizados están el área de la región, máximo y mínimo rectángulo/círculo/elipse que contiene/está contenido en la forma, la compacidad, elongación de la región, momentos de orden p-q, coeficientes de la DFT del contorno.

- Es importante que los descriptores utilizados sean invariantes a rotación, traslación y cambio de escala.
- Se necesitan medidas de similitud que nos permitan comparar las formas de los objetos.
- Algunos de los sistemas asumen que se tienen formas segmentadas correctamente en la BD.

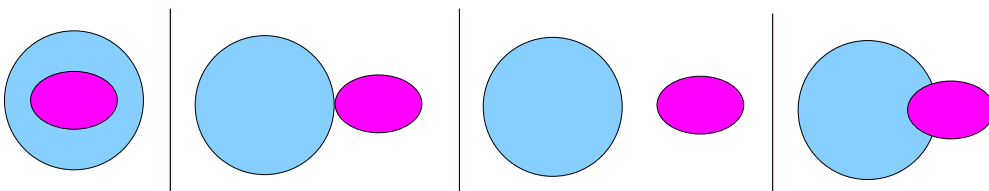
Sistemas de recuperación de contenido (CBMR)

Descriptores de relaciones espaciales:

Las entidades espaciales son puntos, líneas, regiones y objetos.

Estos descriptores capturan las relaciones existentes entre objetos segmentados u otras estructuras que los representan. Las relaciones espaciales se pueden clasificar en

- Relaciones direccionales: se refieren a direcciones relativas (arriba, abajo, derecha, izquierda). Se necesita definir una orientación para calcularlas. Otros atributos relacionados son distancias y ángulos.
- Relaciones topológicas: capturan conceptos de conjunto entre objetos cercanos como inclusión, adyacencia, disyunción, intersección. En las imágenes aparecen objetos que ocupan distintas posiciones en la imagen. Se puede extraer información de cómo se relacionan unos elementos con otros. Las entidades espaciales que normalmente se manejan son los puntos, las líneas, regiones, y objetos

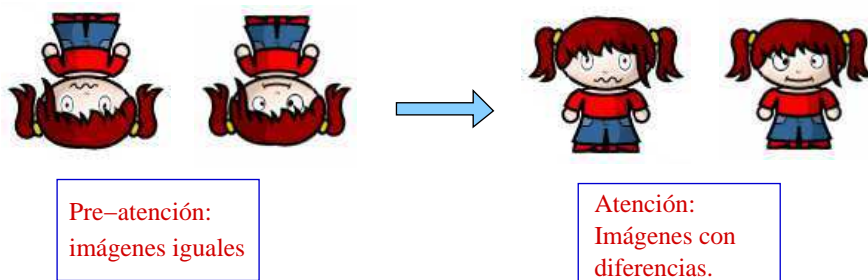


Sistemas de recuperación de contenido (CBMR)

Medidas de similitud (I)

La percepción humana trabaja a dos niveles de semejanza:

- Semejanza de pre-atención: es la similitud entre los estímulos percibidos sin ningún tipo de interpretación. Se utiliza en sistemas donde la recuperación se basa en características como el color, la textura, las relaciones espaciales, etc.
- Semejanza de atención: requiere conocimiento previo y alguna forma de razonamiento. Se utiliza en sistemas como BD de rostros, aplicaciones médicas donde es muy importante la experiencia previa.



Sistemas de recuperación de contenido (CBMR)

Medidas de similitud (II)

Medidas entre distancias comúnmente utilizadas:

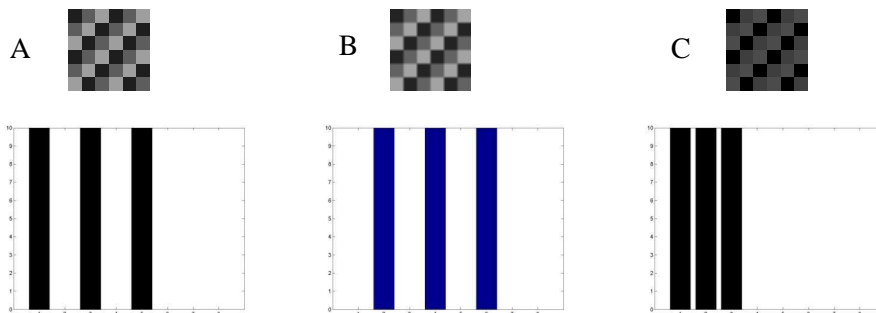
- Distancias entre vectores de características ($\mathbf{v}^1 = \{v_1^1 \dots v_n^1\}$, $\mathbf{v}^2 = \{v_1^2 \dots v_n^2\}$)
 - Distancia euclídea (o distancia euclídea ponderada).
$$d(\mathbf{v}^1, \mathbf{v}^2) = \sqrt{\sum_{i=1}^n w_i (v_i^1 - v_i^2)^2}$$
 - Distancia de Mahalanobis: se utiliza como métrica la inversa de la matriz de covarianza. $d(\mathbf{v}^1, \mathbf{v}^2) = \mathbf{v}^1 C^{-1} \mathbf{v}^2$
- Distancias entre distribuciones (normalizadas) $h = (h_1 \dots h_n)$, $k = (k_1 \dots k_n)$
 - Distancia L1: $d_1(h, k) = \sum_{i=1}^n |h_i - k_i|$
 - Distancia L2: $d_2(h, k) = \sqrt{\sum_{i=1}^n (h_i - k_i)^2}$
 - Intersección entre histogramas: $d_I(h, k) = 1 - \sum_{i=1}^n \min(h_i, k_i)$
 - Divergencia de Kullback-Leibler: $d_{KL}(k, l) = - \sum_{i=1}^n h_i \log \frac{h_i}{k_i}$
 - Divergencia de Jeffrey: $d_J(k, l) = \sum_{i=1}^n h_i \log \frac{h_i}{m_i} + k_i \log \frac{k_i}{m_i}$, $m_i = \frac{h_i + k_i}{2}$

Sistemas de recuperación de contenido (CBMR)

La divergencia de Kullback-Leibler es quizá la más usada. Extraída de la teoría de la información, no es propiamente una distancia por no cumplir la propiedad de simetría ni la desigualdad triangular. Mide cuán ineficiente sería codificar una distribución utilizando la otra; se llama también entropía relativa.

Problemas de estas medidas (intervalo a intervalo):

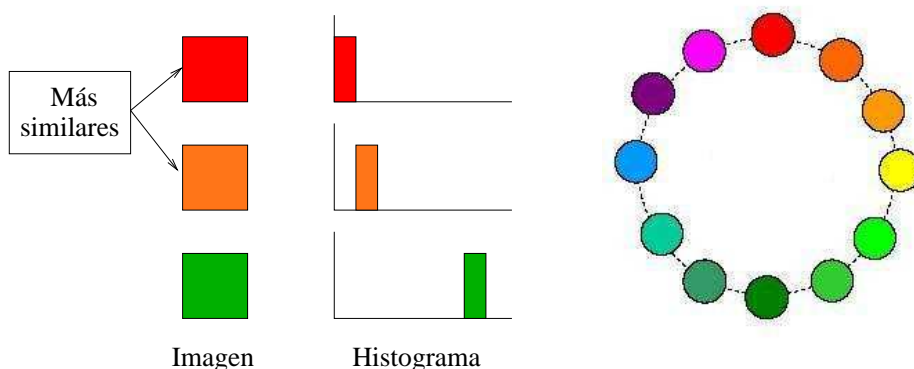
- No se ajustan a la percepción humana, solo tienen en cuenta diferencias elemento a elemento, no entre elementos cruzados.
- Son muy sensibles al tamaño de los intervalos.



A y B se parecen mucho más entre sí que A y C, pero la distancia $dL1(H_A, H_B) > dL1(H_A, H_C)$.

Sistemas de recuperación de contenido (CBMR)

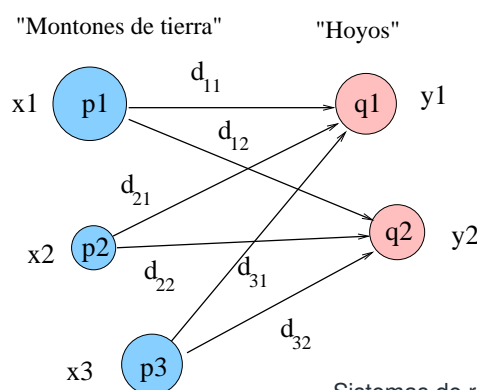
Para evitar esto se puede usar una forma cuadrática entre los histogramas que tenga en cuenta las distancias entre cada par de intervalos ('bins'):



$$d(h, k) = \sqrt{(h - k)^T A (h - k)}$$

Sistemas de recuperación de contenido (CBMR)

También se evita usando la llamada "distancia de movimiento de tierras" ("earth moving distance"). Consiste en representar cada histograma por un conjunto de pares $S = \{(x_i, p_i)\}$ donde x_i es el centro de cada bin con valor no nulo y p_i es dicho valor. Un grafo relaciona cada par (nodo) del primer histograma con cada uno del segundo mediante un arco que contiene el coste de mover una unidad de valor entre uno y otro (d_{ij}). Si vemos el valor en el primer histograma como cantidad de tierra, y en el segundo como capacidad del hoyo, se trata de nivelar al máximo el terreno con el mínimo coste de movimiento de material. Esto es un problema que tiene solución usando técnicas de programación dinámica.



Sistemas de recuperación de contenido (CBMR)

Para acceder a la BD de características no es posible hacer un acceso secuencial para comparar con cada una de las imágenes almacenadas, de modo que es necesario indexar. Sobre la indexación de los atributos de texto, se utilizan los métodos tradicionales de indexación.

Para la indexación de los atributos visuales se usan:

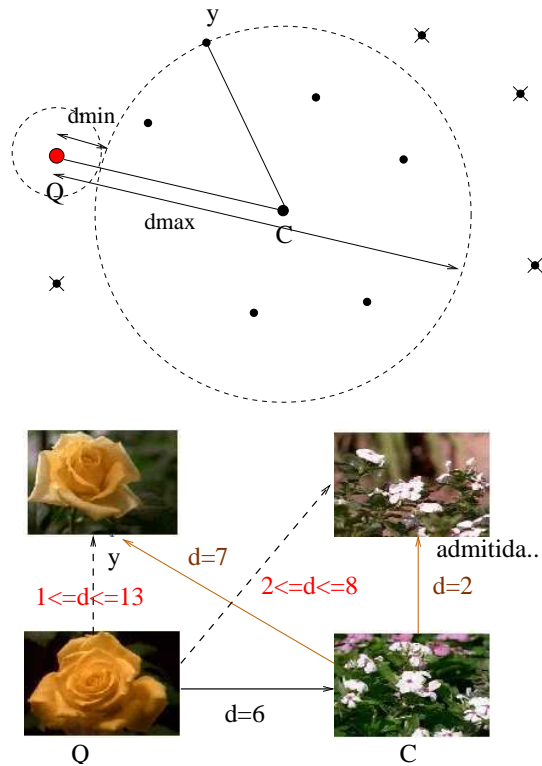
- Desigualdad triangular (método de los pivotes): se crea un índice sobre un campo creado que almacena la distancia de las imágenes de la base de datos con una imagen de referencia también almacenada en la propia BD.
- Reticula fija: es un índice no jerárquico que organiza el espacio de características en rectángulos n-dimensionales.
- Árboles k-d: son árboles binarios, en los que el valor de una de las k características se comprueba en cada nivel.
- Árboles R: Los nodos hoja contienen un número dado de puntos caracterizados por las coordenadas del mínimo rectángulo que los contiene. Los nodos superiores se caracterizan por el mínimo rectángulo que contiene todos los rectángulos de sus nodos hijos. Los rectángulos pueden solaparse.

Sistemas de recuperación de contenido (CBMR)

Desigualdad triangular:

$$d(Q, y) \geq |d(Q, C) - d(y, C)|$$

Q es la imagen "query", y C una imagen clave, cuyas distancias a todas las de la BD hemos precalculado. Calculando sólo $d(Q, C)$ y sabiendo $d(y, C)$ tenemos límites inferior y superior para la distancia buscada, que permiten descartar muchos puntos. Sin embargo, se admiten muchos otros que pueden estar muy alejados...

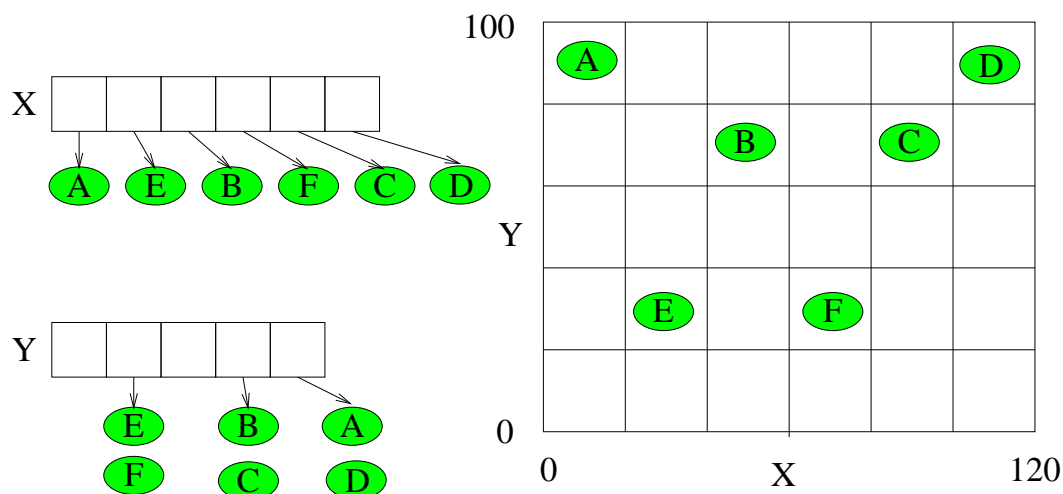


Tecnologías Multimedia en Web

Sistemas de recuperación de contenido (CBMR) (xxv)

Sistemas de recuperación de contenido (CBMR)

Retícula fija:

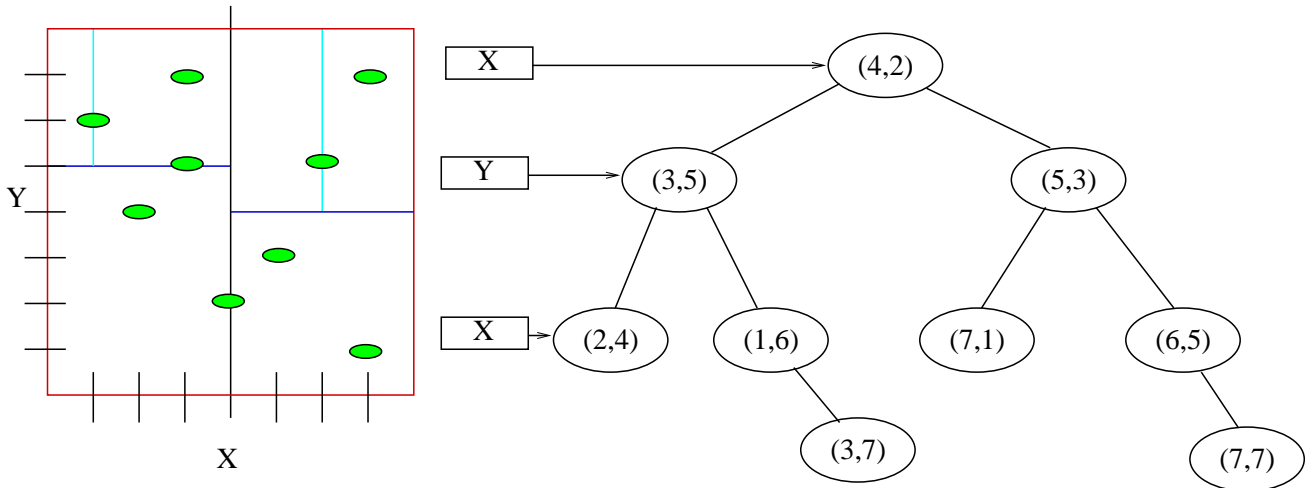


Tecnologías Multimedia en Web

Sistemas de recuperación de contenido (CBMR) (xxvi)

Sistemas de recuperación de contenido (CBMR)

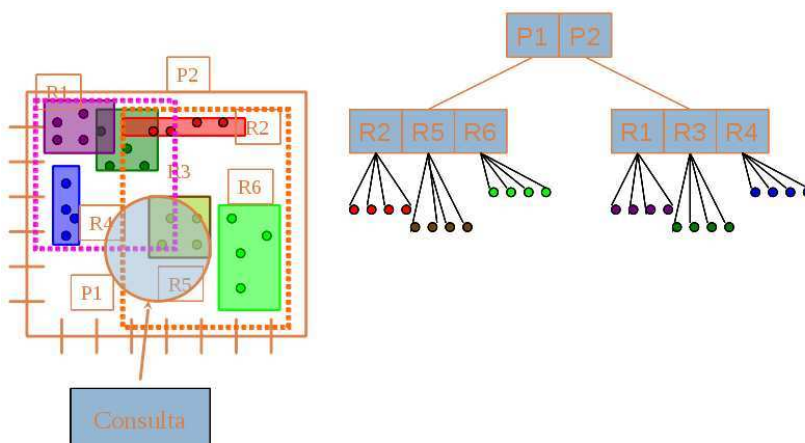
Kd-tree:



En cada paso de división se elige el pivote como la mediana en esa dimensión.

Sistemas de recuperación de contenido (CBMR)

Arboles R:



Los parámetros que usa son el **cubrimiento** (área recubierta por todos los rectángulos en los nodos hoja) y el **solapamiento** (intersección entre dos o más nodos hoja). En la creación del árbol se intenta minimizar simultáneamente estos dos parámetros.

Sistemas de recuperación de contenido (CBMR)

Un árbol R debe cumplir las siguientes condiciones:

- Cada nodo hoja contiene entre m y M puntos, a menos que sea la raíz.
- Cada nodo no hoja tiene entre m y M hijos, a menos que sea la raíz.
- Por cada entrada en el nodo no hoja (t,r) , r es el menor rectángulo que contiene todos los rectángulos en el subárbol t .
- Todos los nodos hoja están balanceados.

Sistemas de recuperación de contenido (CBMR)

Un problema muy importante de todos estos métodos de indexado es que no funcionan bien en espacios de características de dimensionalidad muy alta, tanto es así que se le conoce como "the curse of dimensionality" (la maldición de la dimensionalidad).

Para paliarlo se emplean métodos de reducción del número de dimensiones:

- Análisis de componentes principales: se utiliza el conjunto de vectores propios de la matriz de covarianza de los datos, y se mantienen únicamente aquellas componentes de mayor valores propios. Problema: inserción de nuevas imágenes en la BD
- Escalado multidimensional: se minimizan las diferencias entre la matriz N_{ij} que almacena las distancias entre puntos en el espacio de dimensión n , y la matriz distancias K_{ij} en el espacio de dimensión $k < n$. Esto proporciona un método iterativo de calcular las componentes del vector en el espacio k -dimensional.

Sistemas de recuperación de contenido (CBMR)

Un aspecto muy importante de los sistemas CBIR es la evaluación de su rendimiento. Para ello se usan los valores de la siguiente tabla:

Resultados/BD	Relevantes	No relevantes
Recuperadas	A (correctas)	B (falsas)
No recuperadas	C (perdidas)	D (rechazadas)

que permite definir los dos índices fundamentales:

- Recall: $R = \frac{A}{A+C}$ (proporción de imágenes relevantes que han sido recuperadas)
- Precision: $P = \frac{A}{A+B}$ (proporción de imágenes recuperadas que son correctas)

Sistemas de recuperación de contenido (CBMR)

Aun con todos estos procedimientos, es muy difícil que ningún algoritmo de recuperación genere un resultado con valores altos de los índices R y P. La razón está en que todos ellos operan con características de bajo nivel (color, textura,..) muy alejadas de la intención del usuario, que puede contener un significado muy complejo ("imágenes que representen medios de transporte terrestre").

A esta separación se le llama "**semantic gap**" (brecha semántica), y el medio principal para salvarla es recurrir a la interacción con el usuario mediante el uso de **realimentación**.

Sistemas de recuperación de contenido (CBMR)

Realimentación:

- El sistema responde a una consulta con un conjunto de imágenes resultado.
- La idea de la realimentación es que el usuario pueda indicar la relevancia de las imágenes resultado para mejorar el resultado de forma iterativa.
- El objetivo es llegar a mostrar las imágenes que un determinado usuario quiere obtener.
- Posibles métodos de incorporar 'relevance feedback' en una consulta:
 - El usuario marca aquellas imágenes que considera relevantes $\{x_1, \dots, x_n\}$ dándoles un peso, $\{w_1, \dots, w_n\}$, $\sum_{i=1}^n w_i = 1$
 - Se utilizan esas imágenes como consultas y se reordena el conjunto resultado de acuerdo a la siguiente medida:

$$s_j = \sum_{i=1}^n w_i D(x_i, y_j)$$

Sistemas de recuperación de contenido (CBMR)

Un ejemplo de recuperación con un interface de usuario típico sería:

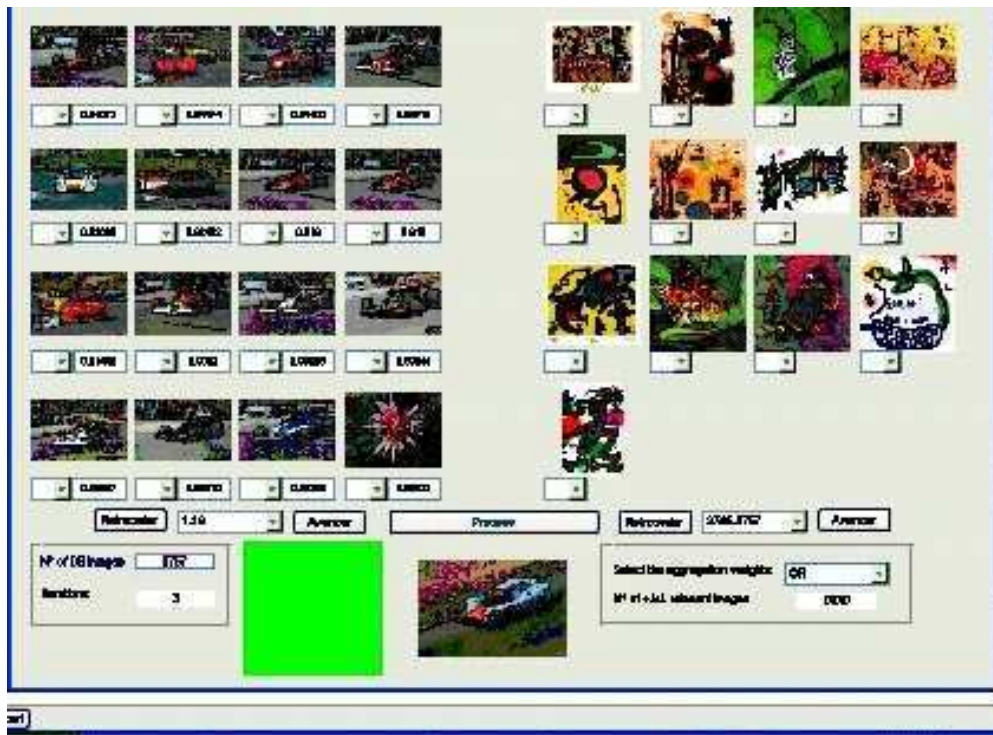
Imágenes patrón



39 Imágenes recuperadas

 flor39.jpg 0.0 <input checked="" type="radio"/> + <input type="radio"/> - 1.0 peso	 flor57.jpg 264.51944 <input type="radio"/> + <input type="radio"/> - peso	 flor58.jpg 268.73535 <input type="radio"/> + <input type="radio"/> - peso	 flor8.jpg 271.39478 <input type="radio"/> + <input type="radio"/> - peso	 flor50.jpg 276.85965 <input type="radio"/> + <input type="radio"/> - peso
 flor75.jpg 277.17743 <input checked="" type="radio"/> + <input type="radio"/> - 0.9 peso	 vidrio14.jpg 284.19867 <input type="radio"/> + <input type="radio"/> - peso	 vidrio18.jpg 286.1201 <input type="radio"/> + <input type="radio"/> - peso	 vidrio9.jpg 295.39938 <input type="radio"/> + <input type="radio"/> - peso	 vidrio3.jpg 295.39938 <input type="radio"/> + <input type="radio"/> - peso
 flor51.jpg 296.43787 <input type="radio"/> + <input type="radio"/> - peso	 flor35.jpg 305.0345 <input checked="" type="radio"/> + <input type="radio"/> - 0.9 peso	 flor17.jpg 337.80893 <input type="radio"/> + <input type="radio"/> - peso	 flor23.jpg 339.03958 <input type="radio"/> + <input type="radio"/> - peso	 flor77.jpg 350.6686 <input type="radio"/> + <input type="radio"/> - peso

Sistemas de recuperación de contenido (CBMR)



Sistemas de recuperación de contenido (CBMR)

Combinación de imágenes consulta: Los sistemas también pueden permitir especificar una consulta con varios imágenes ejemplo.

En ese caso el sistema podría utilizar operaciones de conjunto para devolver...

- Las imágenes que pertenecen a la intersección de los resultados para cada imagen.
- Las imágenes que pertenecen a la unión de los resultados para cada imagen consulta.

Se ordena el conjunto resultado de acuerdo a la suma de las distancias ponderadas a las imágenes consulta (de forma similar a como se realiza en la realimentación).

Imágenes patrón:				
				
flor39.jpg	flor75.jpg	flor35.jpg	flor39.jpg	flor23.jpg
152 Imágenes recuperadas				
				
flor75.jpg	flor35.jpg	flor39.jpg	flor23.jpg	vidrio18.jpg
213.18439	224.12315	229.11143	320.46616	338.3746
				
flor39.jpg	flor75.jpg	flor35.jpg	flor23.jpg	vidrio18.jpg
346.5806	349.88965	368.9118	375.82385	378.21704
				
flor39.jpg	flor75.jpg	flor35.jpg	flor23.jpg	vidrio18.jpg
380.14005	381.4858	381.4858	383.12518	386.97354

Sistemas de recuperación de contenido (CBMR)

Recuperación de imágenes por **color**

Ya vimos que los histogramas de color eran la manera habitual de describir el color en una imagen.

El histograma captura la distribución de colores en toda imagen o en una región de ella.

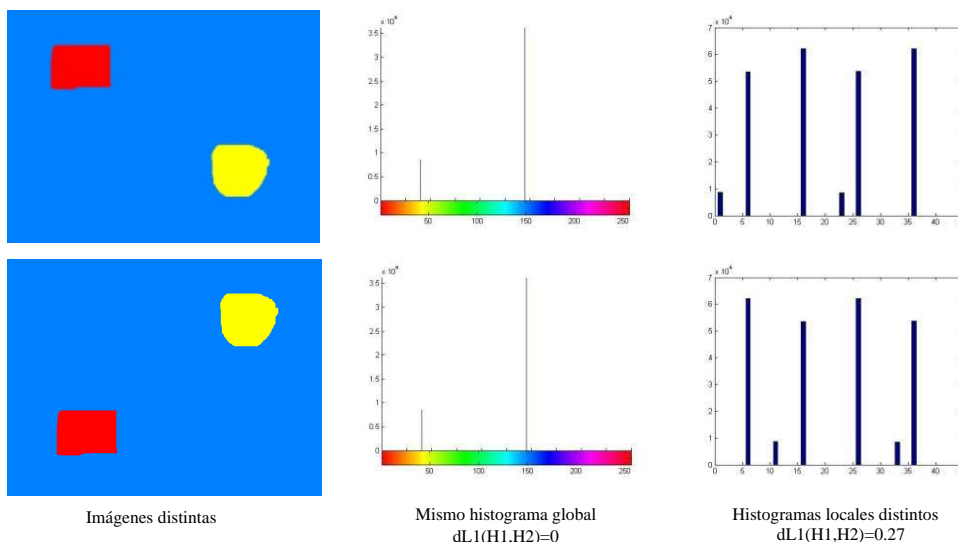
Para que se pueda trabajar con histogramas se debe reducir el número de colores de la imagen (cuantización del color).

Las decisiones a tomar son:

- El espacio de color en el que se calculará el histograma
- La cuantización: reducción del número de colores con los que se representarán las características cromáticas.
 - Cuantización uniforme.
 - Cuantización no uniforme

Sistemas de recuperación de contenido (CBMR)

El problema de trabajar con el histograma de color global es que se pierde la información de localización espacial de los colores. Para evitarlo se usan histogramas locales (p. ej, de cada cuadrado de una retícula). Esto permite buscar imágenes con colores localizados.



Los histogramas de cada subimagen se concatenan en un solo vector.

Sistemas de recuperación de contenido (CBMR)

Las consultas basadas en color pueden implicar alguna de estas cosas:

- Encontrar imágenes con una determinada proporción de color (consulta más sencilla).
- Encontrar imágenes con colores similares a los contenidos en una imagen ejemplo.
- Encontrar imágenes que contienen regiones de un color similar a alguna de las regiones de la consulta.
- Encontrar imágenes que contienen un objeto conocido basado en sus propiedades de color.

En las siguientes transparencias se muestran consultas de color local y global obtenidas con un sistema experimental llamado Virage.

Sistemas de recuperación de contenido (CBMR)

Recuperación por color global:



Sistemas de recuperación de contenido (CBMR)

Recuperación por color local:



Tecnologías Multimedia en Web

Sistemas de recuperación de contenido (CBMR) (xLI)

Sistemas de recuperación de contenido (CBMR)

Otro ejemplo de recuperación por color global:



Tecnologías Multimedia en Web

Sistemas de recuperación de contenido (CBMR) (xLII)

Sistemas de recuperación de contenido (CBMR)

... y su correspondiente ejemplo de color local.



Sistemas de recuperación de contenido (CBMR)

Recuperación de imágenes por **textura**

Es casi imposible describir texturas con palabras. Una buena representación puede basarse en propiedades estadísticas y estructurales de los patrones que aparecen en una imagen. Algunos descriptores ampliamente utilizados en la literatura son:

- La función de autocorrelación. Una textura se puede representar teniendo en cuenta el tamaño de las primitivas en niveles de gris. La función de autocorrelación tomará valores mayores para texturas gruesas que finas.

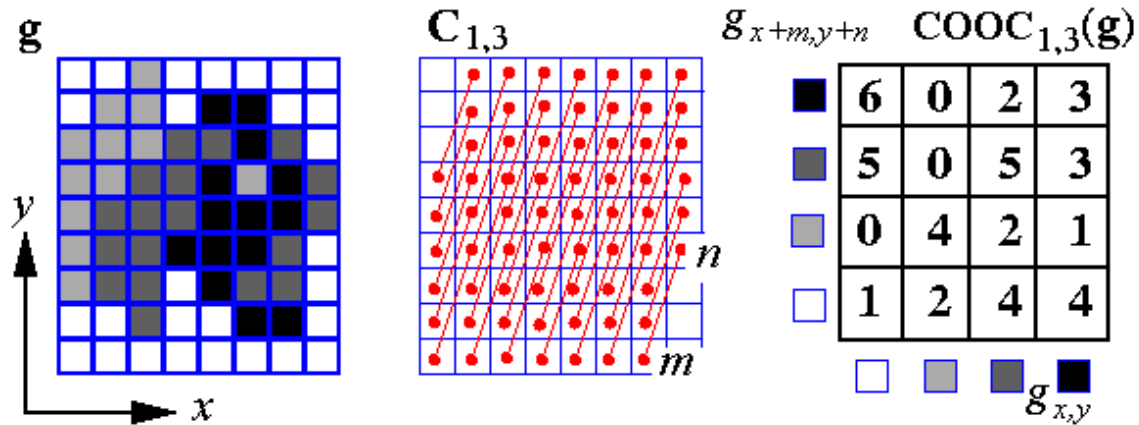
$$C_f(p, q) = \frac{MN}{(M-p)(N-q)} \frac{\sum_{i=1}^{M-p} \sum_{j=1}^{N-q} f(i, j) f(i+p, j+q)}{\sum_{i=1}^{M-p} f(i, j)^2}$$

- La matriz de co-ocurrencias

$$P(i, j) = \frac{\#\{(p_1, p_2) \in I \mid p_1 = i, p_2 = j\}}{\#I}$$

Sistemas de recuperación de contenido (CBMR)

Una gráfica que ayuda a entender el significado de la matriz de coocurrencias es:



$$P(i, j) = \frac{\#\{(p_1, p_2) \in I \mid p_1 = i, p_2 = j\}}{\#I}$$

Sistemas de recuperación de contenido (CBMR)

La matriz de coocurrencias no suele utilizarse tal cual como descriptor de la textura; en lugar de esto se hallan ciertos parámetros estadísticos de ella, entre los que destacan:

- Energía: $\sum_{i,j} P(i, j)^2$
- Entropía: $\sum_{i,j} P(i, j) \log P(i, j)$
- Contraste: $\sum_{i,j} (i - j)^2 P(i, j)$
- Homogeneidad: $\sum_{i,j} \frac{P(i, j)}{1 + |i - j|}$

Otros descriptores de textura muy usados son debidos a Tamura:

- Contraste: depende de la intensidad y de su distribución en la imagen. σ es la desviación típica de los niveles de gris, y μ_1 su primer momento (media)
- Direccionalidad: mide lo concentrada que está la textura, entendida como conjunto de microbordes, en torno a una dirección. Δ_x y Δ_y son las componentes del gradiente.

$$C_T = \frac{\sigma}{(\mu_1 / \sigma^4)^{\frac{1}{4}}} \quad D_T = \arctan \left(\frac{\sum_{i,j} \Delta_x(i, j)}{\sum_{i,j} \Delta_y(i, j)} \right) + \frac{\pi}{2}$$

Sistemas de recuperación de contenido (CBMR)

Otros descriptores de textura son los que usan el concepto de **rugosidad**

La rugosidad relaciona las distancias entre cambios espaciales notables en el nivel de gris, esto es, tiene que ver con el tamaño de las primitivas que forman la textura. Para calcularla:

- Para cada pixel (x, y) se calculan 6 promedios de los niveles de gris para ventanas de tamaño $2k \times 2k$, $k = 0, \dots, 5$ centradas en el pixel, y se calculan las diferencias absolutas $E_k(x, y)$ entre pares de promedios no solapados en las direcciones verticales y horizontales:

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

$$E_{k,h}(x, y) = | A_k(x + 2^{k-1}, y) - A_k(x - 2^{k-1}, y) |$$

- Luego se obtiene el valor de k que maximiza las diferencias $E_k(x, y)$, sea k_0 .
- La rugosidad global es el promedio de los niveles de gris de las ventanas de tamaño $2k_0$ centradas en cada pixel.

Sistemas de recuperación de contenido (CBMR)

Existen otros grupos de descriptores de textura que usan conceptos matemáticos que no podemos explicar en detalle:

- Métodos basados en el espectro de potencia: analizan el espectro de frecuencias de una imagen. Las texturas gruesas tendrán la energía espectral concentrada en las bajas frecuencias y las texturas finas en las altas frecuencias.
- Coeficientes de la transformada wavelet: esta transformada usa simultáneamente las frecuencias espaciales, y la localización de dichas frecuencias, es decir, cuán importantes son la bajas (o las altas) frecuencias, y dónde están en la imagen.
- Granulometrías: usan operaciones de morfología matemática, en concreto la apertura, que es una operación que elimina todas las partículas menores que un tamaño pedido. Se realizan aperturas con varios tamaños y se calcula el área de la imagen resultante.

Sistemas de recuperación de contenido (CBMR)

Resultados de recuperación sin realimentación del usuario, usando sólo color global (ej. 1)



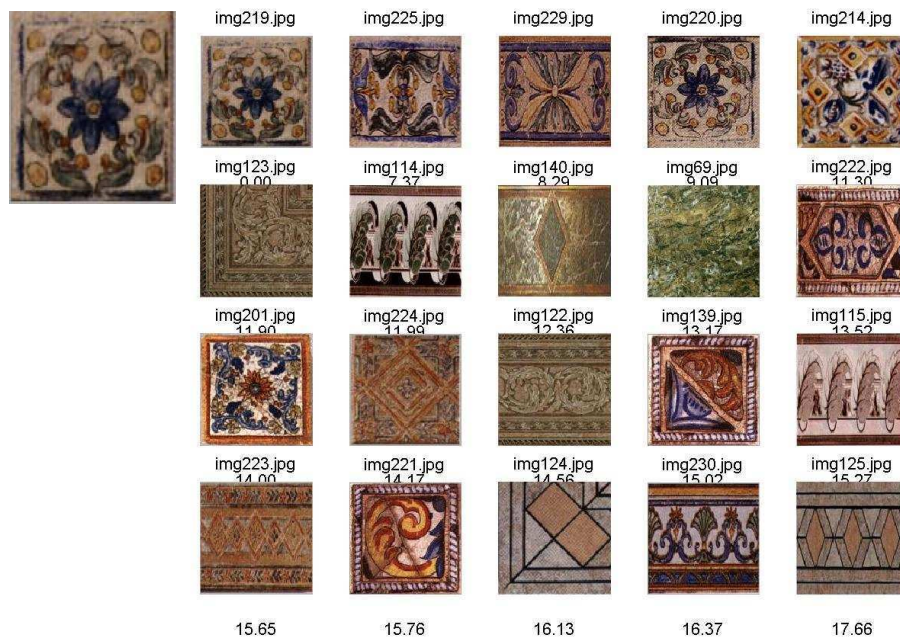
Sistemas de recuperación de contenido (CBMR)

Resultados de recuperación sin realimentación del usuario, usando sólo color global (ej. 2)



Sistemas de recuperación de contenido (CBMR)

Resultados de recuperación sin realimentación del usuario, usando color local y textura



Sistemas de recuperación de contenido (CBMR)

Entre los **campos de aplicación** de los sistemas de recuperación basada en contenido podemos citar:

- Bases de datos de galerías de arte y museos.
- Bases de datos policiales para la recuperación de obras de arte robadas.
- Archivos fotográficos.
- Bases de datos textiles, cerámicas o de otros productos industriales.
- Bases de datos de logotipos.
- Bases de datos de imágenes médicas.
- Bases de datos de rostros para identificación policial o búsqueda de personas desaparecidas.
- ... y muchas más que surgen cada día.

Sistemas de recuperación de contenido (CBMR)

Algunos sistemas y prototipos existentes son:

- Sistemas comerciales:
 - [QBIC](#) de IBM Corporation.
 - [Virage Search Engine](#) de Virage Corporation.
 - [Visual Retrievalware](#) de Excalibur Technologies Corp.
- Prototipos de investigación:
 - Photobook, desarrollado por MIT.
 - VisualSeek desarrollado en Columbia University.
 - Picasso, de la universidad de Florencia.

Sistemas de recuperación de contenido (CBMR)

Bibliografía

"Visual Information Retrieval", Alberto del Bimbo. Morgan Kaufmann.

"Perceptual metrics for image database navigation", Yossi Rubner, Carlo Tomasi. KAP.

"Exploratory Image databases. Content based-retrieval", Simone Santini. Academic Press.

"Image Databases. Search and Retrieval of Digital Imagery", V. Castelli and L. Bergman. Wiley.