National Technical University of Athens, Greece
School of Electrical and Computer Engineering
Department of Computer Science
Image, Video & Multimedia Laboratory
Semantic Multimedia Analysis Group

A Context-based Region Labeling Approach for Semantic Image Segmentation

Thanos Athanasiadis, Phivos Mylonas and Yannis Avrithis

SAMT 2006, Athens

Overview

- Problem statement
- Knowledge representation
- Semantic vs. syntactic segmentation
- Visual context

Problem Statement & Proposed Approach (I)

- Importance of segmentation in computer vision applications (e.g. object recognition, image/video annotation, indexing & retrieval, compression/coding, etc.)
  - Processing of raw data can expose part of image's semantics, i.e. detection of specific concepts in constrained domains
  - How to deal with semantics in various scales (local, composite, global)?
  - Yet unsolved, yet very challenging

Our approach:

- Semantic vs. Syntactic: regions are assigned a fuzzy set of labels instead of numerical features
- Modification of traditional (region-based) segmentation algorithms to operate on labeled regions
- Extraction & exploitation of visual context
- **Simultaneous image segmentation and region labeling**

Problem Statement & Proposed Approach (II)

Target:

- Solve over-segmentation problems
- Assign labels with confidence values to regions
  - Of various scales, from tiny ones to whole image
- Accumulate all labels and link them with concepts existing in ontologies

Knowledge Representation

No semantics can be extracted without "any sort" of knowledge!

Two representation models, for two purposes:

- Graph Representation of an Image
  - Attributed Relational Graphs (ARG) is favored for image representation and analysis
- Ontology-based, domain specific, contextual knowledge representation
  - RDF-based knowledge model is ideal to store in and retrieve from a knowledge base

Common element:

Introduction & employment of fuzziness (fuzzy sets)

Contextual Knowledge Representation

- Ontologies may be described as:

  O : an ontology

  C : set of concepts it describes and

  : semantic relation amongst two concepts

- Define ***ontological context*** in the means of fuzzy taxonomic ontological relations:

  : a "fuzzified" ontology, : fuzzy relation

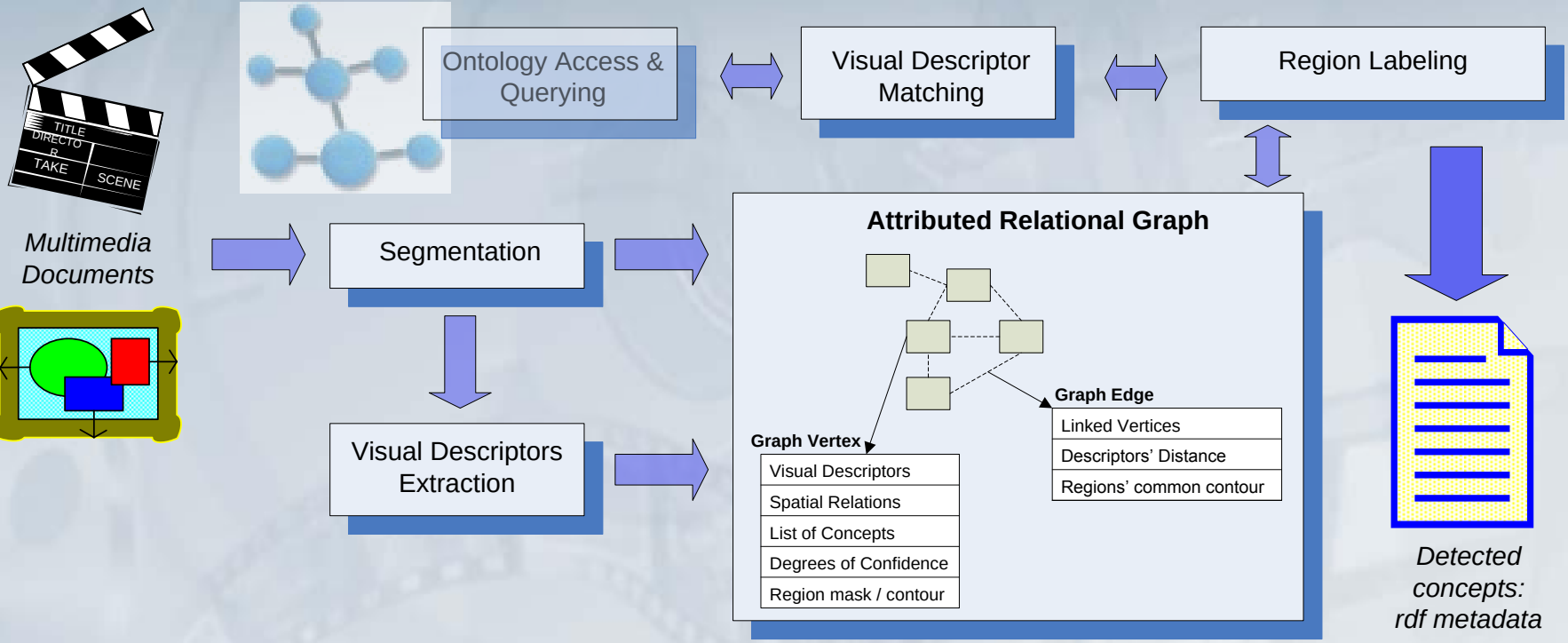
  C : set of all possible concepts it describes and

  denotes a fuzzy relation amongst two concepts

Attributed Relational Graph (ARG)

- Graph structure holds region-based representation of an image during analysis:
  - A graph's vertex represents a segment/region, where both visual (MPEG-7 descriptors, region mask, contour, etc.) and semantic information (region's candidate labels) are stored
  - A graph's edge represents the link between two regions, holding the overall neighboring information (spatial relations)
- Why use graphs?
  - Good for representation of structured objects
  - Image analysis problems can be considered as graph theory problems, inheriting their solid theoretical grounds

Initial Region Labeling



Semantic Segmentation

Novelty relies on:

- Modify traditional (region-based) segmentation algorithm to work on the regions' fuzzy sets of labels, stored in the ARG and not on visual features only
- Segmentation algorithm independent (!?) So far:
  - RSST (modify regions' distance/similarity)
  - Watershed (modify dam's height)

… more precisely … Semantic Region Growing?

Semantic RSST

- Similar to its traditional counterpart:
  - Calculation of distance/similarity between neighbour regions:

$$w(e_{ab}) = 1 - s_{ab}$$

$$s_{ab} = \sup_{c_k \in C}(t\text{-}norm(\mu_a(c_k), \mu_b(c_k))),\ a,b \in G$$

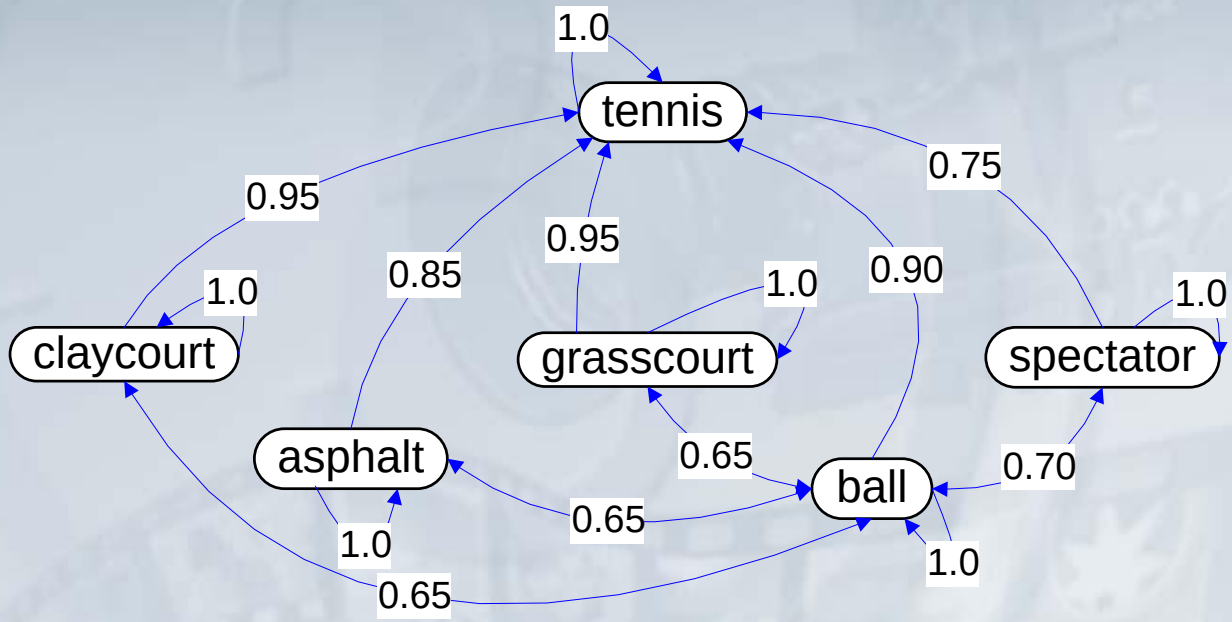
  - Find the edge with the least weight: $w(e^*)$
  - Remove edge e* and update ARG appropriately:
    - Re-evaluate degrees of membership for fuzzy sets of labels
  - Termination criteria: $w(e^*) > T_w$

Semantic Watershed

- Use of markers: Regions with only one dominant label of high confidence
- "Flooding" of neighbour regions (w.r.t. ARG)
  - Regions' semantic similarity sets the dum's height
  - "Wave's" height decreases as moving away from its source
- Flooded regions are merged and set to same fuzzy set of labels
- Initiate a new round of flooding, for the remaining regions (with new region markers)

Context-based Confidence Value Readjustment

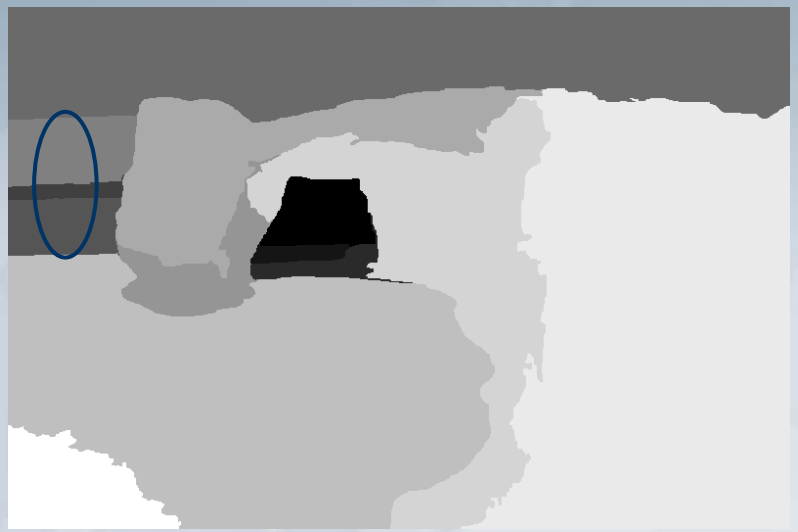
- Readjust initial region labels and confidence values
- Utilize a priori constructed contextual ontological information



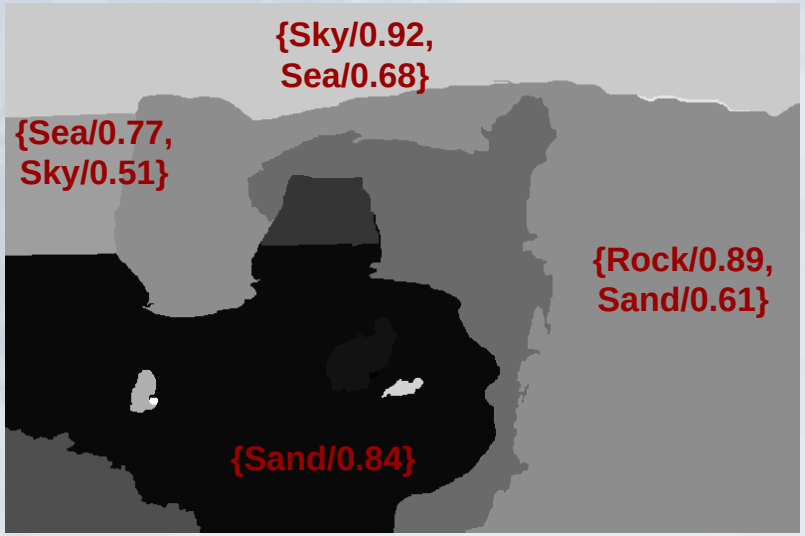
- Use *context relevance* cr_{c_k} to tackle cases that more than one concept is related to multiple concepts
  - *Max* used as compatibility indicator

Indicative Results (I)

RSST

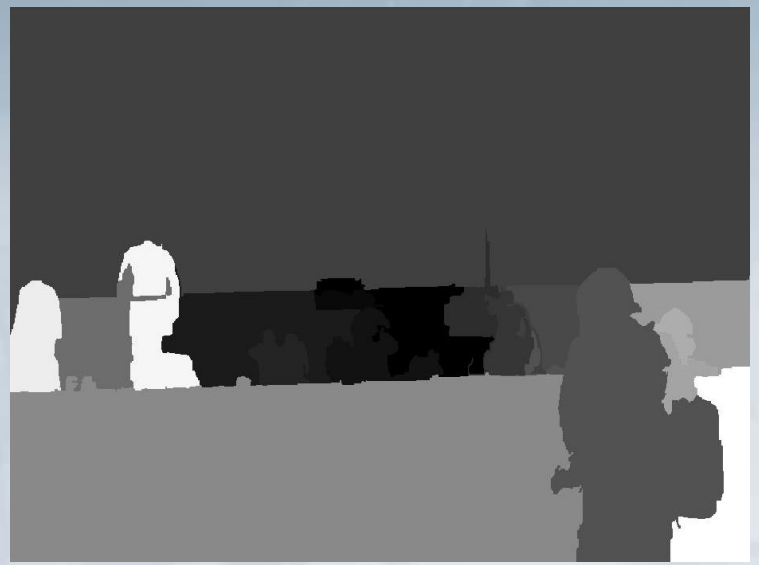


Semantic RSST

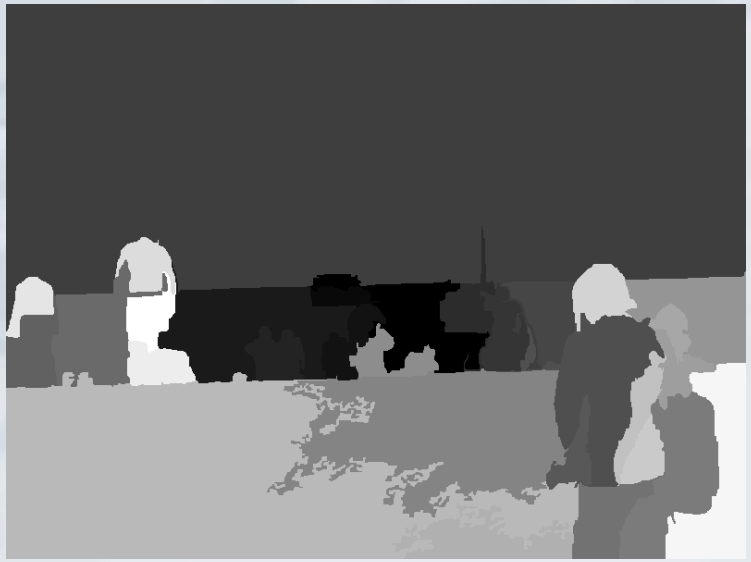


Indicative Results (II)

RSST



Semantic RSST



Thank you for your attention! ☺

Questions

Contact details:
Thanos Athanasiadis
thanos@image.ntua.gr
http://www.image.ntua.gr/~thanos