PÓLO DO I.S.T

INSTITUTO SUPERIOR TÉCNICO

The Automatic Annotation and Retrieval of Digital Images of Prints and Tile Panels using Network Link Analysis Algorithms

Gustavo Carneiro and João P. Costeira

EUROPEAN COMMISSION

MARIE CURIE ACTIONS

FCT
Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA E DO ENSINO SUPERIOR

PrintArt

- Collaboration between
  - Instituto Superior Tecnico (Tech University of Lisbon)
  - Faculdade de Letras (University of Lisbon)
  - National Tile Museum (Lisbon, Portugal)

PrintArt

- System for organizing art image databases
  - Image annotation
  - Image retrieval
    - Text query
    - Image query

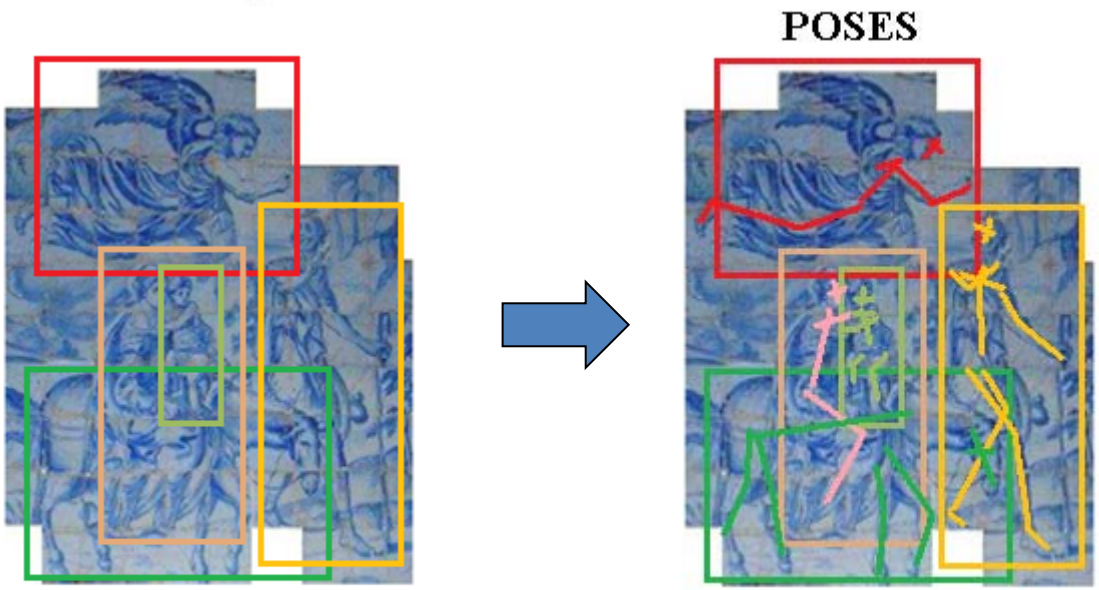
- Focus on Print and Tile panel images
  - Etching, engraving, lithography, etc.
  - Important for arts and art history

Global Image Annotation

Query image:

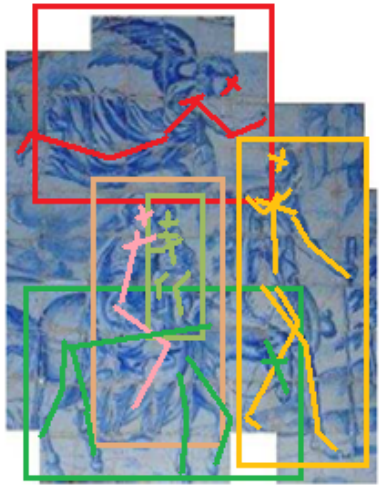
1. Initial annotation: Black and white print, Italy, XV century, Christian Art, Angel, Donkey, Robes, Saint Mary Blessed Virgin, Sain Joseph

Image Annotation



2. User correction: infant Jesus Christ, XVII century
3. Fixed annotation: Black and white print, Italy, XVII century, Christian Art, Angel, Donkey, Robes, Saint Mary Blessed Virgin, Saint Joseph, Infant Jesus Christ, flight into Egypt, Print by Pietro de Po, Paint by Nicolas Poussin

Image Annotation & Print Retrieval



3. Fixed annotation: Black and white print, Italy, XVII century, Christian Art, Angel, Donkey, Robes, Saint Mary Blessed Virgin, Saint Joseph, Infant Jesus Christ, flight into Egypt, Print by Pietro de Po, Paint by Nicolas Poussin

4. Image returned.

Image Retrieval

- Query: "Jesus Christ", "Saint Mary Blessed Virgin",and "Saint Joseph".

The system then returns the following images:



User selects the last two images above as relevant



User selects the last image above as the target image, and search is done.

Sensorial Gap

SCENE:
Crucifixion of
Jesus Christ

photograph

painting

printmaking image

artistic tile panel image

LOSING INFORMATION

3D > 2D, Color > Black and White, Texture > Textureless, Smooth > Tiles, …

Sensorial Gap

a) Changes: background, positions of donkey and angel; similarities: poses and textures of main subjects

b) Changes: background, the whole scene suffered a mirror transform; similarities: poses and textures of main subjects

Sensorial Gap

c) Changes: background, the whole scene suffered a mirror transform, and one of the subjects is positioned at a different place; similarities: poses and textures of main subjects

d) Changes: background; similarities: poses and textures of main subjects

Semantic Gap

TITLE: THE ANNUNCIATION

ARTIST: ABRAHAM DE BRUYN
PIETER VAN DER BORCHT

DATE: 1583

MEDIUM: ENGRAVING

CLASSIFICATION: PRINT

SUBJECTS: BLACK AND WHITE PRINTS
CHRISTIAN ART
RELIGIOUS SYMBOLISMS
NARRATIVE CYCLES
ANNUNCIATION
ANGELS
VIRGINS
INTERIOR VIEWS
CANOPY BEDS
CABINETS
PORTABLE DESKS
SCEPTRES
LILY
VISIONS
DOVE

SUBJECT PERSON: SAINT MARY BLESSED VIRGIN
ANGEL GABRIEL

PERIOD/STYLE: NETHERLANDISH/EUROPEAN

Annotations can change, depending on the Art Historian

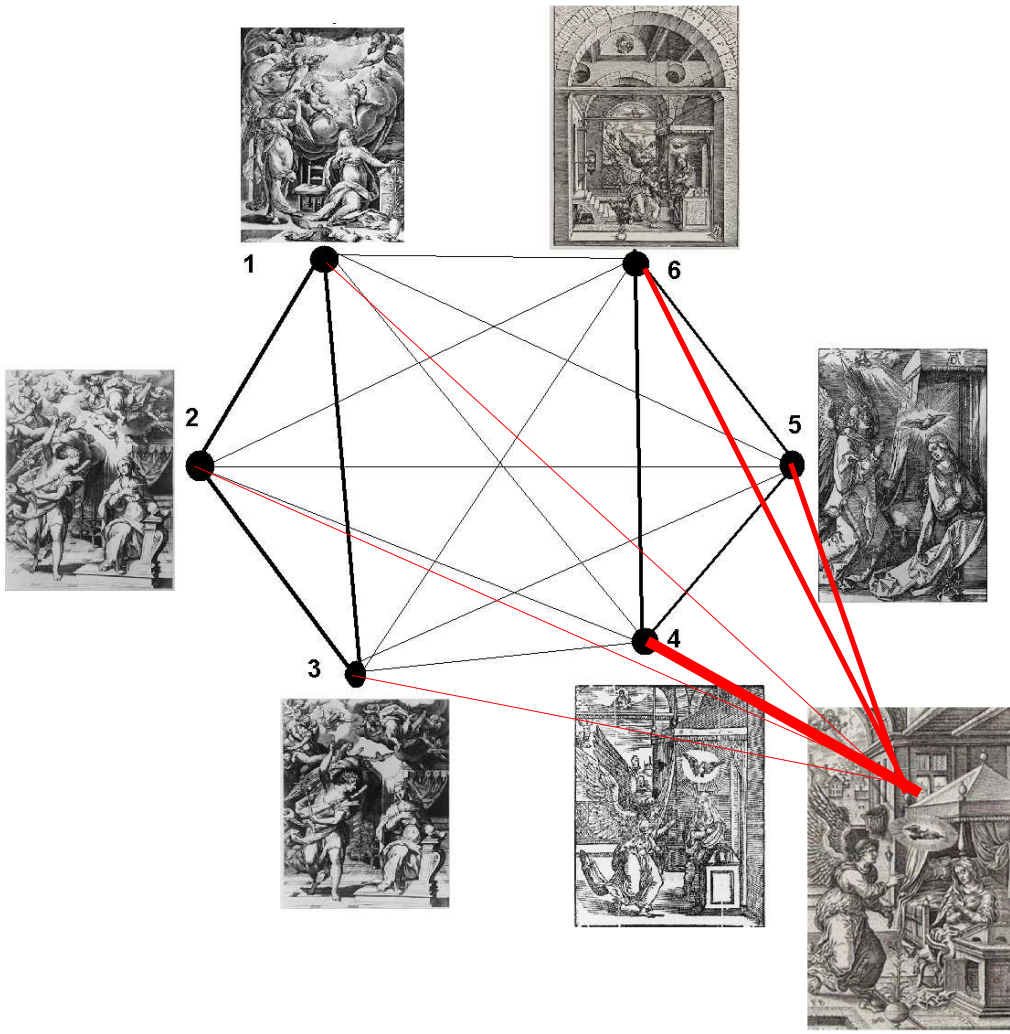
Methodology: Random Walk

- Similarity graph
  - Training image appearance: $\mathbf{x}$
  - Test image appearance: $\widetilde{\mathbf{x}}$
  - Label: $\mathbf{y}$
  - W: adjacency matrix

$$\mathbf{W}(j,i) = I_y(\mathbf{y}_i, \mathbf{y}_j) \times I_x(\mathbf{x}_i, \mathbf{x}_j) \times I_x(\mathbf{x}_j, \widetilde{\mathbf{x}})$$

- Random walk (k steps)

$$T_{r,k} = [(\mathbf{x}^{(r,1)}, \mathbf{y}^{(r,1)}), ..., (\mathbf{x}^{(r,k)}, \mathbf{y}^{(r,k)})]$$



Methodology
Annotation and Retrieval

- Probability of an annotation given a test image

$$p(\mathbf{y}|\widetilde{\mathbf{x}}) = \frac{1}{Z_T}\sum_{r=1}^{R}\sum_{k=1}^{K} p(T_{r,k}|\widetilde{\mathbf{x}})^{\frac{1}{k}} p(\mathbf{y}|\mathbf{x}^{(r,k)})$$

- Where the probability of a random walk sequence (Markov process) is

$$\begin{aligned} p(T_{r,k}|\widetilde{\mathbf{x}}) &= p([(\mathbf{x}^{(r,1)}, \mathbf{y}^{(r,1)}), ..., (\mathbf{x}^{(r,k)}, \mathbf{y}^{(r,k)})]|\widetilde{\mathbf{x}}) \\ &= \prod_{j=2}^{k} p(\mathbf{x}^{(r,j)}|\mathbf{x}^{(r,j-1)}, \widetilde{\mathbf{x}}) p(\mathbf{y}^{(r,j)}|\mathbf{y}^{(r,j-1)}) p(\mathbf{x}^{(r,1)}|\widetilde{\mathbf{x}}) p(\mathbf{y}^{(r,1)}) \end{aligned}$$

- Annotation: weighted sum of annotations
  - likelihood of a given annotation is the probability that the random walk reached the image of that annotation
- Retrieval

$$\mathbf{x}^*_{y_t} = \max_{\widetilde{\mathbf{x}} \in \mathcal{T}} p(y_t|\widetilde{\mathbf{x}}) p(\widetilde{\mathbf{x}}) / p(y_t)$$

Methodology
Image Feature and Annotation

- Image feature: Scale Invariant Feature Transform (SIFT) [Lowe, IJCV04]
- Image Representation: Hierarchical Bag of Features [Nister CVPR06]
- Image Similarity

$$I_x(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{\mathcal{Z}_x} \sum_{d=1}^{1111} \min(x_i(d), x_j(d))$$

- Label representation: binary vector (presence or absence of a visual class)
  - 1 multi-class problem
  - 21 binary problems
- Label Similarity

$$I_y(\mathbf{y}^{(r,j)}, \mathbf{y}^{(r,j-1)}) = \frac{1}{\mathcal{Z}_y} \sum_{l=1}^{M} \lambda_l \times y_l^{(r,j)} \times y_l^{(r,j-1)}$$

Experiments

- Bag of Features
  - Support Vector Machine (SVM) [Vapnik 98]
  - Random Forest (RF) Classifiers [Breiman 01]
  - Trained class by class with one vs. all scheme
  - Same image representation
- Database with 307 images
- 10 fold cross validation (90% training, 10% test)
- Retrieval

$$precision_R = \frac{|\mathcal{A}|}{n}, \text{ and } recall_R = \frac{|\mathcal{A}|}{|\mathcal{B}|}.$$

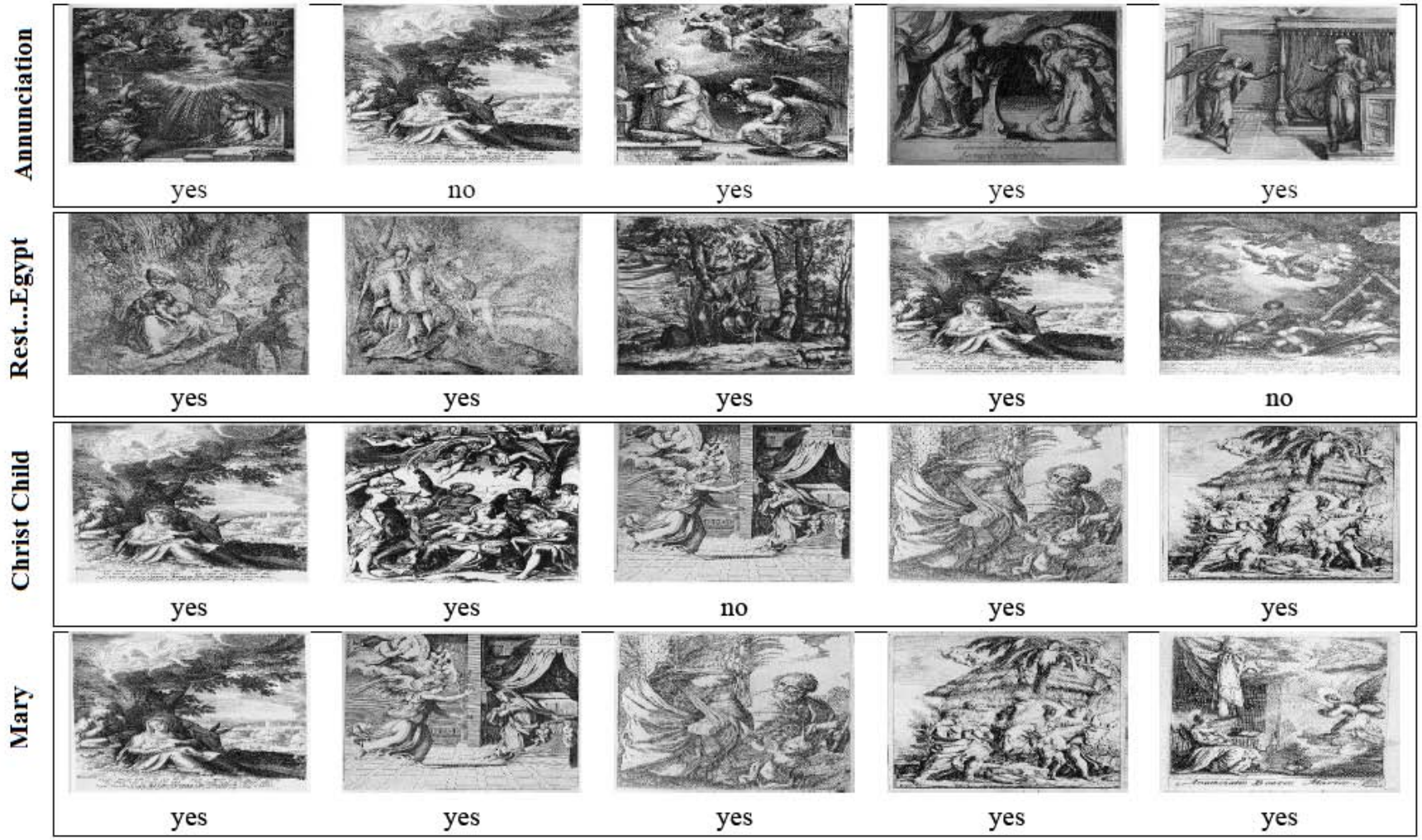
- Annotation

$$precision_A = \frac{w_C}{w_{auto}}, \text{ and } recall_A = \frac{w_C}{w_H}.$$

Retrieval

Table 1. Retrieval performance.

Models	Random Walk	RF	SVM
MAP	0.31	0.30	0.22



Annotation

Table 2. Annotation performance.

Models	Random Walk	RF	SVM
Mean Per-word Recall	0.51	0.40	0.12
Mean Per-word Precision	0.30	0.23	0.52

					
Human Annotation	Theme: Rest flight Egypt angels floating, Christ child, Mary, st. Joseph, wing	Theme: Magi Christ child, Mary, st. Joseph, wise men	Theme: Annunciation Gabriel, Lilly Mary, wing	Theme: Baptism of Christ angels floating, Christ dove, st. Frances, wing	Theme: Flight into Egypt angels floating, Christ child Mary, st. Joseph, wing
Rand Walk Annotation	Theme: Rest Flight Egypt angels, angels floating, Christ child, donkey, Mary, miracle...palm tree st. Joseph	Theme: Magi angels, angels floating, Christ child, Mary, Melchior, shepherd, st. Joseph, wing, wise men	Theme: Annunciation angels floating, dove, Gabriel, Lilly, Mary, Melchior, vase, wing, wise men	Theme: Baptism of Christ angels, angels floating, Christ, dove, shepherd, st. Elizabeth, st. Frances, wing, Zacharias	Theme: Rest Flight Egypt angels, angels floating, Christ child, donkey,Mary miracle...palm tree, st. Joseph, wing

Conclusions

- Better retrieval and annotation performance than state-of-the-art models
- SVM shows better precision than our model for annotation (but worse recall)
- Dependencies between labels (cannot have Melchior in Annunciation, for instance)
- Image features: try others (wavelet, curvelet, etc.)