

VISUAL INFORMATION RETRIEVAL

Iker Vázquez López

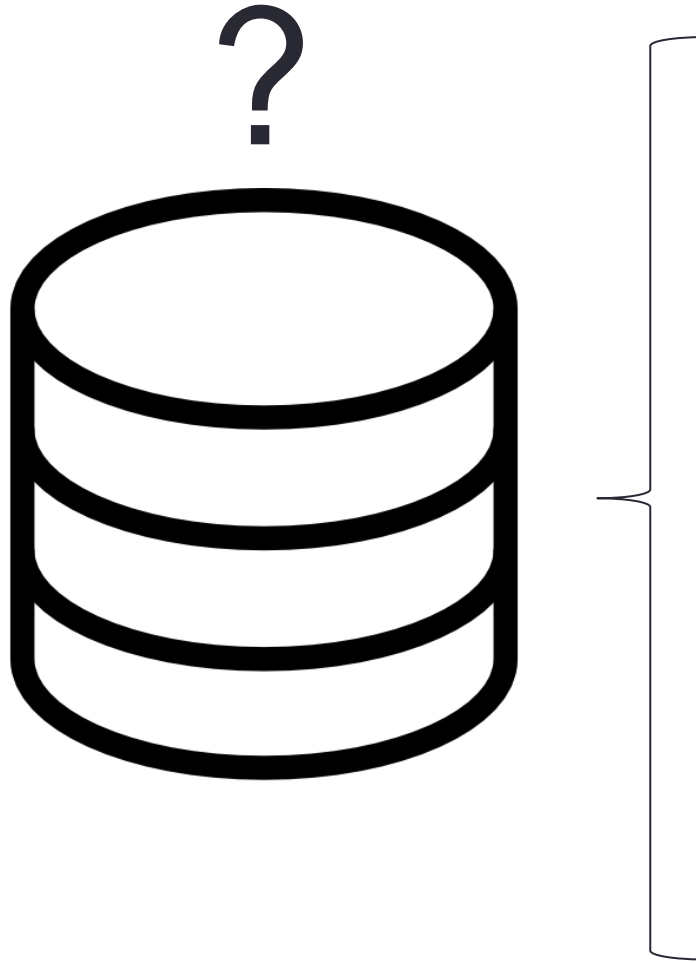
Information Retrieval

- 1951 - Calvin Moore about Information Retrieval:

Information retrieval embraces the intellectual aspects of the description of information and its specification for search, and also whatever systems, techniques, or machines that are employed to carry out the operation.”

- He was thinking about textual document retrieval

Searching for multimedia



Searching for multimedia

- Use queries?

“Find me a clip of about two seconds in which a red car racing along a hillside road on a bright day disappears as the road bends around the hill.”

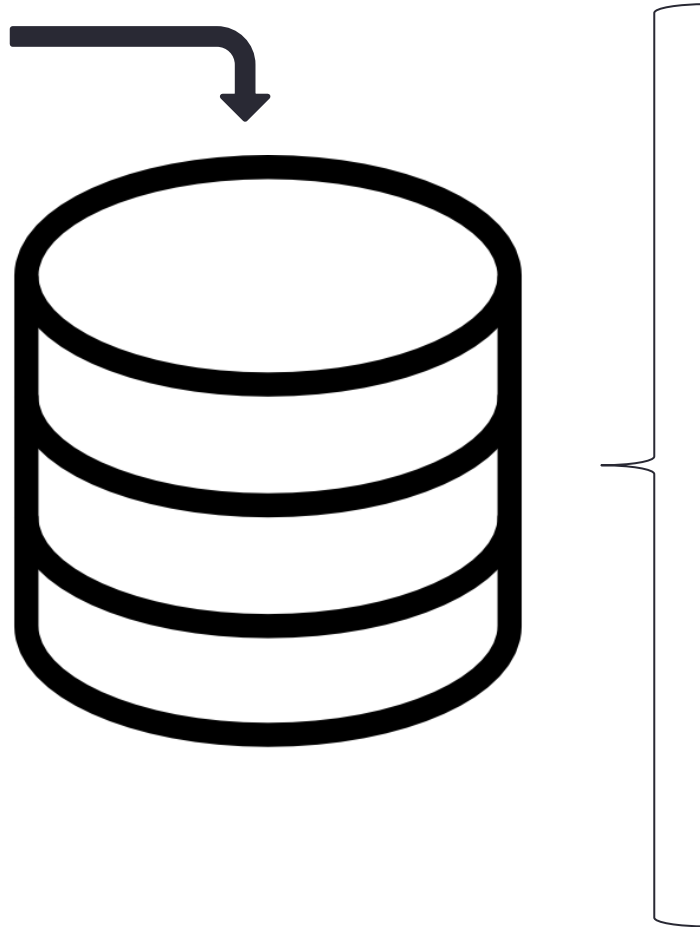
- Time
- Color
- Objects
- Movement
- ...

- Too many things to describe a video...

Searching for multimedia

- All the descriptors may be stored as metadata, but this data is too big to store.
 - Image (Pixel vector) or video (sequential images) + metadata
- Different users may use different queries to describe the same image

Searching for multimedia



- Use images to search similar images or videos.

Visual Information Retrieval

- To introduce VIR issues and techniques, they address three basic questions:
 - What constitutes the “information content” of an image or video in the specific context of any application?
 - With how much meaning can a user specify a search for a desired piece of information?
 - How efficient and accurate is the retrieval process?

What is Visual Information?

- Information about the object (metadata)
- Information contained within the object (visual features)



Flower, red, green, clouds...



Shape, texture, characteristic points...

Simple visual features

- Use image pixels (raw data)
 - Find all images for which the 100th to 200th pixels are orange if orange is defined as having a mean value of (red = 255, green = 130, and blue = 0).
 - Find all images that have about the same color in the central region of the image as this particular one. (Distance measure)
 - Find all images that are shifted versions of a particular image

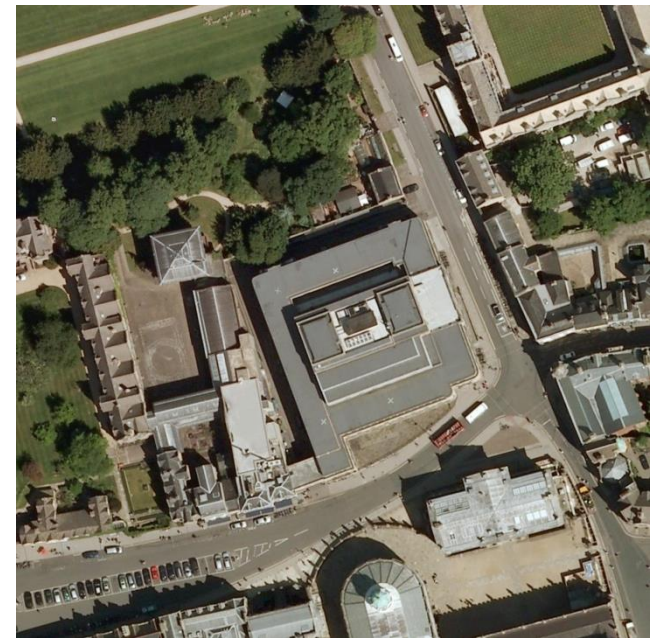
Simple visual features

- Problems:
 - Very sensitive to noise.
- Translation and rotation.
- Illumination variation.



Going further simplicity

- Consider an aerial images database.
 - Buildings.
 - Vehicles.
 - Roads.
 - Terrain.
 - Aircraft..
- Use human interaction for identifying specific areas.
 - Draw rectangles for each region in an image in which one or more of these five kinds of objects appear.

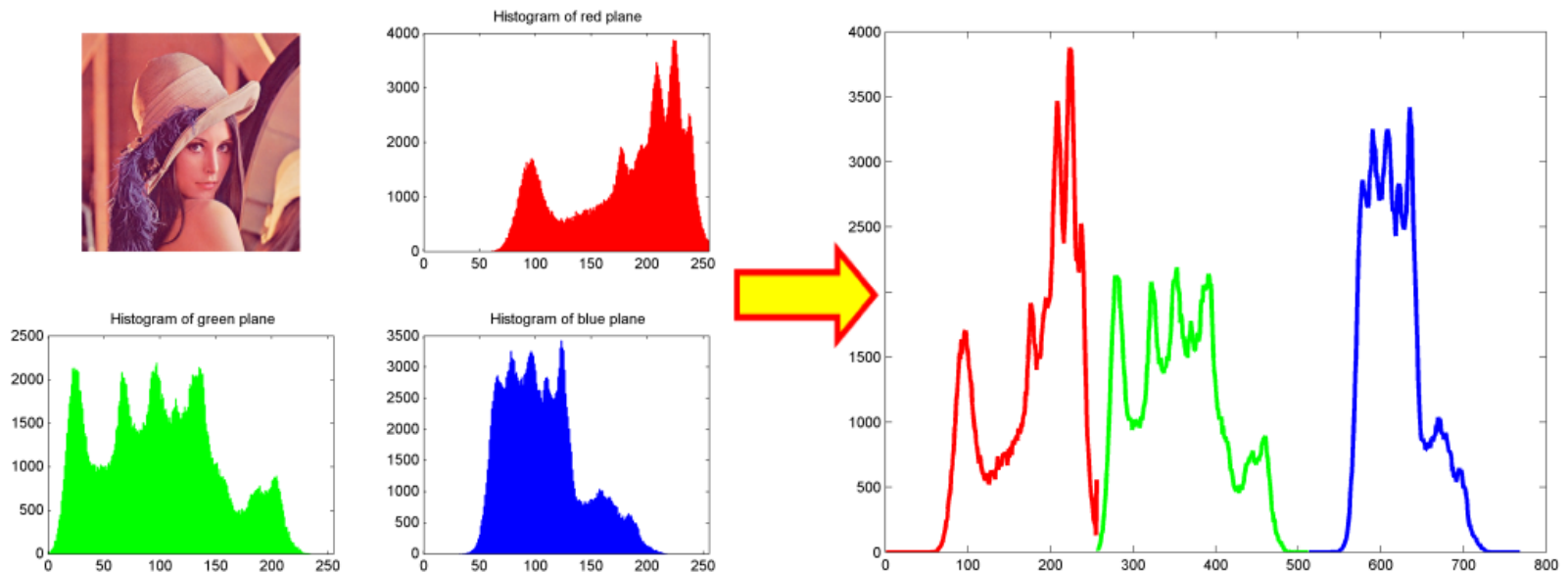


Going further simplicity

- Is there any location where more than five ground vehicles are close to a building located in the middle of the general terrain?
- Have there been any changes in the position of the aircraft at this location in the past couple of hours?
- Which approach roads have been used by ground vehicles over the past few days to come close to the aircraft?

Using the color

- Use the color to generate histograms.

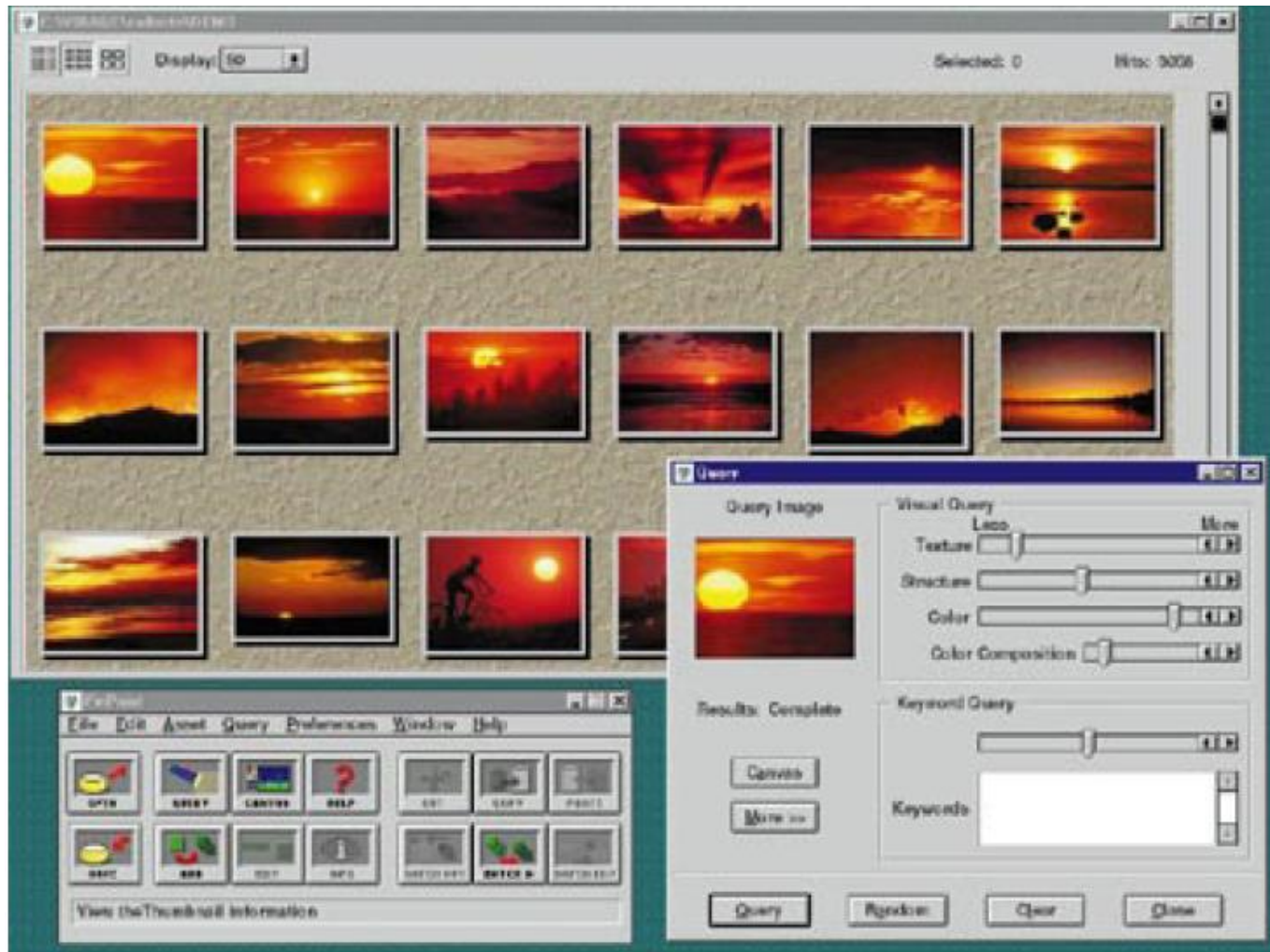


Using the color

- Find all images in which more than 30% of the pixels are sky blue and more than 25% of the pixels are grass green (an outdoor picture?).
- Find all images with the same dominant colors.
- Measure the distance of the different images.



Using the color



Using the color composition

- Color composition:
 - Do the same as before but subdividing the image in quadrants and compute their local histograms.



Using the color composition

- Find all images with more than 20% red-orange pixels in the upper right quadrant, more than 20% yellow pixels in the upper left quadrant, and about 30% brown to dark brown pixels in the lower half of the image (a sunset picture?).



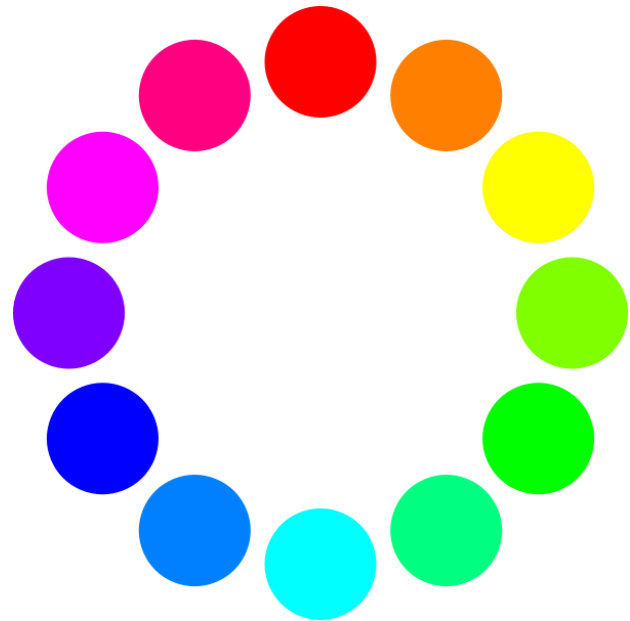
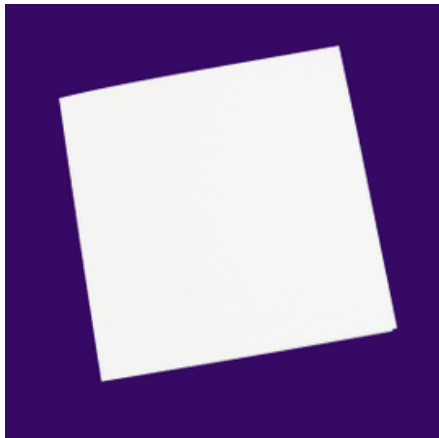
- Find all images with a red patch in the center of the image and with a blue patch around it.

Using shapes

- Shape
 - Segment the image into regions of connected colors.
 - Used most commonly on Clip-art images.
- Compute this four properties:
 - Color
 - Area
 - Elongation
 - Centrality

Using shapes

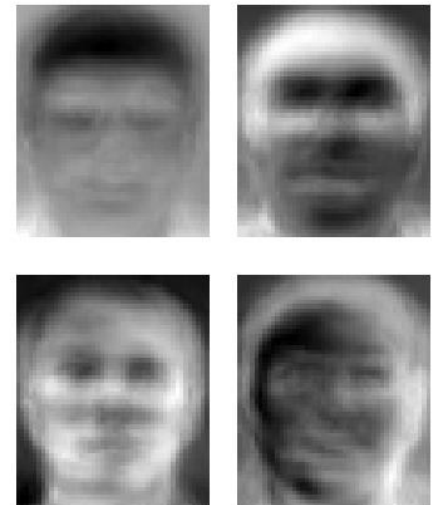
- Find all images having a dominant white square in the center.
- Find all images containing one circle of each color.



Using Face Retrieval

- With a database of face images it is possible to generate the called “eigenfaces”.
 - Although these features do not correspond to significant physical parts of the face they capture enough mathematically robust “information” to find similar faces with good precision.

- It is possible to search people in images



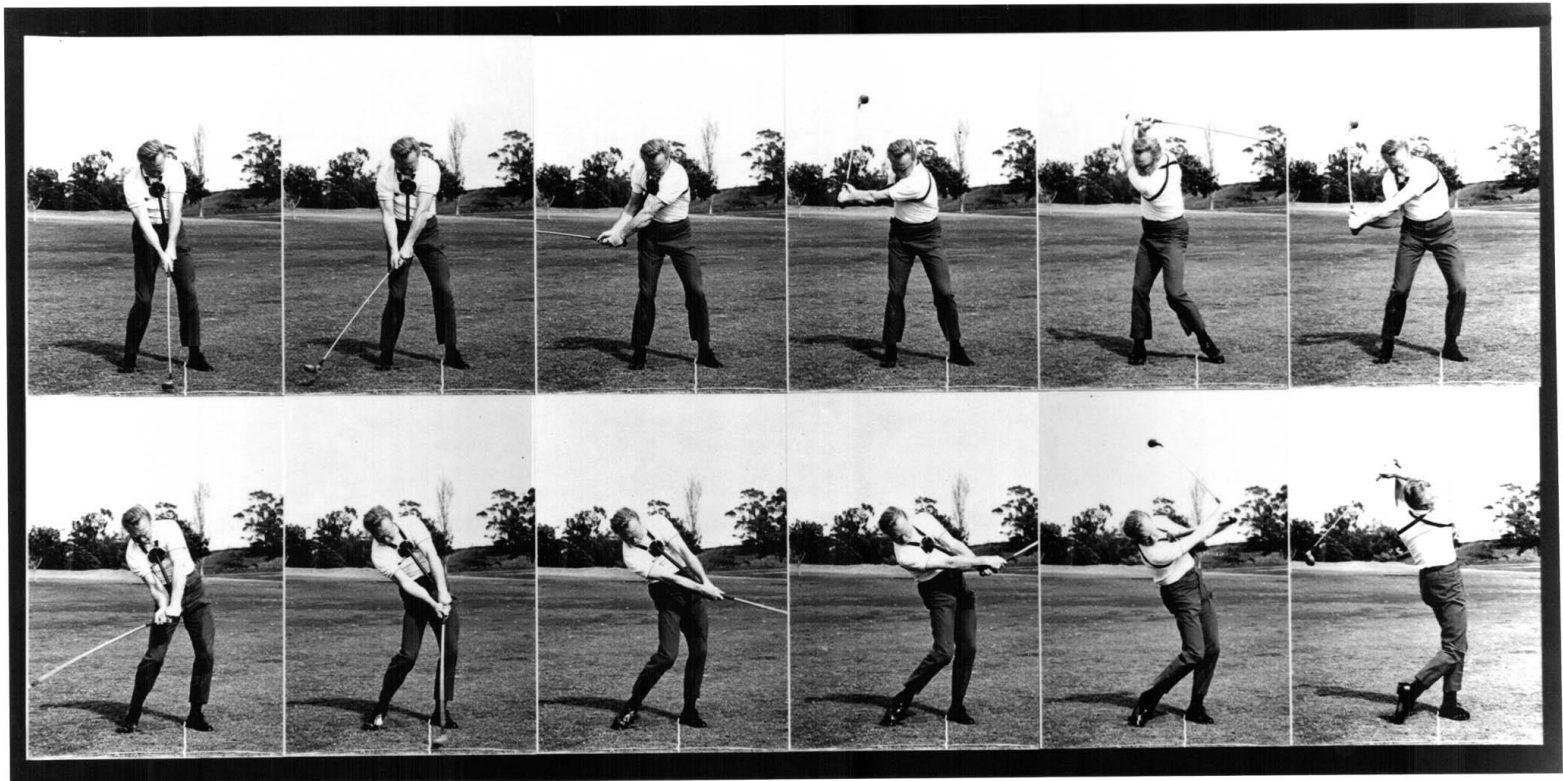
From image to video

What is a video?



From image to video

- Consider a video clip as a large number of image frames with progressively varying image content.



From image to video to image

- Extract some key frames from the video and treat them as normal images.
 - This technique does not make a good use of the time and motion.
- Videos have three types of motion:
 - Object motion.
 - Camera movement.
 - Post-processing motion.
- Some systems use the motion encoding in compressed video formats to extract the motion information. (Like the metadata from images)

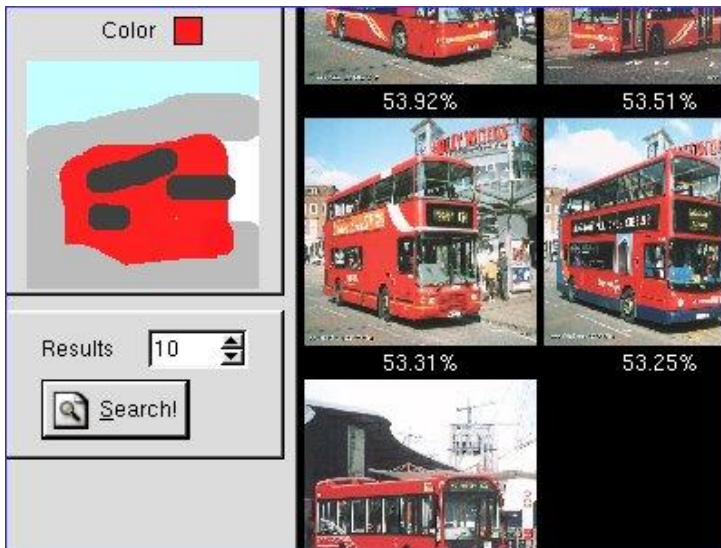
Specifying a Visual Information Query

- The primary intent of a visual query must be to capture the user's mental image of a specific picture or video image.
- A textual query language, such as PICQUERY+, has constructs to “compose” a visual description through textually specified attributes and operators.
- In a visual query language, the user visually places object icons to specify relative locations, orientations, and sizes of objects within the desired image.

Specifying a Visual Information Query

Image as a example

- The user specifies a query providing an image or by drawing what a image.



Value examples of the image

- The user specifies values of the image. For example 30% green and 50% blue with a grass-like texture in the green part.

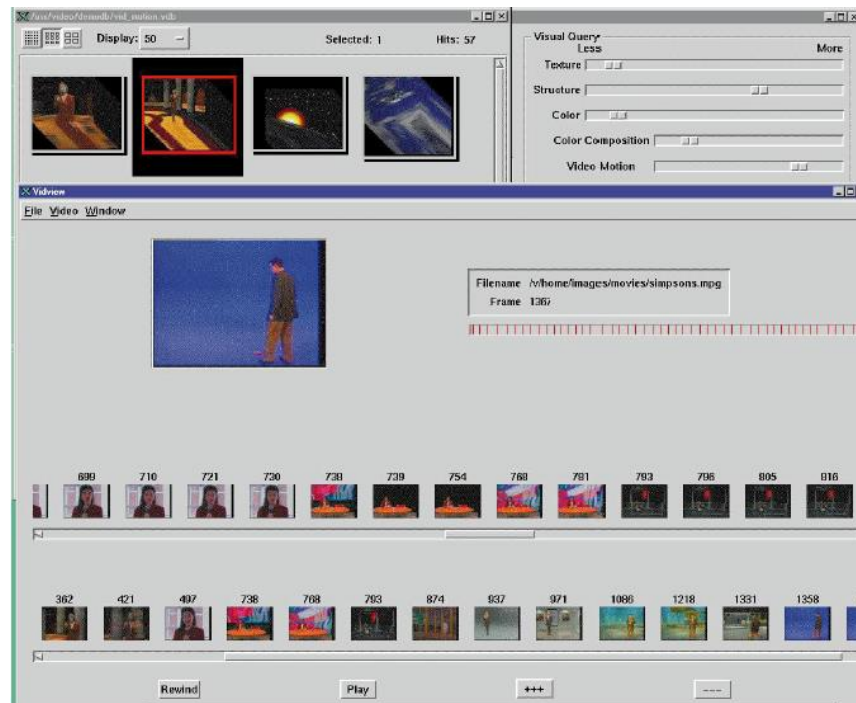


Specifying a Visual Information Query

- For video queries, is more complex to generate a query.
- Possibilities:
 - Find a video collection with a key-frame like a given image.
 - Rank the clips in the video collection in order of their similarity with a given video clip.
 - Cue the given video clip to the frame that is like the given query image.
 - Cue the given video to the frame that has the closest frame-wise similarity to the given query video.

Visual Information Retrieval systems

- The problem of this type of systems are two:
 - How to represent a image/video to search relevant images of the query.
 - How to create queries for this type of search.



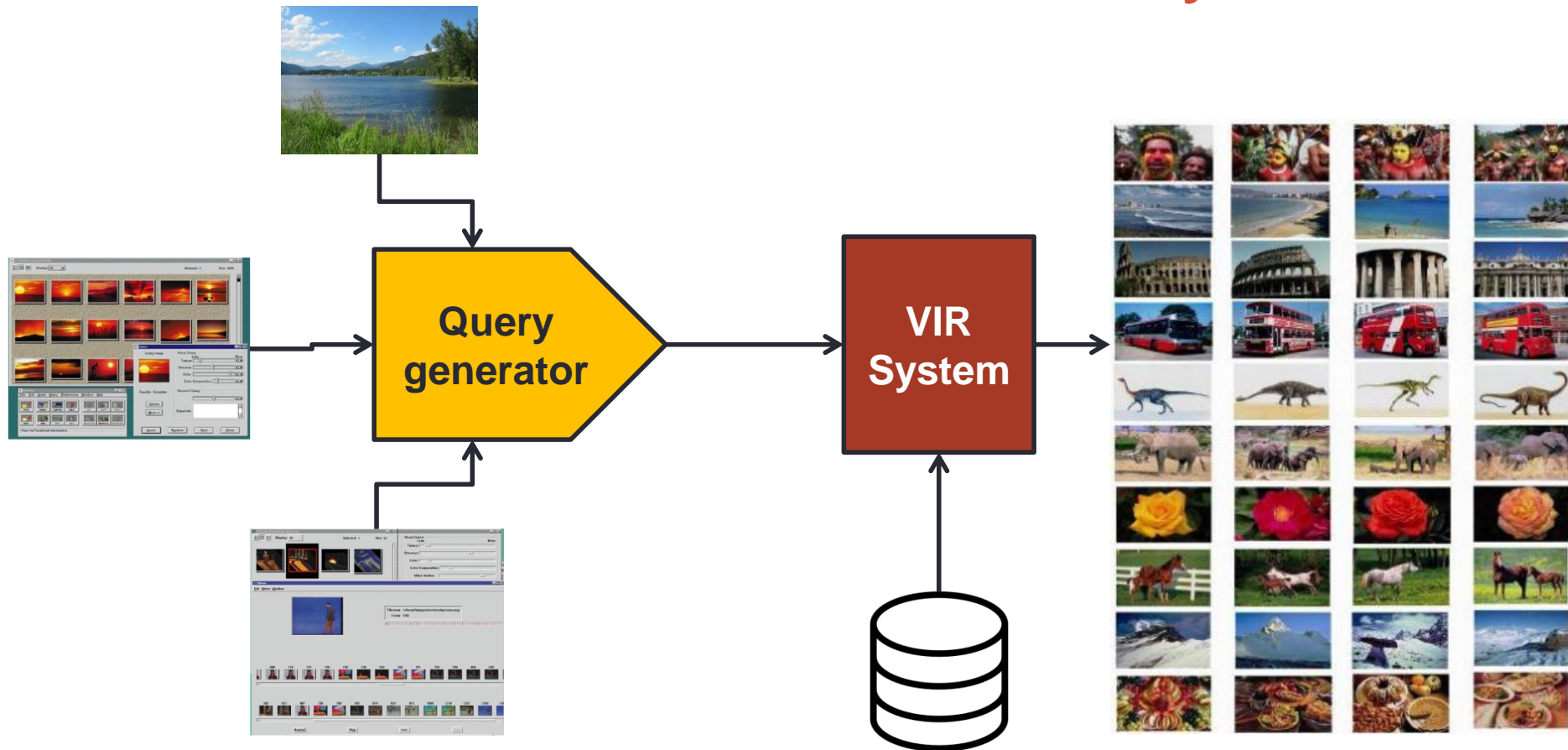
Visual Information Retrieval systems

- To represent an image:
 - Metadata
 - Visual features
 - Using raw data (pixels)
 - Determining objects in the image (human interaction)
 - Color
 - Composition
 - Shapes
 - Face retrieval

Visual Information Retrieval systems

- To create queries we need some tools:
 - An image-processing tool.
 - A feature-space manipulation tool.
 - An object specification tool.
 - A measurement specification tool.
 - A classification tool.
 - A spatial arrangement tool.
 - A temporal arrangement tool.
 - An annotation tool.

Visual Information Retrieval systems



HoloShots: Photorealistic Immersive Virtual Reality

In the HoloShots project we are building technology to let you step into photographs. By capturing stereo-panoramas at resolutions above 100 Megapixels virtual reality becomes reality.

<http://cs.boisestate.edu/~scutchin/holoshots/impano.html>