From Partition Trees to Semantic Trees

Xavier Giró and Ferran Marqués
Outline

1. Introduction
2. Image representation
3. Semantics representation
4. Detection algorithm
5. Undirected detection
6. Examples
7. Conclusions
Introduction

- Bridge the semantic gap

Dangerous curve to the left

Perceptual information (visual)

Semantics

SEMANTIC GAP
Outline

1. Introduction
2. Image representation ←
3. Semantics representation
4. Detection algorithm
5. Undirected detection
6. Examples
7. Conclusions
Image representation

Colour-based segmentation [5]

Image representation

Binary Partition Tree [6]

Image representation

Example: Colour-based BPT
Image representation

- Visual descriptors for each BPT node [7]

Outline

1. Introduction
2. Image representation
3. Semantics representation
4. Detection algorithm
5. Undirected detection
6. Examples
7. Conclusions
Semantics representation

- Dual model (perceptual and semantic) of Semantic Classes

Visual Descriptors (VDs)  Description Graph (DG)
Semantics representation

- Semantic model with Description Graph [9].
- A **Description Graph** (DG) models a semantic class by assigning semantic instances and their Relations to its vertices.

**Example: DG of the semantic class “Curve traffic sign”**

![Description Graph](image)

Hierarchical decomposition in Semantic Trees (STs)
Example: ST of the semantic class “Curve traffic sign”
Outline

1. Introduction
2. Image representation
3. Semantics representation
4. Detection algorithm
5. Undirected detection
6. Examples
7. Conclusions
Detection algorithm

- Detection of a given semantic class (top-down approach).
  - Detection problem ➞ Building a ST on a BPT.

**Semantic Tree (ST)**

- Traffic Sign
  - Red triangular frame
  - White triangular background
  - Curve silhouette

**Binary Partition Tree (BPT)**

**Initial Partition**
Detection algorithm

Preanalysis is a generic and costly step.

[Diagram showing the detection algorithm with INPUT and OUTPUT pathways, including blocks labeled Preanalysis, Detector, Analysis, BPT+, VD, and Storage.]

Class’DG

\( f_{Th} \)

Image

Class instances in image with \( f > f_{Th} \)

INPUT

OUTPUT

Preanalysis

Detector

Analysis

Storage

INPUT

OUTPUT

Class’DG

\( f_{Th} \)

Image

\[
\text{Detector} \quad \rightarrow \quad \text{Class instances in image with } f > f_{Th}
\]

\[
\text{Preanalysis} \quad \rightarrow \quad \text{Analysis}
\]

\[
\text{Preanalysis is a generic and costly step.}
\]
Detection algorithm

Analysis is decomposed in 3 basic steps.

1) Perceptual analysis
2) Semantic analysis
3) Conflict resolution

Only instances composed by BPT nodes can be detected.
Detection algorithm

1) Perceptual Analysis

Example: Detection of classes “H” and “I” based on their perceptual models

- based on VD similarity
- class-specific detectors may also be used (e.g. face).
Detection algorithm

2) Semantic Analysis

- Top-down semantic expansion

Example: Detection of class “E” based on its semantic model

```
Semantic Mesh

class “E” DG
SR
G H

class “G” DG
SR
H I

Semantic Mesh

E

H

BPT nodes

(a) (b) (c) (d) (e)
```
Detection algorithm

2) Semantic Analysis

Graph matching between DG nodes and ST nodes in the mesh.

Semantic Mesh

Description Graph

class “M”
Detection algorithm

2) Semantic Analysis

- Discard matchings that create cycles through the Mesh.
- A cycle is created when a single instance is multiply considered in the same higher semantic instance (e.g., one shoe as part of two different legs for the same person).

Example: Refusal of two nodes (in grey)
Detection algorithm

3) Conflict resolution

- Each ST node can only sustain one ST node.
- Otherwise, one single instance would be part of more than one higher semantic instance (e.g., 1 mouth for 2 faces)
- Keep the highest and most confident node if conflict.

Example: Conflict resolution among three potential instances of class “F”.

\[
\begin{align*}
\text{f}_2 &> \text{f}_1 > \text{f}_3 \\
\text{Semantic Mesh}
\end{align*}
\]
Outline

1. Introduction
2. Image representation
3. Semantics representation
4. Detection algorithm
5. Undirected detection ←
6. Examples
7. Conclusions
Undirected detection

If no previous knowledge about which class to detect, an exhaustive perceptual analysis with all models in the database may launch a bottom-up semantic expansion.

Example: Undirected detection of class “G”
Outline

1. Introduction
2. Image representation
3. Semantics representation
4. Detection algorithm
5. Undirected detection
6. Examples
7. Conclusions
Examples

Example: Syntactic-based BPT [15]

Examples

(a)  
(b)  
(c)
Examples

**Semantic class: Frontal face**

![Diagram of a semantic tree
 for frontal face](image)

- **DG**: Parallel, Triangle, Eyebrows, Eyes, Nose, Mouth, Around
- **ST**: Frontal face, Eyes, Mouth, Eyebrows, Nostrils, Skin

**Segmentation**

- Image of a person
- Image of a detected object

**Detected object**

- Image of a person
- Image of a detected object
Examples

Semantic class: Italian Flag

DG
- Green Bar
- White Bar
- Red Bar
  - On the right of

ST
- Italian flag
- Green Bar
- White Bar
- Orange Bar

Segmentation

Detected object

(a) $f=0.94$, (b) $f=0.97$

www.eurunion.org
Examples

Semantic class: Laptop

$DG_1$

Screen
Chassis

Near

Around

Touchpad

$DG_2$

Screen
Keyboard

Near

Near

Touchpad
Outline

1. Introduction
2. Image representation
3. Semantics representation
4. Detection algorithm
5. Undirected detection
6. Examples
7. Conclusions
Conclusions

- Generic approach.
- Stored preanalysis (BPT+VD) allows fast semantic retrieval.
- Human-intuitive semantic models (Description Graphs) with direct application to analysis algorithms.
- Description Graphs introduce context in analysis.
- Performance depends on models accuracy.

**Future work:**

- Fusion of SVMs feature detectors in perceptual analysis.
- Semi-supervised semantic models creation.
- Evaluation with annotated databases (TRECVid, ImageEval, CHIL ?)
Outline

1. Introduction
2. Image representation
3. Semantics representation
4. Detection algorithm
5. Undirected detection
6. Examples
7. Conclusions