Pattern Recognition Lecture
“Template Matching”

Prof. Dr. Marcin Grzegorzek

Research Group for Pattern Recognition
www.pr.informatik.uni-siegen.de

Institute for Vision and Graphics
University of Siegen, Germany

Overview

Overview

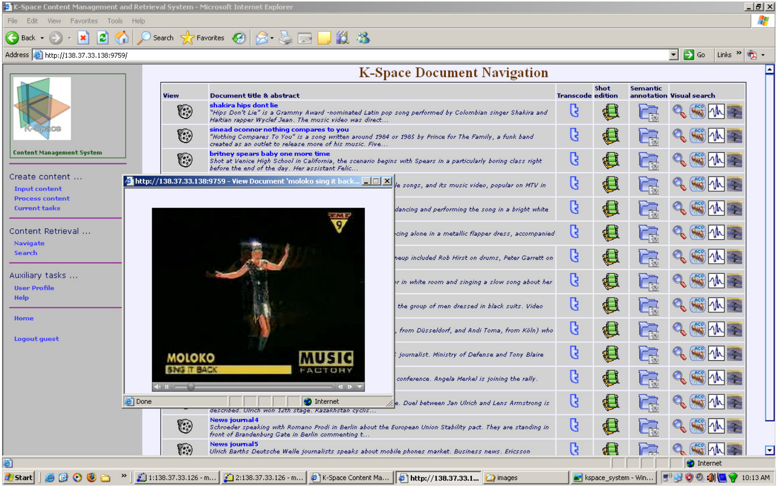
Introduction

- In previous lectures, the major concern was to assign an unknown pattern to one of the possible classes.

- Now, we assume that a set of reference patterns is available to us, and we have to decide which one of these reference patterns matches best the unknown pattern (test pattern).

- A reasonable first step in approaching such a task is to define a measure or a cost measuring the distance or the similarity between the known reference patterns and the unknown test pattern.

An Example Tool for Image Similarity Measure



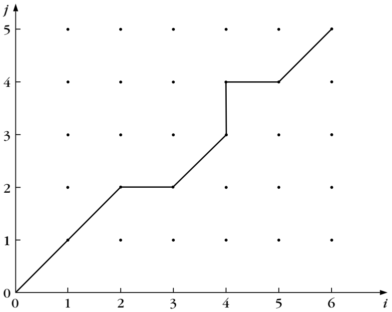
Overview

Introduction

- Here, the focus is on a category of template matching, where the involved patterns consist of strings of identified symbols or feature vectors (string patterns).
- Each of the reference and test patterns is represented as a sequence (string) of measured parameters and one has to decide which reference sequence matches best the test pattern.
- Let $\mathbf{r}(i), i = 1, \ldots, I$ and $\mathbf{t}(j), j = 1, 2, \ldots, J$ be the respective feature vector sequences for a specific pair of reference and test patterns. In general $I \neq J$.
- The objective is to develop an appropriate distance measure between the two sequences.

Approach in General (1)

- We form a two-dimensional grid with the elements of the two sequences as points on the respective axes. Example for $I = 6$ and $J = 5$ looks like follows:



- Each node (i, j) is associated with a cost (distance) $d(i, j)$.

Approach in General (2)

- The path from the initial to a final node is an ordered set of nodes

$$(i_0, j_0), (i_1, j_1), (i_2, j_2), \ldots, (i_f, j_f)$$

- Each path is associated with an overall cost D defined as

$$D = \sum_{k=0}^{K-1} d(i_k, j_k) \equiv D(i_k, j_k); \qquad D(0,0) = 0$$

  where K is the number of nodes along the path.

- The path is complete if $(i_0, j_0) = (0, 0); (i_f, j_f) = (I, J)$.

Approach in General (3)

- The distance between the two sequences is defined as **the minimum D over all possible paths**.
- At the same time, the minimum cost path unravels the pairwise correspondence between the elements of the two sequences.

Overview

Introduction

- The major task to be addressed in this section can be summarised as follows: "Given a block of recorded data, find whether a specific known reference pattern is contained within the block and where it is located."
- A typical application of this is found in scene analysis, when we want to search for a specific objects within the image.

Approach in General (1)

- Given are a reference pattern expressed as an $M \times N$ image array $\mathbf{r}(i, j)$ and $I \times J$ image array $\mathbf{t}(i, j)$, where $M \leq I$ and $N \leq J$.
- The goal is to develop a measure for detecting an $M \times N$ subimage within $\mathbf{t}(i, j)$ that matches best the reference pattern $\mathbf{r}(i, j)$.

Approach in General (2)

- The reference image $\mathbf{r}(i,j)$ is superimposed on the test image $\mathbf{t}(i,j)$ and it is translated to all possible positions (m,n).
- For each of the points (m,n), the mismatch between $\mathbf{r}(i,j)$ and the $M \times N$ subimage of $\mathbf{t}(i,j)$ is computed according to

$$D(m,n) = \sum_{i=m}^{m+M-1} \sum_{j=n}^{n+N-1} |t(i,j) - r(i-m,j-n)|^2$$

- The template matching algorithm looks for the location (m,n) for which $D(m,n)$ is minimum.

Overview

Introduction

- By now we have been looking for the perfect match between the reference and the test pattern.
- However, there are many problems where we know a priori that the available template and the object we search for in the image may not look exactly the same (remember the demo with the system for sketch-based image retrieval).
- Our goal here is to allow the template matching procedure to account for deviations between the reference template and the corresponding test pattern in the image.
- Thus, we will focus on shape information only.

Approach in General (1)

- The basic idea is simple: Deform the prototype and produce deformed variants of it.
- From a mathematical point of view a deformation consists of the application of a parametric transform T_ξ on $\mathbf{r}(i,j)$.
- Different values of ξ lead to different versions.
- From the set of the deformed prototype variants that can be generated, there will be one that best matches the test pattern.

Approach in General (2)

- The goodness of fit is measured via a cost which is called the matching energy $E_m(\xi)$.
- The cost measuring the deformation, which the prototype needs to undergo in order to fit the test pattern is called the cost deformation energy $E_d(\xi)$.
- The optimal vector parameter ξ is chosen so that the best trade-off between these two energy terms is achieved.

Overview

General about CBIR

- The more traditional way of information retrieval is text-based; stored information is manually annotated by text descriptors.
- In CBIR, stored information is indexed and searched based on its content.

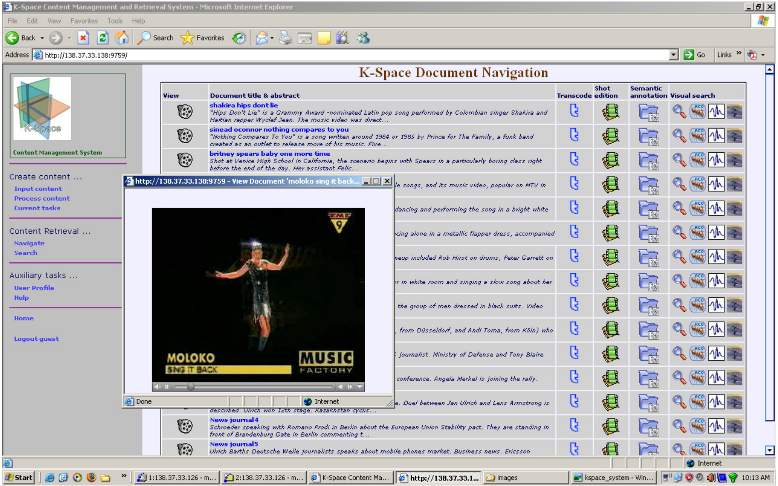
A Popular Metric for CBIR

- A popular metric that has extensively been used for CBIR is

$$d(\mathbf{x}, \mathbf{y}) = \left(\sum_{i=1}^{l} \omega_i |x_i - y_i|^p \right)^{\frac{1}{p}}$$

- Obviously, for $p = 2$ and $\omega_{i=1,2,\ldots,l} = 1$ this becomes the Euclidean distance and for $p = 1$ the so called weighted l_1 (Manhattan) norm.

Content-Based Video Retrieval System



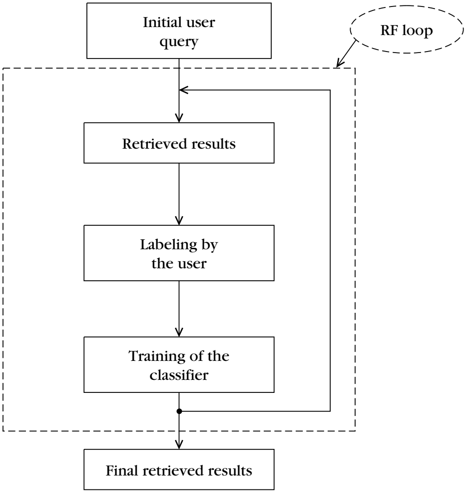
Disadvantages of CBIR Systems

- Search and retrieval are based on low-level features.
- Humans, being much more intelligent that the machines, utilise a number of so called high-level concepts when they recognise objects.
- This discrepancy is called semantic gap.

Relevance Feedback in CBIR - Intro

- The mentioned problem with the semantic gap can partly be solved by involving the human into the retrieval process.
- The search/retrieval session is divided into a number of consecutive loops.
- At every loop, the user provides feedback regarding the results by characterising the retrieved patterns as either relevant or irrelevant.

Relevance Feedback in CBIR - a Typical Scenario

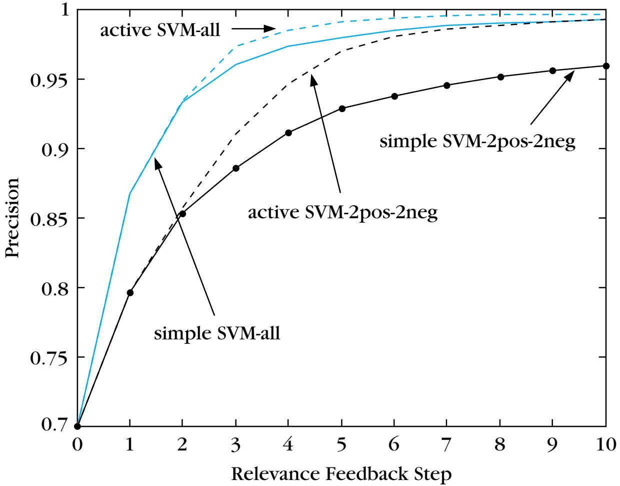


Evaluation of the CBIR Systems

- **Precision** is the ratio of relevant patterns to the total number of patterns in the set of returned patterns P_r.

- **Recall** is the ratio of returned relevant patterns to all relevant patterns in the database.

Evaluation of Different Strategies



Overview