



Semantic Clustering and Querying on Heterogeneous Features for Visual Data

G. Sheikholeslami, W. Chang, A. Zhang
Department of Computer Science
State University of New York at Buffalo

Presented by Yuan Ji



Outline

- Introduction
- Image Database Organization
- Clustering on Heterogeneous Features
- Query on Heterogeneous Features
- Merging Heterogeneous Features

11/12/00

COMPUT 695 Presentation

2



Introduction

Visual data can be considered to contain feature vectors representing the content of the data.

Different features may have different similarity measurements.

The feature vectors of some semantically irrelevant images may be located very close in the feature space.

SemQuery overcomes the problems of traditional distance-based indexing and retrieval approaches

11/12/00

COMPUT 695 Presentation

3



Introduction

Visual data can be considered to contain feature vectors representing the content of the data.

Different features may have different similarity measurements.

The feature vectors of some semantically irrelevant images may be located very close in the feature space.

SemQuery overcomes the problems of traditional distance-based indexing and retrieval approaches

11/12/00

COMPUT 695 Presentation

4



Introduction

Visual data can be considered to contain feature vectors representing the content of the data.

Different features may have different similarity measurements.

The feature vectors of some semantically irrelevant images may be located very close in the feature space.

SemQuery overcomes the problems of traditional distance-based indexing and retrieval approaches



Introduction

Visual data can be considered to contain feature vectors representing the content of the data.

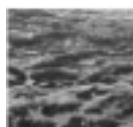
Different features may have different similarity measurements.

The feature vectors of some semantically irrelevant images may be located very close in the feature space.

SemQuery overcomes the problems of traditional distance-based indexing and retrieval approaches



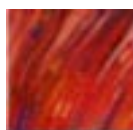
(a)



(b)

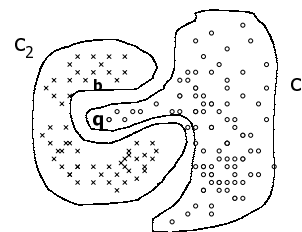


(c)

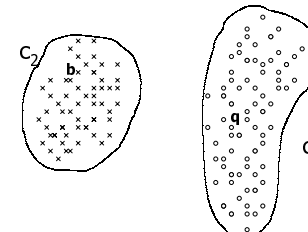


(d)

Semantically different images with similar feature vectors



(a)



(b)

Views of two semantic clusters in two different feature spaces

Introduction

Visual data can be considered to contain feature vectors representing the content of the data.

Different features may have different similarity measurements.

The feature vectors of some semantically irrelevant images may be located very close in the feature space.

SemQuery overcomes the problems of traditional distance-based indexing and retrieval approaches.

Image Database Organization

System architecture for using heterogeneous features in image retrieval

- Extract content of query image
- Compare the features of query image to database image
- Merge the results from individual feature classes

Clusters and templates

Hierarchy of clusters and templates

SemQuery system architecture

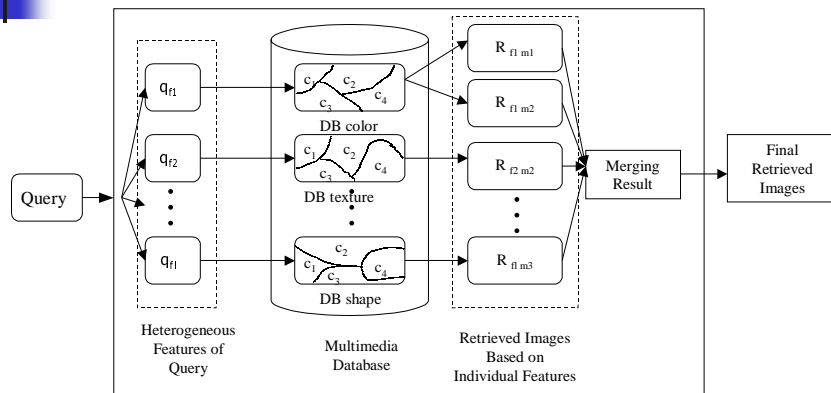


Image Database Organization

System architecture for using heterogeneous features in image retrieval

Clusters and templates

Templates : set of feature vectors that represent each cluster of the images

Hierarchy of clusters and templates

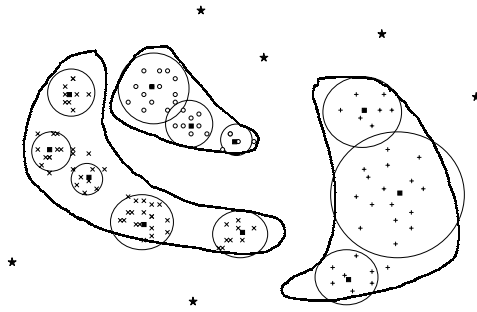
Image Database Organization

System architecture for using heterogeneous features in image retrieval

Clusters and templates

Hierarchy of clusters and templates

- Application
- Semantic cluster
- Feature class
- Template

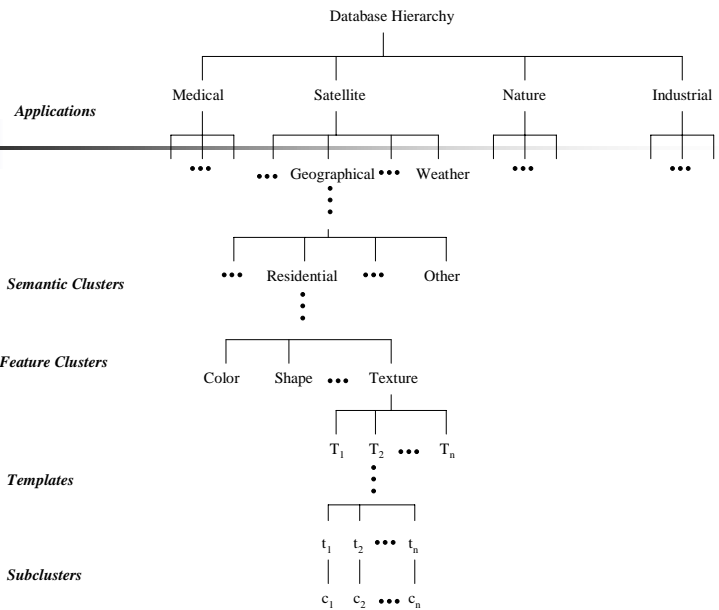


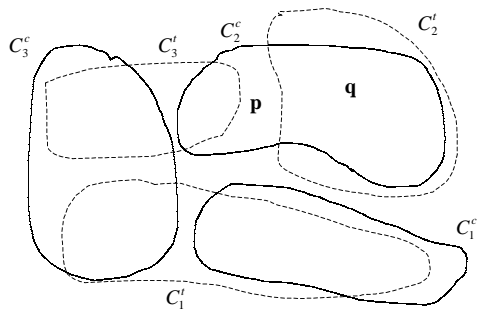
Clustering on Heterogeneous Features

A semantic cluster of each feature class is composed of a set of sub-clusters and its scope is the union of the scopes of the sub-clusters.

The scope of a semantic cluster based on heterogeneous features will be the intersection of the scopes of the clusters of the heterogeneous feature classes

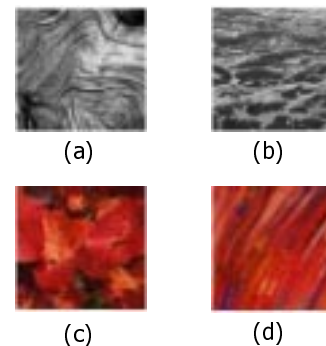
An image will be grouped into a semantic cluster if its heterogeneous features fall within the scope of all the cluster.





----- Texture cluster
 _____ Color cluster

Set representation of image clusters based on color and texture feature classes



Semantically different images with similar feature vectors

Query on Heterogeneous Features

Keyword describes the feature of interest, application domain, or semantics of the query.

With keyword matching the name of a semantic cluster, only images belonging to the specified semantic cluster will be considered in the retrieving process.

Without keyword, the retrieving process will be based on the image features and it will find the matched semantic clusters for the query.

Merging Heterogeneous Features

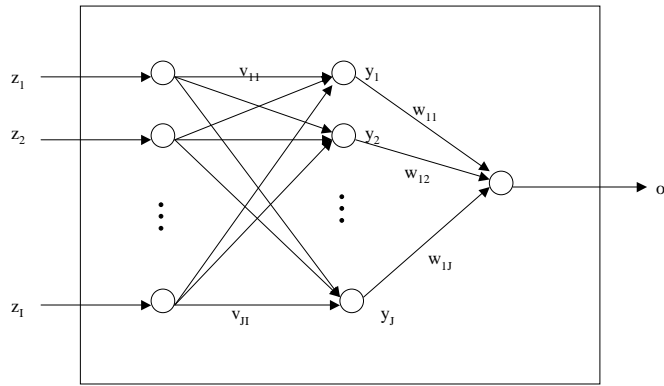
The ranking process must assign weights to the results obtained from the sub-queries.

Linear merging.
$$\sum_{i=1}^l w_i z_i$$

NeuroMerge.

input: set of measurements z_i

output: 0-1



Structure of Neural Network



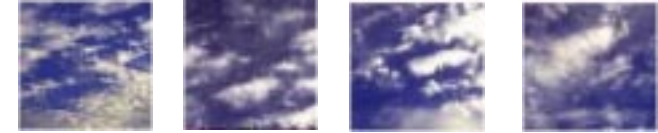
Performance



Query image



Retrieved images using nearest neighbor retrieval



Retrieved images from semantically clustered database