

Diseño, implementación, evaluación y análisis de un Sistem de Recuperación de Información

David Orlando De Quesada Oliva, Javier Domínguez

MATCOM, Universidad de La Habana.

`d.quesada@estudiantes.matcom.uh.cu`, `j.dominguez@estudiantes.matcom.uh.cu`

Abstract. Este artículo aborda sobre la implementación de un Sistema de Recuperación de Información usando el modelo vectorial y el modelo fuzzy para la recuperación de documentos de texto de diversos formatos.

Keywords: pdf, txt, fuzzy, vectorial

Table of Contents

Diseño, implementación, evaluación y análisis de un Sistem de Recuperación de Información	1
<i>David Orlando De Quesada Oliva, Javier Domínguez</i>	
Diseño	3
Herramientas usadas para el el desarrollo del sistema ..	4
Principales métodos para la recuperación de imágenes	6
1 Preprocesamiento del texto:	6
1.1 Text Based Image Retrieval:	9
1.2 Field Based Image Retrieval:	10
1.3 Structure Based Image Retrieval:	10
2 Content Based Image Retrieval(CBIR):	10
2.1 Low-Level Image Feature:	11
2.2 Image Segmentation	11
2.3 Varias características de bajo nivel de las imágenes:	12
3 Semantic Based Image Retrieval:	14
4 Ontology Based Image Retrieval:	14
Evaluación del sistema	17
Aplicaciones	18
Ventajas y desventajas	20
References	22

Diseño

El sistema está diseñado, primero que todo, para pre-procesar los documentos que están en este. Este pre-procesamiento consiste en:

- Remover los stopwords (preposiciones, conjunciones y artículos).
- Remover los signos de puntuación.
- Remover los números.
- Llevar todas las palabras a minúscula.
- Stemming (llevar cada palabra a su palabra raíz).
- Separar por palabras el documento (tokenizar).

Los primeros 4 procesos en la implementación son modificables, o sea, es posible decidir si aplicarlos o no.

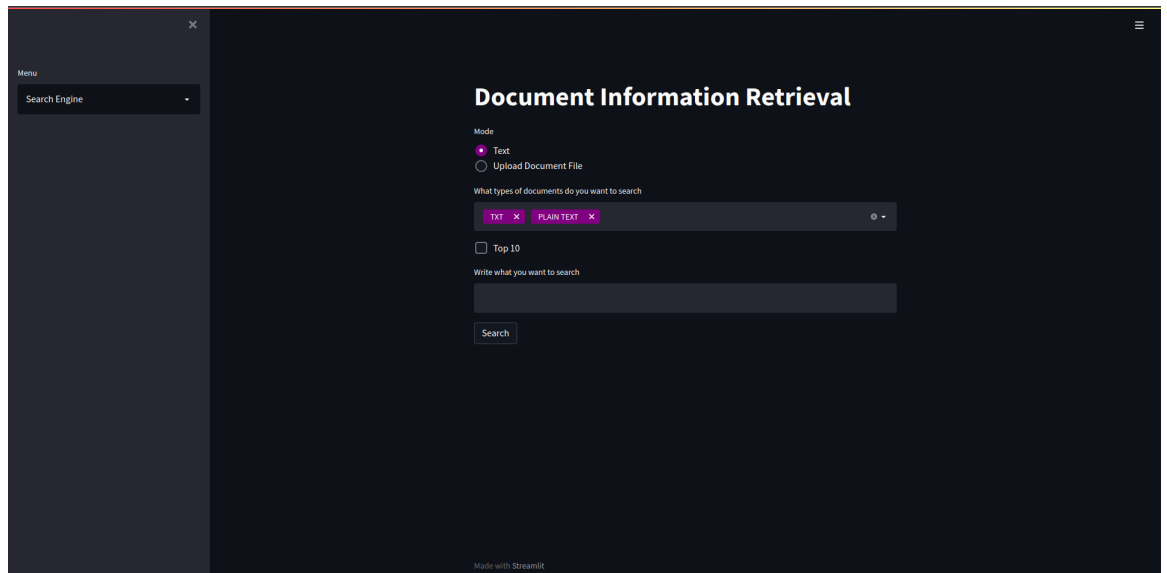
Herramientas usadas para el desarrollo del sistema

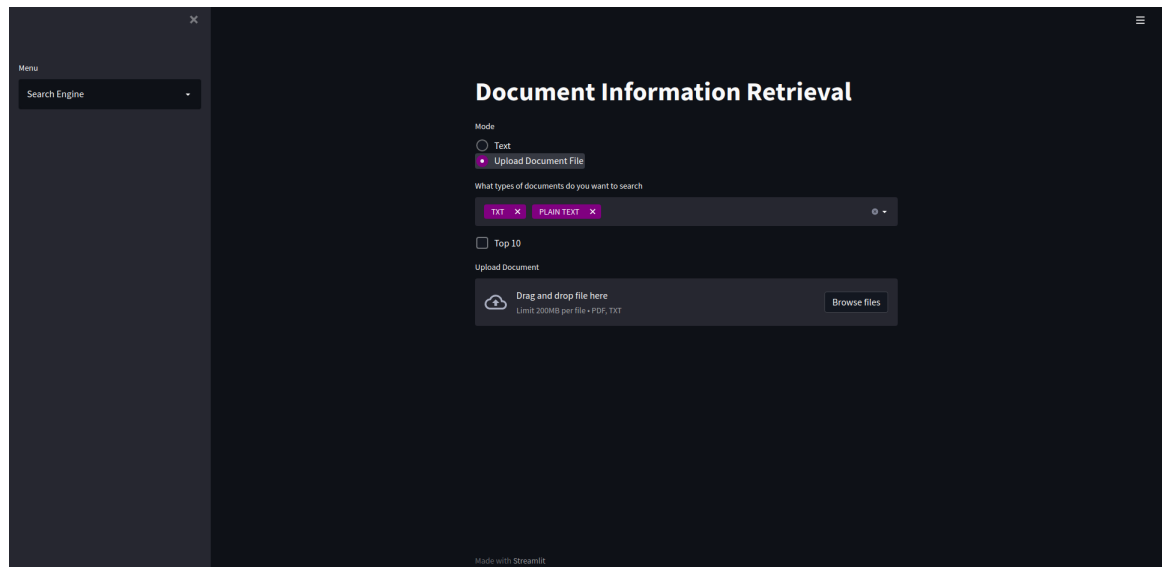
Para el desarrollo del sistema usamos Python como lenguaje de programación. Librerías de Python que utilizamos:

- nltk, para los stopwords, stemming y lemmatizing.
- streamlit para la interfaz gráfica.
- pdfminer y fpdf para procesar archivos .pdf.

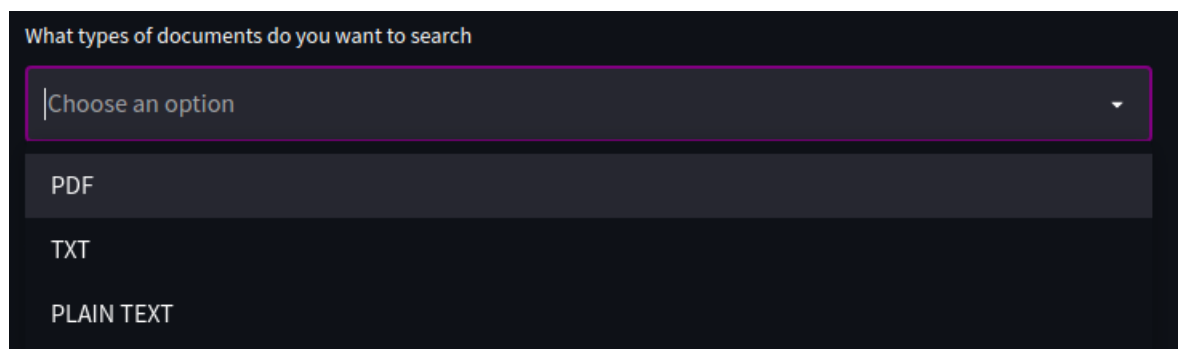
Para correr la aplicación: **streamlit run main.py**

Luego del pre-procesamiento es posible realizar las queries al sistema mediante la interfaz gráfica de dos formas, mediante texto, o mediante otro documento como se muestra en las imágenes siguientes:





Es posible escoger que formato de documento es el que se desea buscar:



Es posible escoger solamente los 10 elementos más relevantes de la query:



Principales métodos para la recuperación de imágenes

1 Preprocesamiento del texto:

Siempre que tengamos datos textuales, debemos aplicar varios pasos de preprocesamiento a los datos para transformar palabras en tokens que permitan un mejor resultado del modelo. En el sistema de recuperación de información que se implementó es posible aplicar los siguientes pasos de preprocesamiento del texto. La implementación de los mismo está en el file **text processing.py**. En muchos pasos hacemos uso de la librería NLTK (Natural Language Toolkit) de Python.

Llevar el texto a minúsculas: Se lleva todo el texto a minúsculas para reducir el tamaño del vocabulario de nuestro texto.

Implementación hecha en Python:

```
def text_to_lowercase(text: str):  
    return text.lower()
```

Remover números:

Podemos remover los números o convertir los números en sus representaciones textuales.

Implementación hecha en Python haciendo uso de expresiones regulares para eliminar los números:

```
import re  
def remove_numbers(text):  
    return re.sub(r'\d+', '', text)
```

Implementación hecha en Python para convertir los números a sus expresiones textuales haciendo uso de la librería inflect:

```
import inflect  
def convert_numbers_into_words(text: str)-> str:  
    result = []  
    for word in text.split():  
        if word.isdigit():  
            result.append((inflect_engine.number_to_words  
                           (word)))  
        else:  
            result.append(word)  
    return ' '.join(result)
```

Remover los signos de puntuación:

Eliminamos los signos de puntuación para no tener diferentes formas de la misma palabra. Si no eliminamos los signos de puntuación, entonces por ejemplo **call**. **call!** ,**call** serán tratados por separado cuando deberían referirse a la misma palabra.

Implementación hecha en Python para remover los signos de puntuación:

```
def remove_punctuation_marks(text: str):
    translator = str.maketrans('', '', string.punctuation)
    return text.translate(translator)
```

Remover los stopwords:

Los stopwords son palabras que no contribuyen al significado de una oración. Por tanto pueden eliminarse con seguridad sin causar algún cambio en el significado de la oración. La biblioteca NLTK tiene un conjunto de stopwords y podemos usarlas para eliminar los stopwords de nuestro texto y devolver una lista de tokens de palabras.

LIST OF ENGLISH STOPWORDS IN NLTK:

their, few, wasn't, has, m, or, did, isn, very, themselves, you've, you'd, do, between, other, t, shan, yourself, does, ours, i, it, should, what, himself, so me, itself, there, weren, most, her, mustn, hers, doesn, won, doesn't, hasn, s, y, wouldn't, didn't, him, couldn, after, a, will, ain, than, for, being, which, during, ll, my, isn't, its, any, hadn't, his, then, don, of, shouldn't, out, ou r, have, such, o, nor, too, re, should've, needn't, same, she's, but, weren't, all, against, down, don't, can, you, under, where, wouldn, only, been, aren't, haven, that, doing, if, up, d, needn, ma, yours, shan't, wasn, because, about, those, he, are, was, at, hasn't, over, until, had, with, you're, below, have n't, mightn, here, own, off, both, whom, while, as, ourselves, they, further, m ightn't, these, from, to, them, she, who, were, more, am, why, your, aren, had n, in, won't, yourselves, no, me, didn, an, so, before, is, on, now, each, how, be, theirs, shouldn, mustn't, above, herself, just, you'll, the, through, agai n, once, having, by, when, myself, we, it's, this, that'll, couldn't, ve, and, into, not,

Fig. 1. Lista de stopwords de NLTK

Implementación hecha en Python para el proceso de eliminar los stopwords

```
from nltk.corpus import stopwords
def text_remove_stopword(text: str):
    stop_words = set(stopwords.words("english"))
    return ' '.join([word for word in text_tokenize(text)
                     if word not in
                     stop_words])
```

Stemming:

El stemming es el proceso mediante el cual se obtiene la raíz gramatical de una palabra. Stem es la parte a la que se le añaden los afijos flexivos (-ed, -ize, -de, -s, etc. En el caso del idioma inglés) El stem de una palabra es creado removiendo el prefijo o sufijo de la palabra. Por lo que el proceso de stemming de una palabra puede no resultar en palabras gramaticalmente correctas del idioma

	words	stemmed words			
0	friend	friend	books	--->	book
1	friends	friend	looked	--->	look
2	friended	friend	denied	--->	deni
3	friendly	friendli	flies	--->	fli

Fig. 2. Ejemplos del proceso de stemming en algunas palabras

Hay principalmente 3 algoritmos para hacer stemming **Porter Stemming**, **Snowball Stemmer** y el **Lancaster Stemmer**. El Porter Stemming es el más usado entre ellos.

Implementación hecha en Python para hacer el proceso de stemming. Se puede pasar como parámetro un string que indica con cual de los tres algoritmos se quiere hacer stemming.

```
from nltk.stem.porter import PorterStemmer
from nltk.stem.snowball import SnowballStemmer
from nltk.stem.lancaster import LancasterStemmer

porter_stemmer = PorterStemmer()
snowball_stemmer = SnowballStemmer(language='english')
lancaster_stemmer = LancasterStemmer()
```

```
def text_stem_words(text: str, stemmer: str):
    if text == 'porter':
        stemmer = porter_stemmer
    elif text == 'snowball':
        stemmer = snowball_stemmer
    else:
        stemmer = lancaster_stemmer

    stems = [stemmer.stem(word) for word in text_tokenize
              (text)]

    return ' '.join(stems)
```

Lemmatizing:

Como el stemming, el proceso de lemmatizing también convierte una palabra a su raíz gramatical. La única diferencia es que el lemmatizing asegura que la raíz gramatical de la palabra pertenezca al lenguaje. Obtenemos palabras válidas si usamos el lemmatizing. En **NLTK**, podemos hacer uso del **WordNetLemmatizer** para obtener los lemmas de las palabras.

Implementación en Python para el proceso de lemmatizing usando el WordNet de NLTK

```
from nltk.stem import WordNetLemmatizer
lemmatizer = WordNetLemmatizer()
def text_lemmatize_words(text: str):
    lemmas = [lemmatizer.lemmatize(word, pos='v') for word in
              text_tokenize(text)]
```



```
return ' '.join(lemmas)
```

Separar el texto en tokens:

El último paso que hacemos siempre (este paso es necesario) es separar el texto en una lista de tokens de palabras removiendo los espacios. Para eso hacemos uso del word tokenize de NLTK.

Implementación en Python para separar el texto en tokens

```
from nltk.tokenize import word_tokenize
def text_tokenize(text: str):
    tokens = word_tokenize(text)
    return tokens
```

De la forma en que está hecho el código se le puede pasar uno o varios pasos de preprocesamiento al texto según se requiera.

```
def text_preprocessing(*, text: str,
    lowercase: bool = False,
    remove_numbers: bool = False,
    convert_numbers: bool = False,
    remove_punctuation: bool = False,
    stopwords: bool = False,
    stem: str = None,
    lemmatize: bool = False):
    ...
```

El método recibe un texto y se le pasan en True los parámetros con los que se quiere filtrar el mismo. En el caso del parámetro stem se le pasa un string que indica con cual de los algoritmos se va a hacer el stemming.

1.1 Text Based Image Retrieval:

Las técnicas básicas de recuperación de documentos pueden ser usadas para la recuperación de imágenes basadas en metadatos sin modificación. En un keyword based image retrieval, los metadatos que describen las imágenes pueden ser categorizados en 2 partes. Una parte se refiere a las herramientas usadas en el proceso de creación de la imagen, estilo de arte de la imagen, artista, precio, y otras propiedades explícitas de la imagen. La otra parte describe lo que realmente hay en la imagen, las propiedades implícitas que pueden entenderse al percibir la imagen en si. En el contexto actual de la recuperación, el texto plano anotado en imágenes responde de manera similar al texto plano en documentos, debido a que ambos contienen texto, lo cual permite que sean explotados por las técnicas convencionales de text-based information retrieval. La recuperación de información basada en texto genérica se realiza de tal manera que inicialmente el usuario realiza una consulta(query) que tiene de 1 a m **keywords**. En los sistemas de recuperación basados en metadatos(metadata based information

retrieval), el buscador compara los keywords con un conjunto de imágenes recopiladas de una base de datos y les da prioridad a los valores. Por ejemplo, si el keyword es **book**, y la imagen A contiene 2 ocurrencias de book y la imagen B solo una ocurrencia, entonces A tiene una prioridad mayor. Las imágenes con palabras claves anotadas son mostradas al usuarios en el orden de reducción de la prioridad. Imágenes irrelevantes son recuperadas y el usuario tiene que gastar tiempo en el filtrado de la información, usualmente navegando a través de los resultados de búsqueda.

1.2 Field Based Image Retrieval:

Field based retrieval es una extensión del text based retrieval donde solo un campo(field) es usado en anotación y recuperación. El enfoque basado en el campo (field based) describe y recupera artículos usando uno o más pares de valores del campo. Regularmente un esquema de metadatos es descrito por un conjunto de campos y pocas indicaciones sobre el tipo de valores que pueden ser elegidos por un campo particular. La plantilla (template) de metadatos y esquemas ampliamente utilizada para describir documentos online en general es la **Dublin Core(DC)**. Los campos de la DC version 1.1 son rights, coverage, relation, language, source, identifier, format, type, data, contributor, publisher, description, subject, creator y title. Versiones calificadas de DC han sido creadas para dominions particulares como la descripción de piezas de arte en museos.

1.3 Structure Based Image Retrieval:

El paradigma de recuperación basado en estructuras. En este método, se utiliza un enfoque basado en el campo(field) que principalmente utiliza una estructura de pares de valores atributo. Este método permite descripciones más complejas implicando relaciones. Por ejemplo, una definición de una parte de un auto puede incluir especificaciones de esos componentes. Cada elemento del objeto se puede especificar de nuevo usando varios atributos como la forma, el tamaño y el material. Los elementos pueden incluso tener elementos ellos mismos, por ejemplo, una mesa tiene patas, y sus subelementos pueden moverse hasta el nivel donde un elemento no puede obtener un subelemento más particular.

2 Content Based Image Retrieval(CBIR):

El Content Based Image Retrieval(CBIR) es uno de los métodos de visión por computadoras para la recuperación de imágenes, lo que significa que para poder recuperar es necesario imágenes digitales de una base de datos de imágenes. La búsqueda basada en contenido(Content based search) realizará el análisis con el contenido real de la imagen, en lugar de metadatos como etiquetas(tags), palabras clave(keywords), o descripciones anotadas con la imagen. La palabra contenido aquí puede referirse a formas, color, texturas o algún otro detalle que se puede obtener dentro de la propia imagen. El motor de búsqueda de

imágenes relacionadas con la web se basa en metadatos, por lo que genera una gran cantidad de resultados basura. Por lo tanto CBIR es deseable en este caso. Dándole palabras clave (keywords) de forma manual a las imágenes de búsqueda en una larga base de datos se pueden obtener resultados incorrectos. Además el proceso es costoso y puede que no identifique todas las palabras clave(keywords) que especifican la imagen y, por tanto, es ineficiente. Al proporcionar una buena técnica de indexación basada en el contenido real de las imágenes, se puede recuperar y producir resultados precisos.

2.1 Low-Level Image Feature:

Para poder realizar el CBIR las características de bajo nivel de la imagen (low-level image feature) deben ser extraídas primero. La extracción de características puede hacerse en toda la imagen o solo en una región de interés. La técnica simple usada en la recuperación de imágenes depende de las características globales. La percepción humana coincide estrechamente con la representación de imágenes a nivel de región. Para realizar la recuperación de imágenes basada en regiones el paso más importante es la segmentación de imágenes. De la región segmentada, las características de bajo nivel como textura, el color, la forma o la ubicación espacial se pueden extraer fácilmente. Basado en las características de la región, se puede encontrar fácilmente la coincidencia entre dos imágenes

2.2 Image Segmentation

El proceso automático de la realización de la segmentación de una imagen es una tarea difícil. Las técnicas académicas usadas en la segmentación de imágenes son curva de difusión de energía(curve energy diffusion), evolución(evolution) y particionamiento de grafos(graph partitioning). La mayoría de los métodos pueden ser apropiados solo para imágenes que tienen regiones con colores similares, como los métodos de direct clustering en el espacio de color. Tales métodos pueden adaptarse para la recuperación de sistemas que funcionen con colores. Pero las escenas naturales contienen tanto colores como texturas. Aplicar segmentación en texturas resulta difícil. Incluso en la segmentación basada en texturas la estimación del parámetro del modelo de textura es difícil. Para superar esto el algoritmo 'JSEG' es usado. Otro algoritmo llamado segmentación Blobworld es ampliamente utilizado. Algunos algoritmos de segmentación hacen uso de segmentación basada en color, en textura o en ambas. Estos algoritmos usan k-means para propósitos de clasificación. Los bloques de una misma clase se agrupan dentro de una misma región. El algoritmo k-means con restricción de conectividad (KMCC) es un trabajo de segmentación para segmentar objetos en las imágenes. Esta utilización del algoritmo se basa en la confianza, en la necesidad del sistema y el uso del conjunto de datos. Es difícil determinar que algoritmo proporciona mejores resultados. El resultado del JSEG es la textura y el color de regiones similares, pero el resultado de KMCC produce objetos

que son diferentes. El algoritmo KMCC es computacionalmente mucho más exhaustivo que el JSEG. Por tanto, Blobworld y JSEG son principalmente los algoritmos usados.

2.3 Varias características de bajo nivel de las imágenes:

■ **Color:** El color es la más común de las características adoptadas en la recuperación de imágenes. Varios espacios de color son usados para definir colores. Esos espacios de color son usados en dependencia de las diferentes aplicaciones. Los espacios de color más usados son RGB, LAB, LUV, HSV (HSL), YCrCb, y el hue-min-max-difference (HMMD). La covarianza del color, el histograma de color, y los momentos de color (color moments) son principalmente las características de color usadas en RBIR (Region Based Image Retrieval). El color principal (leading color), el color escalable (scalable color) y el diseño de color (color layout) son las características de color que se utilizan principalmente en MPEG-7. Con el origen las características de 3 colores, el par matiz-matiz y matiz se construyen las invariantes de color. La semántica de alto nivel no está directamente relacionada con las características de color mencionadas anteriormente. Para mapear los colores de una región a nombres de colores en semántica de alto nivel, el promedio de color de todos los píxeles en una región podría usarse como su característica de color. Si la segmentación es errónea terminará porque la región original es visualmente diferente al color promedio. Dependiendo de los resultados de la segmentación solo se seleccionan las características de color. El color promedio no es una opción deseable si la segmentación da como resultado objetos que no tienen colores similares. En la mayoría de los trabajos CBIR, las imágenes en color no están preprocesadas. Los filtros de color adecuados son esenciales para mejorar la eficiencia de recuperación debido a que el color en las imágenes siempre está dañado por el ruido.

■ **Textura:** Pocos sistemas no utilizan la textura como el color en la recuperación de imágenes. La textura es una característica importante para describir la semántica de alto nivel en la recuperación de imágenes porque proporciona detalles esenciales en un catálogo de imágenes, ya que define el contexto de muchas imágenes del mundo real como nubes, ladrillos, árboles y telas. El resultado de aplicar **la transformada de Wavelet** o **el filtrado de Gabor**, medidas estadísticas confinadas así como las seis características de textura de Samura, son las características de textura más utilizadas en el proceso de recuperación de imágenes. Las características de textura de Samura son:

- ▶ La regularidad
- ▶ Semejanza de línea
- ▶ La rugosidad
- ▶ La direccionalidad
- ▶ El contraste
- ▶ La aspereza

De estas características la aspereza, la direccionalidad y la regularidad son las más importantes. Estas tres están relacionadas con otras que son menos eficaces con respecto a la descripción de la textura. Los descriptores de navegación de texturas son obtenidos desde **MPEG-7**. Estos son regularidad, aspereza, direccionalidad. Se ha encontrado que la textura de Brodatz funcionará de manera excelente con características de palabras como aleatoriedad, direccionalidad y periodicidad. Las características de Tamura no funcionan para múltiples resoluciones que se consideren para la medición. Las características globales se ven afectadas por distorsiones de la imagen como diferencias de orientación debido a la distorsión del punto de vista y la escala. Si las regiones de textura en la imagen no están organizadas y son similares, se produciría una respuesta de recuperación deficiente para imágenes de escenas naturales. El estudio de la visión humana puede coincidir bastante con las características de Wavelet y Gabor en la mayor parte de la recuperación de imágenes. Pero el diseño actual del filtro de Gabor y la transformada de Wavelet solo está destinado a imágenes rectangulares. Pero en un RBIR la región de la imagen tiene formas erráticas. Por tanto en tal caso las características de la textura se utilizan eficazmente. Pero para la representación de imágenes naturales el descriptor de histogramas de borde (EHD) es el más adecuado y eficaz.

■ **Forma:** Uno de los conceptos más distintivos es la forma. Esta característica tiene un límite consecutivo de segmentos, una relación de aspecto, descriptores de Fourier, circularidad e invariantes de momento. El color y la textura son más útiles en imágenes particulares de dominio, como objetos artificiales. Aún así las características de la forma son características esenciales pero no tienen tanta popularidad en los RBIR como las características de textura y color para explorar los beneficios inherentes de RBIR, algunos sistemas podrían utilizar las características de forma como evaluadores. Por ejemplo, las características de orientación y excentricidad se utilizan para este propósito. La forma de una imagen puede definir la configuración de la superficie, características de un objeto, un contorno. Esto permite distinguir a un objeto de su entorno por su contorno. La representación de la forma puede generalmente dividirse en dos categorías: basado en los bordes y basados en regiones. La representación de la forma basada en los bordes solo usa los bordes exteriores de la forma. Esto se hace describiendo la región considerada usando sus características externas, como los píxeles a lo largo del límite del objeto. Pero la representación de la forma basada en la región es totalmente diferente al método anterior. Utiliza la región de forma completa describiendo la región considerando sus características internas; es decir los píxeles contenidos en esa región.

■ **Localización espacial:** No solo la textura, el color, o la forma son características importantes sino también la localización espacial en la catalogación de la región. Por ejemplo una imagen que contenga árboles con césped en el suelo podría tener características de color y textura similares pero su localización espacial ser diferente, normalmente las hojas de los árboles aparecen en

la parte superior de una imagen, mientras que las hojas que caen en la parte inferior. Las ubicaciones espaciales se definen intuitivamente como **:izquierda(left)**, **derecha(right)**, **top(arriba)**, **bottom(abajo)**, según el lugar de la región en la imagen. El rectángulo delimitador mínimo (minimum bounding rectangle) y el centroide de la región se utilizan para encontrar la ubicación espacial.

3 Semantic Based Image Retrieval:

Ni una, ni la combinación de varias características visuales o de bajo nivel (color, textura, forma, relación espacial) pueden capturar completamente los conceptos de alto nivel de las imágenes. Además, debido a que el rendimiento de la recuperación de imágenes basado en características de bajo nivel no es satisfactorio, hay una necesidad de que la investigación vaya dirigida hacia la recuperación de imágenes basada en el significado semántico, tratando de utilizar el concepto cognitivo del ser humano para traducir esas características de bajo nivel a conceptos semánticos de alto nivel (vacío semántico). Este acercamiento permite a los usuarios acceder a imágenes a través de consultas por texto, la cual es más intuitiva, fácil y preferida por los usuarios para expresar su deseo. Como mencionamos previamente en el CBIR, las imágenes pasan por un proceso de extracción de características de bajo nivel que se almacena junto con las imágenes. Ahora aquí, se hace necesario traducir esas características de bajo nivel a conceptos de alto nivel que no es capaz de comprender una máquina. Esta traducción usualmente se lleva a cabo utilizando herramientas de aprendizaje supervisado o no-supervisado, para asociar las características de bajo nivel con conceptos de alto nivel, los cuáles serán apuntados con palabras, durante el proceso de anotación de imágenes. Como podemos ver la recuperación de imágenes basada en la semántica utiliza técnicas de los 2 mecanismos que surgieron previamente, la extracción de características de bajo nivel utilizada en CBIR y la anotación de imágenes utilizada en el mecanismo basado en palabras claves, para almacenar en la imagen palabras **keywords** que se obtengan del proceso de conversión de características de bajo nivel a conceptos de alto nivel propio de este nuevo mecanismo.

4 Ontology Based Image Retrieval:

Ontología significa una descripción particular de una conceptualización. Diseña un dominio de manera formal. Con la ayuda de información textual en los alrededores solo la recuperación de imágenes web es lograda. Hay algunos motores de recuperación de imágenes dependientes de texto todavía disponibles en la web como Yahoo y Google. Estos usan características de texto como los nombres de archivos como índices para buscar imágenes en la web. Muchos motores de recuperación de imágenes están todavía bajo construcción. Los descriptores de bajo nivel de estos motores están lejos de las nociones semánticas. El otro tipo de sistemas solo se basa en anotaciones humanas. Por tanto, es necesario definir y hacer un enfoque intermedio para la compresión de imágenes. Algunos sistemas

pueden definir un dominio específico a partir de un dominio experto identificando vocabularios utilizados para describir objetos de interés. La cosa más deseable en la recuperación de imágenes es la ontología del concepto visual independiente del dominio. Este tipo de ontología admite el reconocimiento automático basado en técnicas de procesamiento de imágenes. En esto un dominio específico se especifica usando una estructura de árbol con herarquía de clases de sus subelementos en cada nivel.

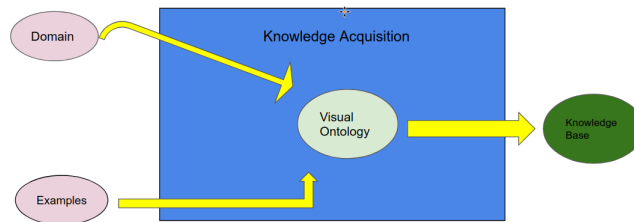


Fig. 3. Idea visual de la Ontología

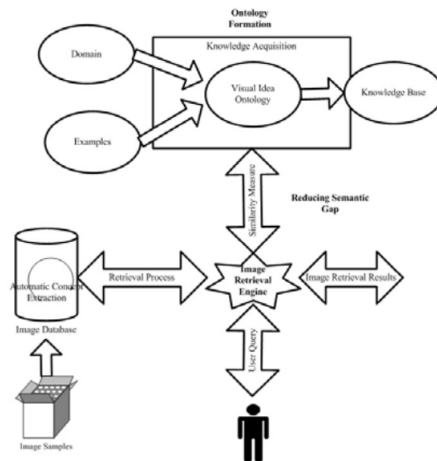


Fig. 4. Arquitectura de la ontología aplicada al proceso de recuperación de imágenes

La extracción de imagen hecha por una computadora puede resultar con conceptos de imagen significativos. Estos conceptos puede ser color, textura, forma o localización espacial. Mapear uno o más de esos conceptos resultantes en ontología interpretará el significado conceptual de una imagen. Si la query de recuperación captura la intención actual de los usuarios a través de la representación ontológica, definitivamente reducirá la disparidad semántica entre el hombre y la máquina.

¿Cual sería el mejor método?:

En un proceso inteligente de recuperación de imágenes, diferentes esquemas de indexación son aplicados comenzando por el text based, anotaciones de keyword, basada en compos(field based), basado en estructuras, basado en contenido(CBIR), basado en semántica hasta el basado en ontología. Por tanto la combinación de varios métodos sería lo más efectivo si se quiere obtener una respuesta bastante cercana a la query del usuario.

Evaluación del sistema

La recuperación de imágenes es esencialmente un problema de recuperación de información. Por tanto, las métricas de evaluación han sido adoptados de forma bastante natural a partir de la investigación de recuperación de información. Dos de las las medidas de evaluación más populares son:

- Precisión: número de imágenes relevantes recuperadas /total de imágenes recuperadas. .
- Recobrado: número de imágenes relevantes recuperadas /número de imágenes relevantes en la colección

Es importante tener en cuenta que cuando la consulta en cuestión es una imagen, la relevancia es extremadamente subjetiva. Mientras mayor sea el recobrado y la precisión , más eficiente será el algoritmo de recuperación en cuestión. El recobrado refleja la capacidad de recuperación de imágenes relacionadas, mientras que la precisión refleja la capacidad de rechazar las imágenes no relacionadas. Tradicionalmente, los resultados se resumen como curvas de recuperación de precisión o curvas de alcance de precisión. Una critica por la precisión se deriva del hecho de que se calcula para todo el conjunto recuperado y esta no se ve afectada por las clasificaciones respectivas de las entidades relevantes en la lista recuperada. Una medida que aborda el problema anterior y es muy popular en la comunidad de recuperación de imágenes, es la precisión media (AP). En una lista clasificada de entidades recuperadas con respecto a una consulta, si la precisión se calcula en la profundidad de cada entidad relevante obtenida, la precisión promedio se da como la media de toda la precisión individual. Como es obvio, esta métrica está muy influenciada por elementos relevantes de alto rango y no tanto por los que se encuentran al final de la lista recuperada.

Aplicaciones

El campo de la recuperación de imágenes ha demostrado ser uno necesario para los tiempos en que vivimos, donde tantas imágenes se generan diariamente, desde satélites orbitando la Tierra, misiones espaciales en lo más lejos del cosmos, animales en la naturaleza, paisajes, cámaras de seguridad y mucho más. Las imágenes son una fuente de información muy útil en la actualidad, y por tanto almacenarlas y organizarlas es una tarea necesaria, para facilitar el acceso a la información. De ahí que este campo tenga una gran cantidad de aplicaciones en la actualidad, a continuación mencionamos algunas de las más importantes:

■ El problema de la anotación automática de imágenes.

El propósito principal de un sistema de recuperación de imágenes basado en contenido es descubrir imágenes que pertenecen a algún concepto, en la ausencia de meta-datos, todos los intentos de automatizar el proceso de creación de estos meta-datos tiene ese objetivo. La anotación de imágenes puede facilitar la búsqueda de imágenes utilizando texto. Si el mapeo resultante imagen-palabras clave es confiable, la búsqueda de imagen basada en texto puede tener semánticamente más sentido que buscar en la ausencia de texto. Uno de los métodos para resolver este problema de la anotación automática es usar aprendizaje supervisado para categorizar las imágenes. La detección de conceptos simples como: paisaje, ciudad, animales, etc, alcanza una alta precisión.

■ El arte y la cultura siempre han sido importantes en la vida del ser humano. A lo largo de la historia, los museos y galerías de arte del mundo se han encargado de preservar nuestra diversa herencia cultural para utilizarlos como fuentes educación y aprendizaje. Es por esto que recientemente se ha expresado la preocupación por digitalizar todos los materiales antiguos, históricos y culturales, para la posteridad. Esto es muy importante por dos razones, primero las computadoras se han convertido en el principal medio de aprendizaje y se supone que así sea durante los próximos años, por tanto la representación digital de los artefactos culturales y las imágenes es algo que facilitará su popularidad, además de que sería accesible desde cualquier rincón del mundo, y segundo al contrario de la información almacenada de forma digital, los artefactos y pinturas antiguas están a merced a la degradación con el paso del tiempo, a los desastres y al vandalismo.

■ Las interacciones entre CBIR y la seguridad de la información ha sido prácticamente nulo, hasta que recientemente ciertas perspectivas han emergido para unir ambos campos, las pruebas de interacción humana (HIPs por sus siglas en inglés) y el cumplimiento de la protección de los derechos de autor. Mientras por un lado constantemente estamos ampliando las fronteras de la ciencia para diseñar sistemas que pueda imitar las capacidades humanas, no podemos negar los riesgos de seguridad inherentes asociados con programas extremadamente "inteligentes". Uno

de dichos riesgos es cuando un sitio web o algún servidor es atacado por programas maliciosos que solicitan servicios a escalas masivas. Pueden ser escritos programas que consuman una gran cantidad de recursos web o que influyan en los resultados de votaciones. En este caso los HIPs también conocidos como CAPTCHAs, son la solución. Estas interfaces están diseñadas para diferenciar entre humanos o programas, basados en la respuesta a algunas preguntas.

■ **Aplicaciones en la medicina** : Una aplicación que recupera cortes de 2D MRI(imágenes de resonancia magnética) de volúmenes cerebrales en 3D. La aplicación aceptaría query de imágenes 2D MRI del usuario e identificaría el volumen 3D que coincide de acuerdo con la región del cerebro relacionada con la query. Los cortes coincidentes se recuperan mediante una búsqueda adicional en los volúmenes 3D coincidentes.

■ **Aplicaciones en el comercio electrónico** : La compra electrónica es la tendencia actual en el mundo más que compras tradicionales debido a la conveniencia de internet. Cuando llega el tiempo de la compra, los usuarios se encuentran con un problema en seleccionar los items que necesitan comprar. Un sistema de recuperación de imágenes puede asistir a la decisiones del usuario en ropa y ayudarlo a tener una mejor selección de ropa durante la compra electrónica. El método usado enriquece la búsqueda y recuperación de productos de información usando CBIR con características como el color o la textura en la aplicación de la venta minorista de ropa electrónica.

Ventajas y desventajas

Keyword Based Image Retrieval

Ventajas:

- La realización de queries mediante texto suele ser más intuitiva y fácil para los usuarios.

Desventajas:

- No tiene en cuenta conceptos semánticos.
- Solo usando texto para describir el contenido de una imagen a menudo causa ambigüedad e insuficiencia en el rendimiento de la búsqueda de imágenes en una base de datos y el procesamiento de las queries de los usuarios.

Content Based Image Retrieval

Ventajas:

- Las queries se pueden realizar utilizando alguna imagen en algunos casos.
- Se computan las características de bajo nivel para lograr una mayor precisión a la hora de responder a una query.

Desventajas:

- Los sistemas basados en esta técnica sufren de lo que se denomina brecha semántica. Es una brecha entre los complejos conceptos semánticos que puede percibir el ser humano, y que las máquinas no son capaces de comprender fácilmente.

Semantic Based Image Retrieval

Ventajas:

- Es capaz de utilizando herramientas de machine learning lograr una mayor precisión a la hora de representar conceptos semánticos.

Desventajas:

- A pesar de que estos sistemas son mucho mejores que los existentes tiempo atrás, aún no existen mecanismos lo suficientemente potentes.

Ontology Based Image Retrieval

Ventajas:

- Usando la representación ontológica de imágenes se puede interpretar las ideas de alto nivel de la mente del usuario.

Desventajas:

- Si se define una ontología compleja el proceso de anotación también será complejo. El modelo de dominio ontológico es tan bueno como la mano experta que decide la ontología.

References

1. Ritika Hirwane(2012). Fundamental of Content Based ImageRetrieval. (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 3 (1)
2. Hui Hui Wang, Dzulkifli Mohamad, N.A. Ismail(2010). Approaches, Challenges and Future Direction of Image Retrieval. Journal of Computing, Volume 2, Issue 6
3. Chandra Mouli P.V.S.S.R.,Mohd Khalid Vijayan Vijayarajan(2012). A review: from keyword based image retrieval to ontology based image retrieval.International Journal of Reviews in Computing
4. John Eakins,Margaret Graham(1999). Content-based Image Retrieval. JISC Technology Applications Programme (October 1999)
5. Yu Xiaohong, Xu Jinhua(2008) The Related Techniques of Content-based Image Retrieval. 2008 International Symposium on Computer Science and Computational Technology
6. Elsaheed E. Abdelrazek(2017). A Comparative Study of Image Retrieval Algorithms for Enhancing a Content-based Image Retrieval System. Global Journals of Computer Science and Technology:Inc Volume 17 Issue 3 Version 1.0
7. Abby A. Goodrum(2000). Image Information Retrieval: An Overview of Current Research. College of Information Science and Technology Drexel University.
8. Unknown Author(2021,November 26). Image retrieval.Wikipedia.https://en.wikipedia.org/wiki/Image_retrieval.