Digital Content Retrieval – 6 credits (a.y. 23-24)
In RED solutions and comments

1) In a database of a professional photographer there are 11,000 digital photos. They are partitioned into the following classes:
Pictures of Interior: 2000 images
Natural landscape photos: 2500 images
Portraits and faces: 1500 images
Animals: 4500
Pictures of cars: 500

You make a query-by-example search submitting to the search engine the picture on the left:

The query returns the following photos: a "Ferrari" car, a dog, a cat, a mouse, a picture of the river Ticino, 10 passport photos, 5 photos of interiors. Calculate Precision and Recall of the query in question.

The image submitted in the query-by-example belongs to the class "Natural landscape photos". Therefore, only returned images of this class are "relevant" images to compute Precision and recall. How many relevant images are in the returned query? Only ONE (a picture of the river Ticino). The query returns 20 images. Therefore
Precision = number of returrned relevant images / number of returned images = 1/20 = 0.05

To compute Recall:

Recall = number of returned relevant images/Cardinality of the class "Natural landscape photos"
Recall = 1/2500 = 0.0004
If the output of the query should be:

a "Ferrari" car, a dog, a cat, a mouse, a picture of the river Ticino, 10 passport photos, 4 photos of interiors, a photo of a cat on a sofa near the Christmas tree. Is it possible to compute Precision and Recall in this case? If so, make the computation, otherwise motivate the answer.

The Precision and Recall values can be computed and are the same of thee previous computation because thee only ambiguity in the returned images (does "a photo of a cat on a sofa near the Christmas tree" belong to class Animals or class Interiors?) does not refer to thee class of the relevant images.

2) List the advantages of using wavelet transform with respect to the cosine transform in compression of still images of a multimedia repository of images.

Wavelet transform and cosine transform are both widely used techniques in image compression, each with its own set of advantages. Here are the advantages of using wavelet transform compared to cosine transform in the compression of still images:

1. **Multiresolution Representation**: Wavelet transform provides a multi-resolution representation of the image, capturing both frequency and spatial information at different scales. This allows for more efficient compression by focusing on important details while discarding redundant information.
2. **Localization of Information**: Wavelet transform localizes information in both frequency and spatial domains, which enables better preservation of image details. This localization property helps in reducing artifacts and maintaining image quality during compression.
3. **Adaptability to Image Characteristics**: Wavelet transform adapts well to images with varying characteristics and textures. It can handle both smooth areas and regions with sharp transitions effectively, resulting in better compression performance for diverse types of images.
4. **Energy Compaction**: Wavelet transform tends to compact most of the signal energy into a few coefficients, often referred to as the "energy compaction property." This allows for higher compression ratios while retaining perceptual image quality.
5. **Progressive Transmission and Reconstruction**: Wavelet-based compression methods support progressive transmission and reconstruction, allowing for the reconstruction of lower quality images from a subset of compressed data. This feature is beneficial for applications where images need to be transmitted over networks with varying bandwidth or displayed gradually.
6. **Ease of Implementation**: Wavelet transform algorithms are relatively straightforward to implement and can be efficiently computed using fast algorithms such as the Fast Wavelet Transform (FWT).
7. **Performance in Lossy and Lossless Compression**: Wavelet-based compression techniques perform well in both lossy and lossless compression scenarios. They offer a trade-off between compression ratio and image quality, allowing users to adjust parameters to suit their specific requirements.
8. **Flexibility in Compression Schemes**: Wavelet transform provides flexibility in designing various compression schemes by using different wavelet families, levels of decomposition, and coding strategies, allowing for optimization based on specific application needs.

While the cosine transform, like in JPEG compression, has its own advantages, such as simplicity and widespread support, wavelet transform, especially in methods like JPEG 2000, often provides superior compression performance with better image quality for various types of images, making it a preferred choice in many multimedia repositories and applications.

Mark the exact answers (one or more than one is possible) IN RED THE CORRECT ANSWER

3) The standard JPEG 2000 is based:
• On predictive coding
• On the Wavelet transform
• The cosine transform
• On encryption

4) In the biometric recognition of the iris the Hamming distance is used to:
• Pull out the position of the pupil
• Calculate the difference between features extracted from the iris
• Take only the comparison inter-class
• Take only the intra-class comparison

5) In the compression of still images, the entropy of the image is:
• the minimum number of bits / pixel reachable by a lossy compression algorithm
• the maximum number of bits / pixel reachable by a lossy compression algorithm
• the minimum number of bits / pixel reachable by a lossless compression algorithm
• the maximum number of bits / pixel reachable by a lossless compression algorithm

6) Please, list at least 3 names of metrics for image quality evaluation

Here are three common metrics used for evaluating image quality:

1. **Peak Signal-to-Noise Ratio (PSNR)**: PSNR is a widely used metric that measures the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. It quantifies the quality of an image by comparing the original image with a compressed or processed version, providing a numerical value indicating the level of distortion or loss.
2. **Structural Similarity Index (SSI or SSIM)**: SSIM is a perception-based image quality metric that compares the structural information of the original and distorted images. It evaluates the perceived change in structural information, luminance, and contrast, aiming to simulate human perception in judging image quality. SSIM often correlates well with perceived image quality.
3. **Mean Squared Error (MSE)**: MSE calculates the average squared differences between the pixels of the original and processed images. While it's a straightforward metric to measure the average squared difference in pixel values, it doesn't always correspond well to human perception as it doesn't consider structural information or other characteristics that affect perceived quality.

These metrics, among others, are used to quantitatively evaluate and compare the quality of images, providing numerical assessments of the level of distortion or degradation caused by compression, processing, or other manipulations.

7) JPEG2000 is not useful for:

- Compression of naturalistic images

- Compression of videos
- Compression of images of faces
- None of the previous ones

7) Which are the three paradigms of search by content in a multimedia database? List them with a brief explanation:
1) Search by metadata: metadata (very often they are textual metadata, are compared between the metadata suggester by the query and the metadata stored in the repository of indexes.
2) Search by example: the user submit an example of the content (for example an image, or a scketch of an object) and the search engine extract by this example the indexes according to the feature extraction algoritm. The indexes are compared to the indexes of each content instance stored in the repository.

In the first two paradigms, results are presented in a ranked order basing upon a distance metric.

3) Search by profile: the use is characterized by a profile and the system provide a set of results of queries which are typical for the users of the same profile. This paradigm is used in recommendation systems.

9) In a watermarking scheme for copyright protection in an image database:
- The watermark is always the same for all the images
- The watermark is provided by the user in run time during a query
- The watermark is a public key
- None of the previous ones

10) Draw the scheme of a one level Wavelet decomposition as filter banks.
See slides of TOPIC #6,7 (Slides3) and reference material at
http://csu.unipv.it/lucidimoda-copy/

In slides3, the scheme is at slide 4 (for 1D signals) and slides 5 (for 2D signals).
11) Which is the difference between the orthogonal and the biorthogonal Wavelets

Orthogonal and biorthogonal wavelet transforms are two types of wavelet transforms used in signal and image analysis. The main difference between these two is related to the orthogonality property.

1.
Orthogonal Wavelet Transforms:
•
Orthogonal wavelet transforms are characterized by the wavelet basis forming an orthogonal set. This means that the basis functions of the wavelet transform and their dilations and translations are orthogonal to each other.
•
Orthogonal wavelets can be symmetric or asymmetric and are used in many applications, including image compression, signal processing, and data analysis.
2.
Biorthogonal Wavelet Transforms:
•

Biorthogonal wavelet transforms have two sets of basis functions: one for decomposition and another for reconstruction. These two sets are not necessarily orthogonal to each other, but each of them can be orthogonal within its own space.

- The main characteristic of biorthogonal transforms is that the wavelet basis used for decomposition is not the same as that used for reconstruction. This can lead to greater flexibility in transform design and may be advantageous in some signal and image processing contexts.

In summary, the main difference between orthogonal and biorthogonal wavelet transforms is the property of orthogonality and the structure of the wavelet bases used for the decomposition and reconstruction of signals or images. Both variants have specific advantages and applications in different contexts, and the choice often depends on the needs of the application and the characteristics of the signal or image in question.