

# Making colors worth more than a thousand words

Jurandy Almeida   Anderson Rocha   Ricardo Torres   Siome Goldenstein

Universidade Estadual de Campinas  
Instituto de Computação  
CEP 13084-851, Campinas, Sao Paulo, Brazil

{jurandy.almeida, anderson.rocha, rtorres, siome}@ic.unicamp.br

# The problem

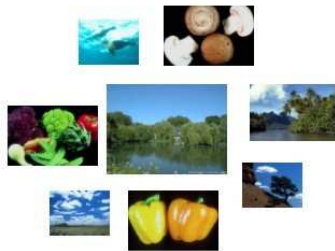
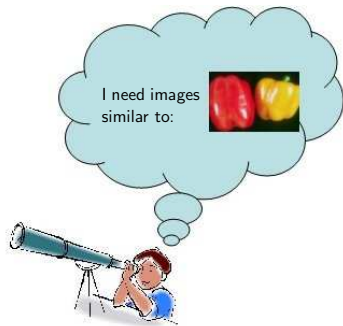
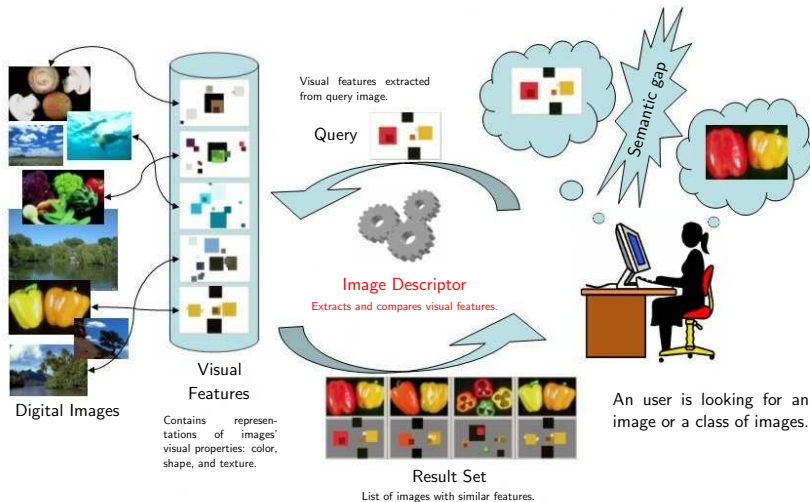


Image Collection



An user is looking for an image or a class of images.

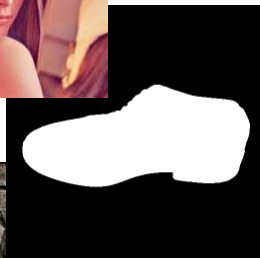
# Motivation



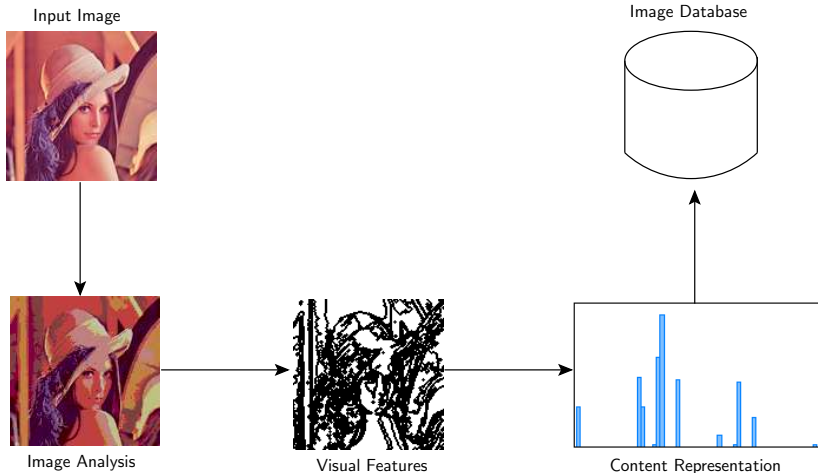
Content-based image retrieval system.

## Visual properties

- Color
- Shape
- Texture



# Color properties led to ...



Color-based image retrieval schema.

# Color image descriptors can be ...

## Global-based

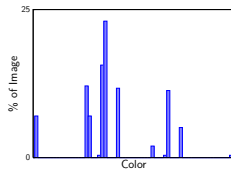
- GCH
- CCV
- BIC

## Partition-based

- LCH

## Region-based

- CBC



# Color image descriptors can be ...

## Global-based

- GCH
- **CCV**
- BIC



## Partition-based

- LCH

## Region-based

- CBC

# Color image descriptors can be ...

## Global-based

- GCH
- CCV
- **BIC**

## Partition-based

- LCH

## Region-based

- CBC



# Color image descriptors can be ...

## Global-based

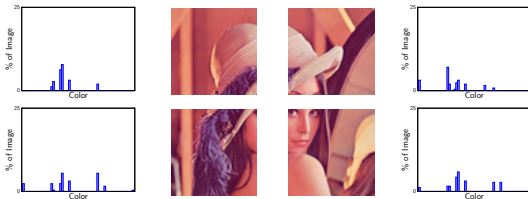
- GCH
- CCV
- BIC

## Partition-based

- LCH

## Region-based

- CBC



# Color image descriptors can be ...

## Global-based

- GCH
- CCV
- BIC

## Partition-based

- LCH

## Region-based

- **CBC**



# What if they fail?

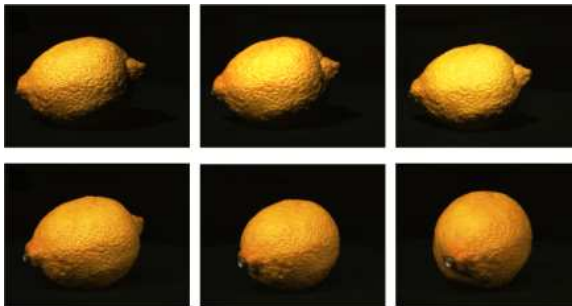


Image under presence of viewpoint, occlusion, and illumination changes.

# How can we address this problem?

## Color and Scale representative Image Regions (CSIR)

- 1 Robust to viewpoint, occlusion, and illumination changes.
- 2 Invariant to image transformations such as rotation and translation.
- 3 Uses an effective metric to compare images.

# CSIR meta-description approach

## Feature region detection

Search for local scale and rotation invariant feature regions  $\mathcal{R}$ .

## Description

- **Construct** a separate Gaussian pyramid  $\mathcal{G}_R$ ,  $\mathcal{G}_G$ , and  $\mathcal{G}_B$  for each color channel (R,G,B) of the input image.
- **For each** feature region  $r \in \mathcal{R}$ 
  - **Extract** local scaled and oriented patches  $\mathcal{P}$  from the Gaussian color pyramids  $\mathcal{G}_R$ ,  $\mathcal{G}_G$ , and  $\mathcal{G}_B$ .
  - **For each** patch  $p \in \mathcal{P}$ 
    - **Calculate** a local low-level color descriptor (e.g., BIC, GCH).

## Comparison metric

Use an appropriate metric to compare the images.

# CSIR meta-description approach

## Feature region detection

Search for local scale and rotation invariant feature regions  $\mathcal{R}$ .

## Description

- **Construct** a separate Gaussian pyramid  $\mathcal{G}_R$ ,  $\mathcal{G}_G$ , and  $\mathcal{G}_B$  for each color channel (R,G,B) of the input image.
- **For each** feature region  $r \in \mathcal{R}$ 
  - **Extract** local scaled and oriented patches  $\mathcal{P}$  from the Gaussian color pyramids  $\mathcal{G}_R$ ,  $\mathcal{G}_G$ , and  $\mathcal{G}_B$ .
  - **For each** patch  $p \in \mathcal{P}$ 
    - **Calculate** a local low-level color descriptor (e.g., BIC, GCH).

## Comparison metric

Use an appropriate metric to compare the images.

# CSIR meta-description approach

## Feature region detection

Search for local scale and rotation invariant feature regions  $\mathcal{R}$ .

## Description

- **Construct** a separate Gaussian pyramid  $\mathcal{G}_R$ ,  $\mathcal{G}_G$ , and  $\mathcal{G}_B$  for each color channel (R,G,B) of the input image.
- **For each** feature region  $r \in \mathcal{R}$ 
  - **Extract** local scaled and oriented patches  $\mathcal{P}$  from the Gaussian color pyramids  $\mathcal{G}_R$ ,  $\mathcal{G}_G$ , and  $\mathcal{G}_B$ .
  - **For each** patch  $p \in \mathcal{P}$ 
    - **Calculate** a local low-level color descriptor (e.g., BIC, GCH).

## Comparison metric

Use an appropriate metric to compare the images.

# CSIR meta-description approach

## Feature region detection

Search for local scale and rotation invariant feature regions  $\mathcal{R}$ .

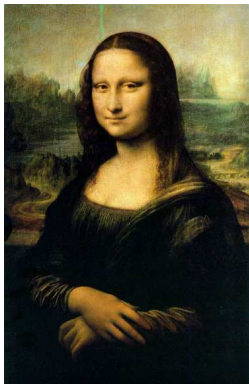
## Description

- **Construct** a separate Gaussian pyramid  $\mathcal{G}_R$ ,  $\mathcal{G}_G$ , and  $\mathcal{G}_B$  for each color channel (R,G,B) of the input image.
- **For each** feature region  $r \in \mathcal{R}$ 
  - **Extract** local scaled and oriented patches  $\mathcal{P}$  from the Gaussian color pyramids  $\mathcal{G}_R$ ,  $\mathcal{G}_G$ , and  $\mathcal{G}_B$ .
  - **For each** patch  $p \in \mathcal{P}$ 
    - **Calculate** a local low-level color descriptor (e.g., BIC, GCH).

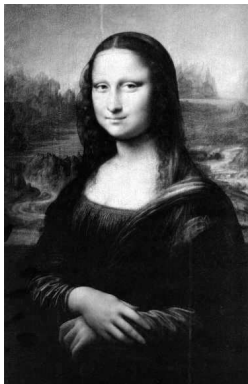
## Comparison metric

Use an appropriate metric to compare the images.

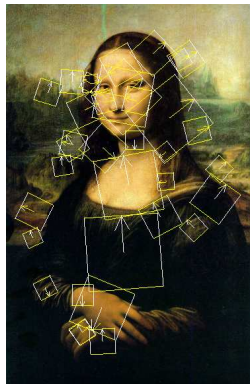
# Feature region detection



Input Image

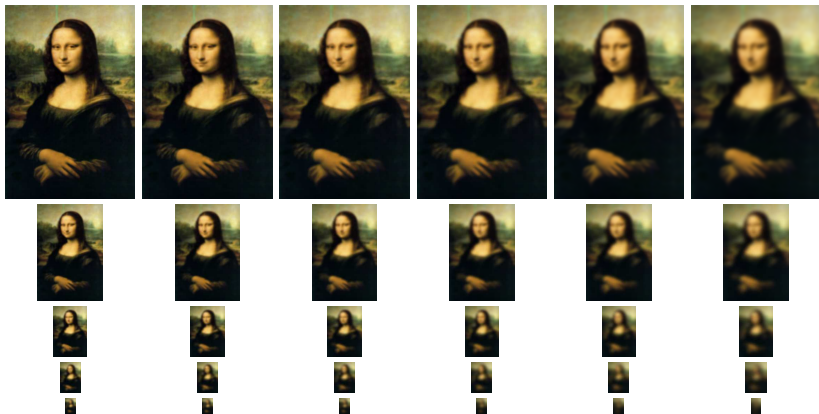


V channel

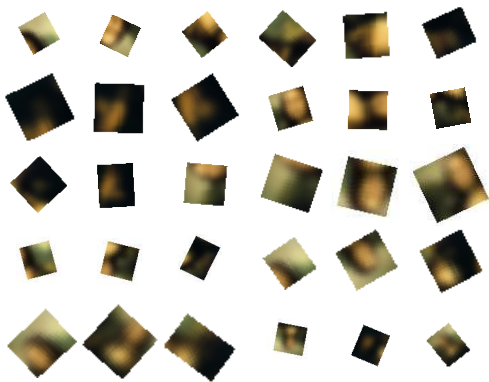


Regions

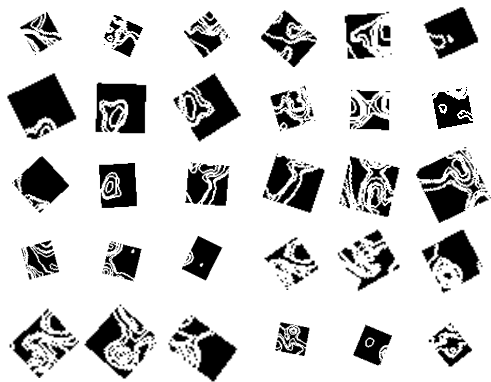
# Description (Gaussian Pyramid)



# Description (Patches)

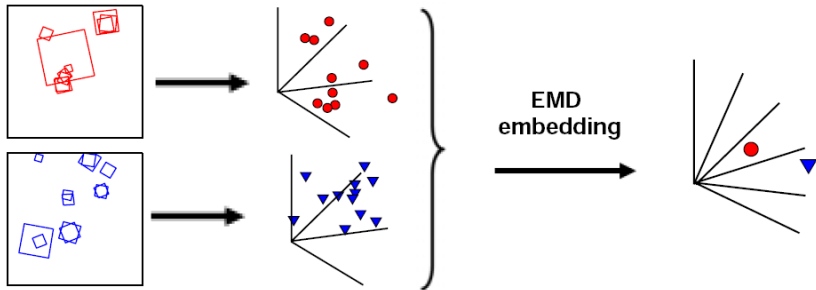


# Description (Descriptors)



## Earth Mover's Distance (EMD)

$$EMD(\mathbf{B}_p, \mathbf{B}_q) = \min_{\pi: \mathbf{B}_p \rightarrow \mathbf{B}_q} \sum_{s \in \mathbf{B}_p} \mathcal{D}(s, \pi(s)).$$



## Image databases

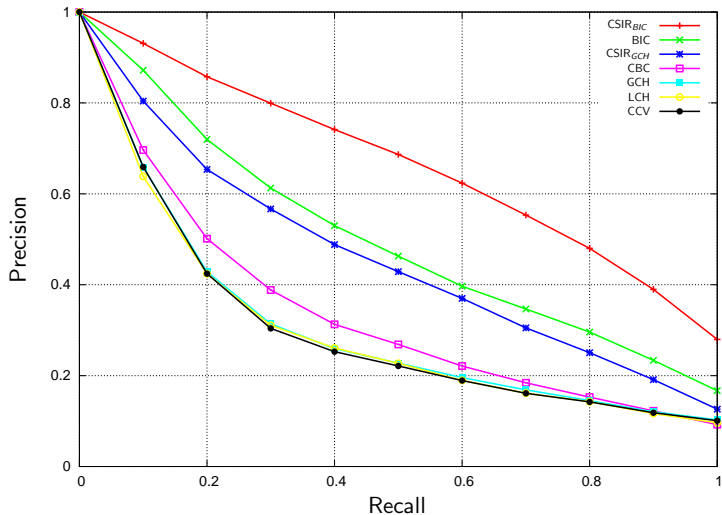
- ETHRel72 = Relevants + ETH80
- 1,320 images into 72 classes

## Retrieval effectiveness measures

- Precision vs. Recall ( $P \times R$ )
- $3P$ ,  $11P$ ,  $p_{30}$ ,  $r_{30}$ ,  $p_{100}$ ,  $r_{100}$ ,  $pR$



# Overall results



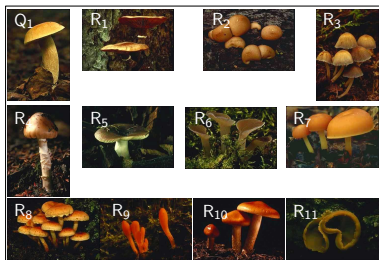
CSIR<sub>BIC</sub>, GCH vs. existing approaches.

# Overall results (cont.)

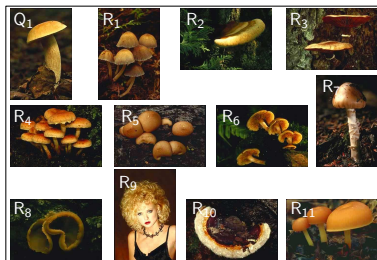
Approach	$3P$	$11P$	$p_{30}$	$r_{30}$	$p_{100}$	$r_{100}$	$p_R$
<b>CSIR<sub>BIC</sub></b>	<b>0.67</b>	<b>0.58</b>	<b>0.53</b>	<b>0.66</b>	<b>0.32</b>	<b>0.90</b>	<b>0.28</b>
BIC	0.49	0.42	0.39	0.52	0.23	0.76	0.17
<b>CSIR<sub>GCH</sub></b>	<b>0.44</b>	<b>0.38</b>	<b>0.37</b>	<b>0.49</b>	<b>0.21</b>	<b>0.70</b>	<b>0.13</b>
CBC	0.31	0.27	0.27	0.38	0.16	0.58	0.09
GCH	0.27	0.24	0.23	0.34	0.14	0.50	0.10
LCH	0.26	0.23	0.23	0.34	0.14	0.51	0.10
CCV	0.26	0.23	0.23	0.34	0.14	0.50	0.10

Unique values measurements results.

# Visual examples



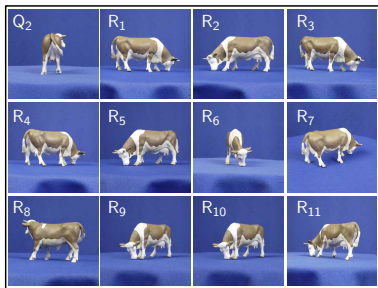
(a) CSIR<sub>BIC</sub>



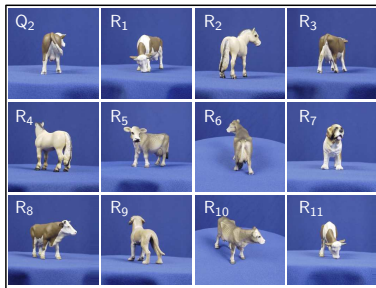
(b) BIC

Q<sub>1</sub> top-11 results.

# Visual examples (cont.)



(c) CSIR<sub>BIC</sub>



(d) BIC

Q<sub>2</sub> top-11 results.

## CSIR remarks

- 1 Robust to viewpoint, occlusion, and illumination changes.
- 2 Invariant to image transformations such as rotation and translation.
- 3 Uses an effective metric to compare images.

### Contribution

Instead of using the color pattern analysis in the whole image (as previous approaches) we use distinctive color and scale representative patterns that can be repeatedly found amongst similar images independent of some affine transformations.

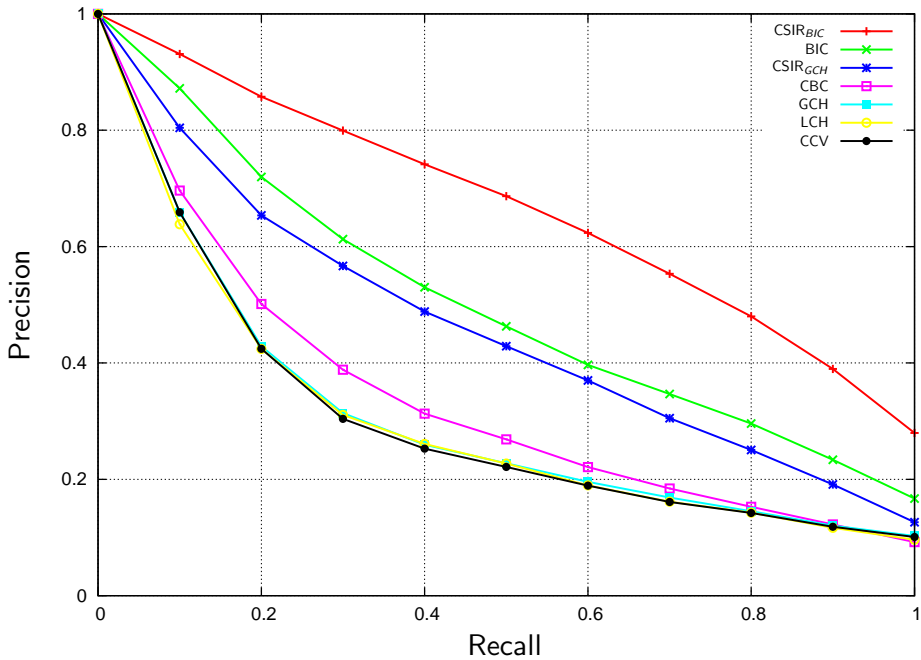
## Future work

Evaluation of other feature region operators and low-level image descriptors to improve the image representation.

# Acknowledgments

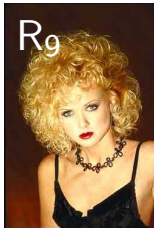








(e) CSIR<sub>BIC</sub>



(h) BIC

$Q_2$



$R_1$



$R_2$



$R_3$



$R_4$



$R_5$



$R_6$



$R_7$



$R_8$



$R_9$



$R_{10}$



$R_{11}$



Q<sub>2</sub>



R<sub>1</sub>



R<sub>2</sub>



R<sub>3</sub>



R<sub>4</sub>



R<sub>5</sub>



R<sub>6</sub>



R<sub>7</sub>



R<sub>8</sub>



R<sub>9</sub>



R<sub>10</sub>



R<sub>11</sub>

