# Analysis of Object Features in Terms of the Dissimilarity of Pattern Recognition

**Tatiana Jaworska**
Systems Research Institute, Polish Academy of Sciences
Tatiana.Jaworska@ibspan.waw.pl

**Abstract**

This article analyses object features in terms of their dissimilarity for a content-based image retrieval (CBIR) system. The first part is devoted to the mathematical fundamentals of recognition procedures. The main part serves the survey of low-level object features in relation to our CBIR system, which is dedicated to supporting estate agents. In the database there are images of houses and bungalows. Hence, the dissimilarity measures are scrutinized with reference to architectural object features. We present a brief overview of (dis)similarity measures for various types of object features, together with their characteristics. At the middle-level image analyses based on spatial relations must also employ comparisons to evaluate the object spatial localization. At this level the spatial similarity gives crucial information to respond to the query concerning the best matching of whole images. With the aim of modelling dissimilarity judgments some tree structures are introduced. Therefore, the approach to the dissimilarity in this CBIR system is hybrid. The basic experimental research concentrates on examining dissimilarities among house images.

**Keywords**: dissimilarity measure, content-based image retrieval system, feature space.

## 1 Introduction

Pattern recognition is both art and science. For the purpose of automatic grouping and classification it is difficult to determine proper features. The formalized representation of objects (usually in mathematical terms) and the definition of classes determine how the object recognition level should be modelled.

While many papers are concerned with various algorithmic procedures, we would like to focus on the issue of representation. This work is devoted to par-

Developments in Fuzzy Sets, Intuitionistic Fuzzy Sets, Generalized Nets and Related Topics. Applications. (K. Atanassov, P. Chountas, J. Kacprzyk, M. Krawczak, P. Melo-Pinto, E. Szmidt, S. Zadrożny, Eds.), EXIT, Warsaw, 2008.

ticular representations, namely dissimilarity representations. Below we will examine the nature of basic problems in pattern recognition (see fig. 1) with the use of dissimilarity representations.

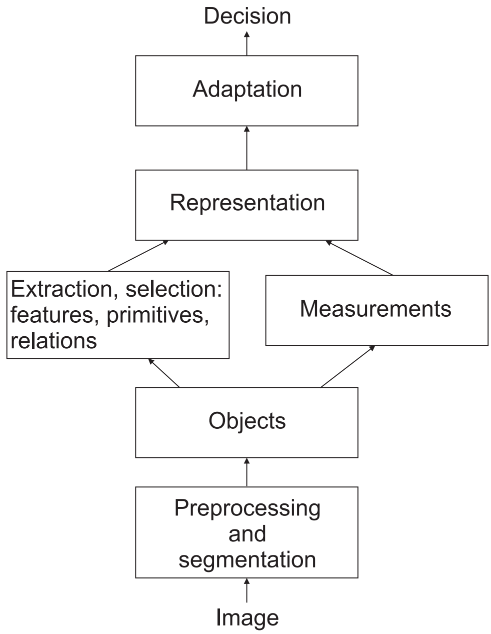


Figure 1: Components of a general pattern recognition system. A representation is either a numerical description of objects and/or their relations or their syntactical encoding by a set of primitives together with a set of operations on objects and spatial relationships. Adaptation relies on a suitable change of a representation, e.g. by reducing the number of features, relations or primitives describing objects, or some nonlinear transformations of the features, to enhance the class or cluster descriptions

There are two principal groups of methods in pattern recognition, statistical and structural (or syntactic) [18], [23], [3]. Both approaches use features to describe objects, but these features are defined differently. In general, features are functions of measurements performed on objects.

## 1.1. CBIR conception overview

Pattern recognition is applied as one of the methods for image content analysis in image retrieval systems (CBIR). Our CBIR system is dedicated to support estate agents. In the estate database there are images of houses, bungalows, and other buildings. To be effective in terms of presentation and choice of houses, the system has to be able to find the image of a house with defined architectural elements, for example: windows, roofs, doors, etc [19]. This work uses pattern recognition as an element of content-based image retrieval system (CBIR). For this purpose it is difficult to determine proper features, i.e. shape descriptors that would precisely discriminate among many different elements, in our case architectural elements. Figure 2 shows the block diagram of our CBIR system. The hierarchical structure of the *image content analysis* block from the segmentation level to the object recognition level covers the present approach to a multi-layer image description model in order to reach progressive image analysis and understanding (Gao [13]). In this structure, image content is analyzed and represented on three levels. There is a context between adjacent levels, i.e. the representation for the upper level is directly extracted from the lower level.

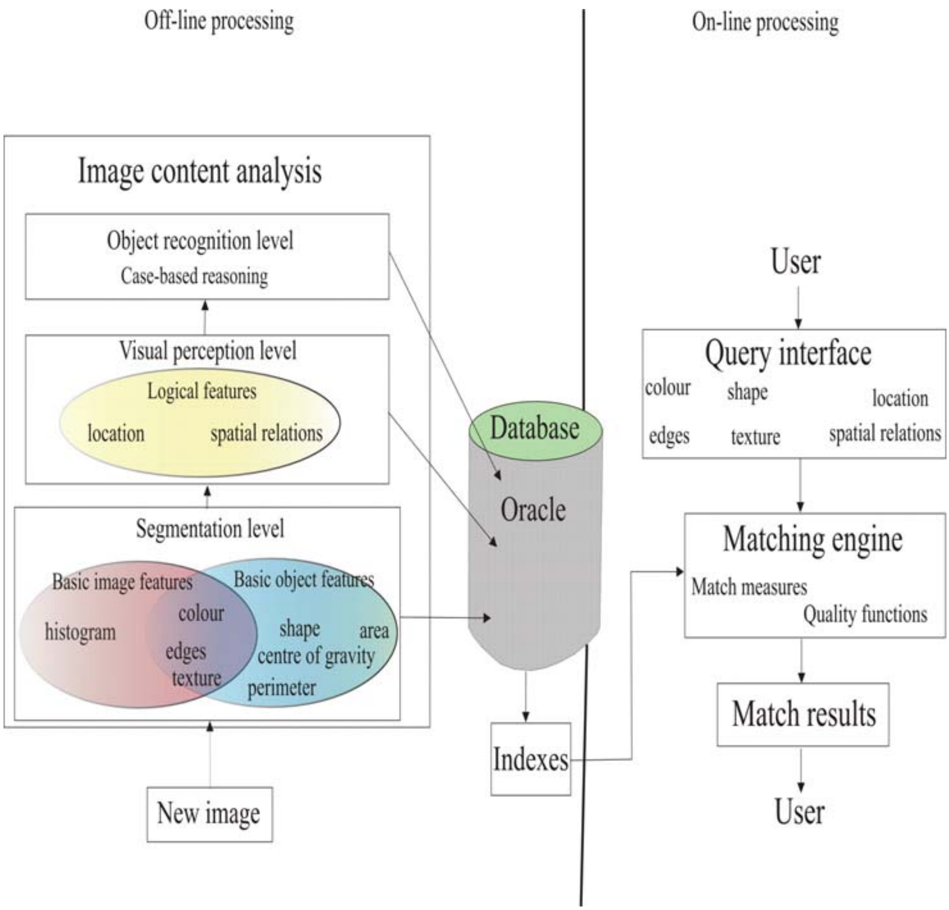


Figure 2: Block diagram of our content-based image retrieval system

The first stage of our analysis amounts to splitting the original image into several meaningful clusters; each of them provides certain semantics in terms of human understanding of image content. Next, separate objects are extracted from these clusters and our efforts have been put into extracting elements from an image in the unsupervised way. With reference to the processes described above the selection of image features should be carefully fitted into dissimilarity methods in terms of pattern recognition. Another important part of the *image content analysis* block is the object recognition level logically preceded by the visual perception level.

New images (in JPEG format), as well as all results achieved on each level of *image content analysis,* are stored in the related database. For our project the Oracle DB system has been chosen on account of its flexibility and completeness. In order to enforce the process of image retrieval many multi-dimensional indexes have to be implemented for this database. Some of them are proposed by Amsaleg and Gros [1] and also by Głomb [14].

The right half of the diagram is dedicated to users, hence, it is processed on-line. It comprises the user's query interface and the matching engine. In our system the user's interface is offered on the WWW site. Users can select separate features such as colour or texture, they can also select an architectural element, which automatically defines many features as a query. Indeed, if the users have a vague target image in mind, they have the possibility to compose their imaginary house and the system presents them some optional houses based on these chosen elements [10].

The next element of the system is the matching engine which uses the dissimilarity measures to search for best-matched images. A careful image content analysis is crucial for the whole on-line process dedicated to the user, otherwise, the proper matching result will not be reached.

## 2 Dissimilarity measures

Pattern recognition relies on the description of regularities in the observed classes of objects. Each object is described by a vector of features. We assume that there exists a resemblance between dissimilarity and feature-based representations (e.g. in their matrix notation).

The main issue is understanding of the structure and complexity of a dissimilarity data representation. Relative similarity can be defined as a relationship among entities which are of the same nature or possess the same characteristics, but to a different degree or measure. The larger the similarity value, the greater the resemblance among objects. Relative dissimilarity, on the other hand, focuses on the differences; the smaller the dissimilarity, the more alike the objects are.

The dissimilarity space approach addresses a dissimilarity representation as a data-dependent mapping specified by the representation set $\mathcal{R}$. A mapping $\psi(\cdot,\mathcal{R}) : \chi \rightarrow \mathbf{R}^n$ is defined as $\psi(x,\mathcal{R}) = [d(x, p_1)\ d(x, p_2)\ \ldots\ d(x, p_n)]$. We assume a collection of objects $\mathcal{R} = \{p_1, p_2, \ldots , p_n\}$, called a representation set, or a set of prototypes and $d$ is dissimilarity. Note that denotes either objects themselves (e.g. a set of convex subsets of a finite-dimension space), or a feature-based vectorial representation of objects. Note that might not be given explicitly. The dimension of such a space is controlled by the cardinality of $\mathcal{R}$.

## 2.1 Measures depending on feature types

The two determining factors for image retrieval performance are the features used to represent the images and the distance function used to measure the similarity between a query image and the images in the database. For a specific feature representation chosen, the retrieval performance depends critically on the similarity measure used. We distinguish the following feature types: binary, categorical, ordinal, symbolic and quantitative.

***Definition 1. (Feature types)***

Let $=\{f_1, f_2,\ldots, f_m\}$ be a set of features, also called variables or attributes, and $D_f$ - a set of valid values for a feature $f$. The following features $f$ can be considered:

- *binary* if $D_f$ is a set of two numbers. They represent either the presence (1) or absence (0) of particular characteristics or some opposite qualities.
- *categorical* if $D_f$ is a finite, discrete set of numbers, e.g. from 1 to 3 to encode RGB triples. Here, we also include the case of a *discrete* feature, i.e. a feature with distinct and separate values, which can be counted, such as grey level degrees from 0 to 255, etc.
- *quantitative* if $f$ is measured on an interval and $D_f$ is a convex subset of $\mathbf{R}$, e.g. an area of an object in pixels.
- *original* if $D_f$ is a finite, discrete set of ordered symbols, e.g. a scale from 1 to 5 representing the answers of ‘strongly dislike’, ‘dislike’, ‘neutral’, ‘like’, and ‘strongly like’, after comparing CBIR results with an imaginary house by a potential client.
- *symbolic* or *nominal* if $D_f$ is a finite, discrete set of symbols; e.g. nationality. Symbolic features represent a set of possible values, symbols or modalities. Their values can be counted, but not ordered.

**Measures for binary features**

Let $i$-th object be represented by a binary vector $\mathbf{x}_i \in \boldsymbol{B}^m$, where $\boldsymbol{B}=\{0,1\}$.

For $\quad \mathrm{x}_i\, \mathrm{x}_j \in B^m, \qquad \mathrm{x}_i^{\mathbf{T}} \mathrm{x}_j = \sum_{k=1}^{m} x_{ik} x_{jk} \qquad (1)$

is the binary scalar product and $(1 - \mathbf{x})$ is the complementary vector of $\mathbf{x}$. This allows us to define the following counters:

- $a_{ij} = \mathbf{x}_i^{\mathbf{T}} \mathbf{x}_j$ - the number of properties common for both objects
- $b_{ij} = \mathbf{x}_i^{\mathbf{T}} (1 - \mathbf{x}_j)$ - the number of properties which $i$ has and $j$ lacks
- $c_{ij} = (1 - \mathbf{x}_i)^{\mathbf{T}} \mathbf{x}_j$ - the number of properties which $j$ has and $i$ lacks
- $d_{ij} = (1 - \mathbf{x}_i)^{\mathbf{T}} (1 - \mathbf{x}_j)$ - the number of properties that both objects lack

where $a_{ij} + b_{ij} + c_{ij} + d_{ij} = m^2$. For various definitions of similarity measures, a 2×2 contingency table is considered for each pair of objects $i$ and $j$ as presented in Tab. 1.

Table 1 Counters for binary features

| | | Object $j$ 1 | 0 |
|---|---|---|---|
| Object $i$ | 1 | $a_{ij}$ | $b_{ij}$ |
| | 0 | $c_{ij}$ | $d_{ij}$ |

**Measures for categorical features**

Let $X$ be a categorical $n \times m$ data matrix and let the feature $f_k$ take values in $c_k$ categories, so that $c = \sum_{k=1}^{m} c_k$ . Dissimilarity measures defined for binary data, Tab. 1 can now be adapted for the categorical data, as well. To achieve that, one has to code each $m$-dimensional data vector $\mathbf{x}_i$ into a $c$-dimensional binary vector $\tilde{\mathbf{x}}_i = [\tilde{\mathbf{x}}_{(1)},...,\tilde{\mathbf{x}}_{(m)}]$. $\tilde{\mathbf{x}}_{(k)}$ is a vector of length $c_k$ consisting of all zeros except for 1 at the $j$-th position assuming that $x_{ik}$ belongs to the $j$-th category [8].

**Measures for ordinal features**

Let $X$ be an ordinal $n{\times}m$ data matrix, so that the feature $f_k$ has $c_k$ categories, so that $c = \sum_{k=1}^{m} c_k$ . In case of ordinal variables, the dissimilarity measure should take into account the positions of categories in the ordering, and it should be larger for more distant categories than for the close ones. Here, a generalization of the Jaccard dissimilarity can be used for a comparison of the objects $p_i$ and $p_j$, as follows:

$$d(p_i, p_j) = \frac{\sum_{k=1}^{m} x_{ik} + \sum_{k=1}^{m} x_{jk} + 2\sum_{k=1}^{m} \min(x_{ik}, x_{jk})}{\sum_{k=1}^{m} x_{ik} + \sum_{k=1}^{m} x_{jk} - \sum_{k=1}^{m} \min(x_{ik}, x_{jk})} \tag{2}$$

**Measures for quantitative features**

Many measures exist for quantitative variables, mostly constructed in an additive way after counting the differences for each variable separately [8], [9], [15], [16]. Some of them are presented in Tab 2. The basic measures come from the family of $l_p$-distances. The $l_p$ metric, for p ≥ 1 is defined as

$$d_p(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|_p = \left[\sum_{i=1}^{m} (x_i - y_i)^p\right]^{1/p} \tag{3}$$

which for $p$=1 becomes the city block distance and for $p$=2, the Euclidean distance.

Table 2 Dissimilarity measures for quantitative data in $\mathbb{R}^m$.

| Kind of distance | D | Dissimilarity $d$ (**x**, **y**) |
|---|---|---|
| Euclidean | $D_E$, $D_2$ | $\sqrt{(\mathbf{x}-\mathbf{y})^{\mathrm{T}}(\mathbf{x}-\mathbf{y})}$ |
| Weighted Euclidean | $D_{wE}$ | $\sqrt{(\mathbf{x}-\mathbf{y})^{\mathrm{T}}\,\mathrm{diag}(w_i^2)(\mathbf{x}-\mathbf{y})}$ |
| City block | $D_1$ | $\sum_{k=1}^{m} \lvert x_i - y_i \rvert$ |
| Max norm | $D_{max}$ | $\max_i \lvert x_i - y_i \rvert$ |
| $l_p$ or Minkowski | $D_p$ | $\left[\sum_{i=1}^{m} (x_i - y_i)^p\right]^{1/p}$, $p \geq 1$, $p \neq 2$ |
| Cosine | $D_{cos}$ | $\frac{1}{2}\left(1 - \frac{\mathbf{x}^{\mathrm{T}}\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}\right)$ |

### Measures for symbolic data

Symbolic objects are described by *m* variables $f_i$, each on the domain $D_{f_i}$ and a logical statement of the form $[f_i \in \chi_i]$, where $\chi_i \subseteq D_{f_i}$, e.g. [colour ∈ {red, green, blue}] or [texture ∈ (20, 40)]. A symbolic object *x* is expressed as the Cartesian product of the values $x_i = f_i(x)$ with the total event being a conjunction of all the feature events. The dissimilarity between two objects $x = [f_1 \in \chi_1] \ \ldots \ [f_m \in \chi_m]$ and $y = [f_1 \in \ _1] \ \ldots \ [f_m \in \ _m]$ can be defined with respect to the components due to position ($d_p$), span ($d_s$) and content ($d_c$), all normalized to [0,1], as [18]:

$$d(x,y) = \sum_{i=1}^{m} [d_p(x_i - y_i) + d_s(x_i - y_i) + d_c(x_i - y_i)] \tag{4}$$

The component $d_p$, valid for quantitative variables only, indicates the relative positions of two variable values. By writing $\chi_i = [x_i^l, x_i^u]$ and $\gamma_i = [y_i^l, y_i^u]$ with the lower $x_i^l$ and upper $x_i^u$ limits, one has

$$d_p(x_i, y_i) = \frac{\left|x_i^l - y_i^l\right|}{\left|D_{f_i}\right|} \tag{5}$$

where $\left|D_{f_i}\right|$ is the range of $f_i$ over all the objects. The remaining two measures, $d_s$ and $d_c$ are defined for quantitative, symbolic or ordinal attributes.

## 2.2 Dissimilarity measures between sets

Dissimilarities can also be considered between two closed and bounded subregions of an Euclidean space, sets of points or elements. Let us first formally introduce the Hausdorff distance.

### Hausdorff metric

Let (X,ρ) be a metric space and C(X) ⊆ X be a space of non-empty, closed and bounded subsets of X. Let $N_\varepsilon(A) = \bigcup_{x \in A} B_\varepsilon(x)$ be the cover of A ∈ X by open ε-balls $B_\varepsilon(x) = \{y \in X : \rho(x,y) < \varepsilon\}$, where A, B are two compact subsets of metric space X. Since Bε(x) is a neighbourhood of x, then Nε(A) is a neighbourhood of A according to the definition of natural topology in metric spaces. The Hausdorff distance between A and B is defined as the smallest ε-neighbourhood of A which covers B and the other way round (see fig. 3). On the other hand, the direct Hausdorff distance between A and B, $d_H^{\triangleright}(A,B)$ can be expressed as the maximum taken over from the collection of minimum distances between elements of the set A and B.

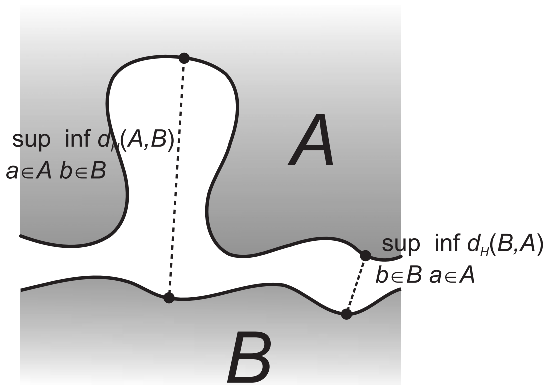


Figure 3: Illustration of the Hausdorff distance between sets A and B: $d\mathrm{H}\ (A,B) = \varepsilon$.

***Definition 2. (Hausdorff distance)***

In a metric space $(X, \rho)$, the Hausdorff distance with the base $\rho$ is defined for all $A, B$  $C(X)$ in one of the following ways:

(1) $d_H(A,B) = \inf\limits_{\varepsilon>0}\{A \subset N_\varepsilon(B)\ \&\ B \subset N_\varepsilon(A)\}$

(2) $d_H(A,B) = \max\{d_H^{\triangleright}(A,B), d_H^{\triangleright}(B,A)\}$, where $d_H^{\triangleright}$ is a directed Hausdorff distance $d_H^{\triangleright}(A,B) = \sup\limits_{a\in A}\inf\limits_{b\in B} \rho(a,b)$.

**Variants of the Hausdorff distance**

Let $(X, \rho)$ be a metric space (usually Euclidean) and $C(X)$  $X$ be a space of non-empty, closed and bounded subsets of $X$. Let $A, B$  $C(X)$ be sets of $n_A$ and $n_B$ elements correspondingly. The distance between an element $a$  $A$ and the set $B$ can be defined as:

$$d(a,B) = d(\{a\},B) = \min_{b\in B} \rho(a,b) \qquad (6)$$

The directed dissimilarity between two sets can be then found as [7]:

$$
\begin{aligned}
d_{\min}(A,B) &= \min_{a \in A} d(a,B) \\
d_{\max}(A,B) &= \max_{a \in A} d(a,B) \\
d_{avr}(A,B) &= \frac{1}{n_A} \sum_{a \in A} d(a,B)
\end{aligned}
\tag{7}
$$

A Hausdorff-like distance can also be defined for fuzzy sets [4, 5]. Consider two non-empty fuzzy sets $A_f$ and $B_f$ on support set $S$ in a metric space and a discrete set of the membership values $\mu_1$, $\mu_2$, …, $\mu_c$. Let $d_H(A_{\mu_i}, B_{\mu_i})$ be a fuzzy Hausdorff distance between sets $A_{\mu_i}$ and $B_{\mu_i}$. Then, the fuzzy Hausdorff-like distance between $A_f$ and $B_f$ is defined as:

$$
d_{H_f}(A_f, B_f) = \frac{\sum_{i=1}^{c} \mu_i d_H(A_{\mu_i}, B_{\mu_i})}{\sum_{i=1}^{c} \mu_i}
\tag{8}
$$

which is metric [5]. $d_{H_f}$ can be seen as membership-weighted average of the Hausdorff distances between the modified level sets of the fuzzy sets considered.

## 3 Feature survey

In terms of the CBIR system discussed in subsec. 1.1 and the above analysed dissimilarities the features of objects are examined here separately. An image/object feature is a distinguishing primitive characteristic or attribute of an image in general or an object in our case. There are two quantitative approaches to the evaluation of image features: a prototype performance and a figure of merit.

In the prototype performance approach for image classification, a prototype image with regions (segments) that have been independently categorized is classified by a classification procedure using various image features to be evaluated. The classification error is then measured for each feature set. The best set of features is, of course, that which results in the least classification error.

The figure of merit approach to feature evaluation involves the establishment of some functional distance measurements between sets of image features such that a large distance implies a low classification error and vice versa.

Below we will show selected examples of features useful for the CBIR.

### 3.1 Colour – second-order histogram

Second-order histogram features are based on the definition of the joint probability distribution of pairs of pixels. Let us consider two pixels $F(j,k)$ and $F(m,n)$ that are located at coordinates $(j,k)$ and $(m,n)$ respectively, and, as it is shown in Fig. 4 these pixels are separated by $r$ radial units at an angle $\theta$ with respect to the horizontal axis. The joint distribution of image amplitude values is then expressed as $P(a,b) = P_R\{ F(j,k) = r_a, F(m,n) = r_b \}$, where $r_a$ and $r_b$ represent quantized pixel amplitude values. As a result of the discrete rectilinear representation of an image, the separation parameters $(r,\theta)$ may only assume certain discrete values. The histogram which estimates the second order distribution is

$$P(a,b) \approx \frac{N(a,b)}{M} \tag{9}$$

where $M$ is the total number of pixels in the measurement window and $N(a,b)$ denotes the number of occurrences for which $F(j,k) = r_a$ and $F(m,n) = r_b$.

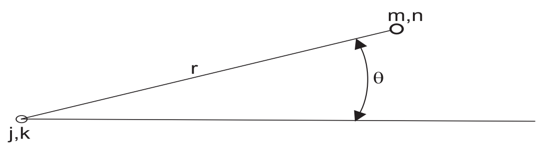


Figure 4: Relationship of pixel pairs

When we need a fast estimation of colour for an architectural element we can use the average colour of this element, computed separately as an average of each colour component of the RGB triple $(r,g,b)$.

### 3.2 Texture parameters

The texture information presented in images is one of the most powerful features available. There are many methods which can be used for texture characterization. One of them is the two-dimensional frequency transformation. For CBIR purposes, such methods could be applied as the classical Fourier transformation or several spatial-domain texture-sensitive operators, for instance, the Laplacian 3x3 or 5x5, the Gaussian 5x5, Hurst, Haralick, or Frei and Chen (Russ, 1995). Generally, all of them are useful for relatively small neighbourhoods.

Later methods are based on the transformation domain algorithms. In 2001 Balmelli and Mojsilović [2] proposed the wavelet domain for texture and pattern. They found features such as directionality, symmetry and regularity for regular textures, geometrical patterns and floral ornaments. Unfortunately, they have not proposed any application of their method for real images. So far only Lewis and Fauzi have managed to perform an automatic texture segmentation algorithm for CBIR based on Discrete Wavelet Transform (DWT). They have applied their method for image retrieval in museum collections [11].

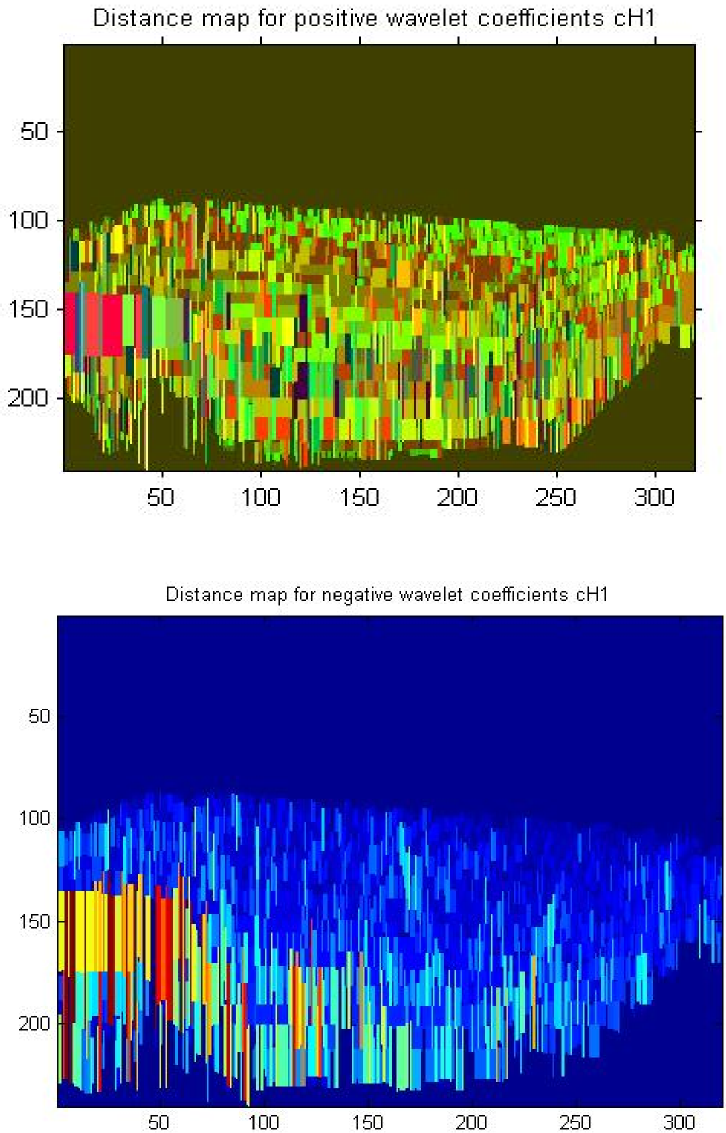


Figure 5: Distance map for positive and negative horizontal wavelet coefficients

In our work we decided to use the Fast Wavelet Transform (FWT) [6], [21], [22]. It is efficient and productive enough for frequent use for our purpose. One of the most important features of wavelet details is their directionality. It means that the convolution of the horizontal wavelet and the horizontal elements of an image results in big values of horizontal details $d_1^1$ , whereas the convolution of the vertical wavelet and the vertical elements of image results in big values of vertical details $d_1^2$ .

If we use this feature and compute the convolution of an image consisting of regular tiles or bricks and the relevant wavelet, we obtain a 2D transform whose maximum values are placed in the connection spots among these tiles or bricks. We have applied the Haar wavelet for all our counts because it is the simplest wavelet and it seems to be the most suitable kind of wavelet for analysis of geometrical elements. Having computed horizontal details, we measured the distances between the maxima for each column of this matrix as well as the distances between the minima for each column of this matrix. After counting the distances we have created two distance maps for all positive and negative horizontal coefficients (Fig. 5) [20].

### 3.3 Shape descriptors

In pattern recognition and CBIR shape (similar to colour and texture) is one of the primary low level features widely exploited. Therefore, many shape description techniques have been developed for both quantitative and qualitative measurements. There are generally two types of shape representation methods in the literature: the region-based and contour-based methods.

Here, we are interested in the comparison of objects, hence in measures of their similarity. For the purpose of matching of region-based shapes variants of Hausdorff distances can be used. These measures are, in fact, measures between sets of points. They are invariant to either affine transformations or similarity transformations of the sets.

The dissimilarity between image contours can be studied as a cost of matching by summing up the costs of local deformations that reflect the differences between two contours. A cost function is proposed, which depends on the local curvature and obeys the constraints of continuity, metric properties and invariance under some classes of transformations.

To be more specific, let us have a look at some kinds of shape descriptors. Shape can be described in a structural way. A chain code representation of a digital boundary as a sequence of direction vectors based on the 4- or 8-connectivity principle was proposed by Freeman and Class in 1961 [12] (see fig. 6).

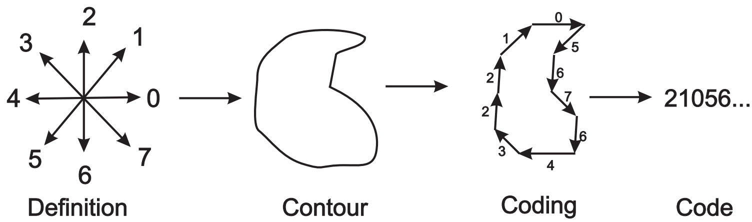


Figure 6: Chain code representation based on the 8-connectivity

This approach is unique and reconstructs a shape when a starting point is given. Unfortunately, chain codes become very long for complex objects, but more importantly, they reflect all the noise present on the boundary.

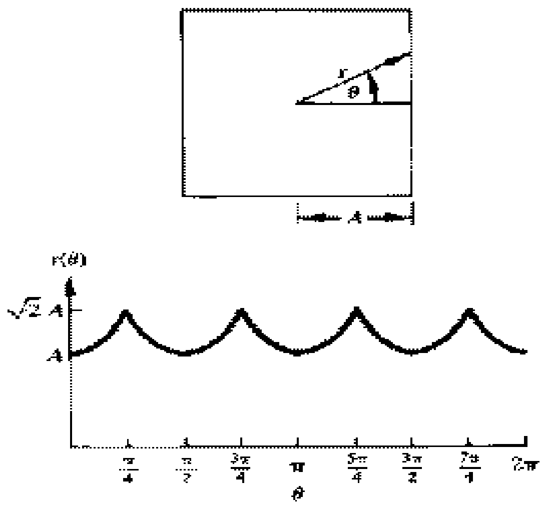

Figure 7: A square and its centroid distance function

Alternatively, shapes can be represented in transform spaces such as Fourier or wavelet space. In general, the set of normalized Fourier transformed coefficients is called the Fourier descriptor of shape. The shape signature is a one-dimensional function, which is derived from shape boundary coordinates. The first step in computing a Fourier descriptor is the obtainment of the boundary coordinates ($x(\theta)$, $y(\theta)$), $\theta = 0,1, \ldots, N-1$, where $N$ is the number of boundary points (see fig. 7). The centroid distance function $r(\theta)$

$$r(\theta) = \sqrt{[x(\theta) - \bar{x}]^2 + [y(\theta) - \bar{y}]^2} \tag{10}$$

is expressed by the distance of the boundary points from the centroid ($\bar{x}, \bar{y}$) computed as follows:

$$\bar{x} = \frac{1}{N}\sum_{\theta} x(\theta) \quad \text{and} \quad \bar{y} = \frac{1}{N}\sum_{\theta} y(\theta)\,. \tag{11}$$

An example of such a function for a square object is presented in figure 7.

Next, we transform function $r(\theta)$ and we obtain a set of Fourier coefficients $a_n$.

$$a_n = \frac{1}{N}\sum_{\theta=0}^{N-1} r(\theta)\exp\left(\frac{-j2\pi\, n\theta}{N}\right) \tag{12}$$

for $n = 0,1, \ldots, N-1$. The acquired Fourier coefficients are translation, rotation and scaling invariant after further normalization $\mathbf{a} = \{a_1, a_2, .., a_N\}$.

## 3.4 Spatial location

Other important criteria for assessing image similarity are: region treated as a location of objects, spatial information about objects in the whole image and spatial relationships among objects.

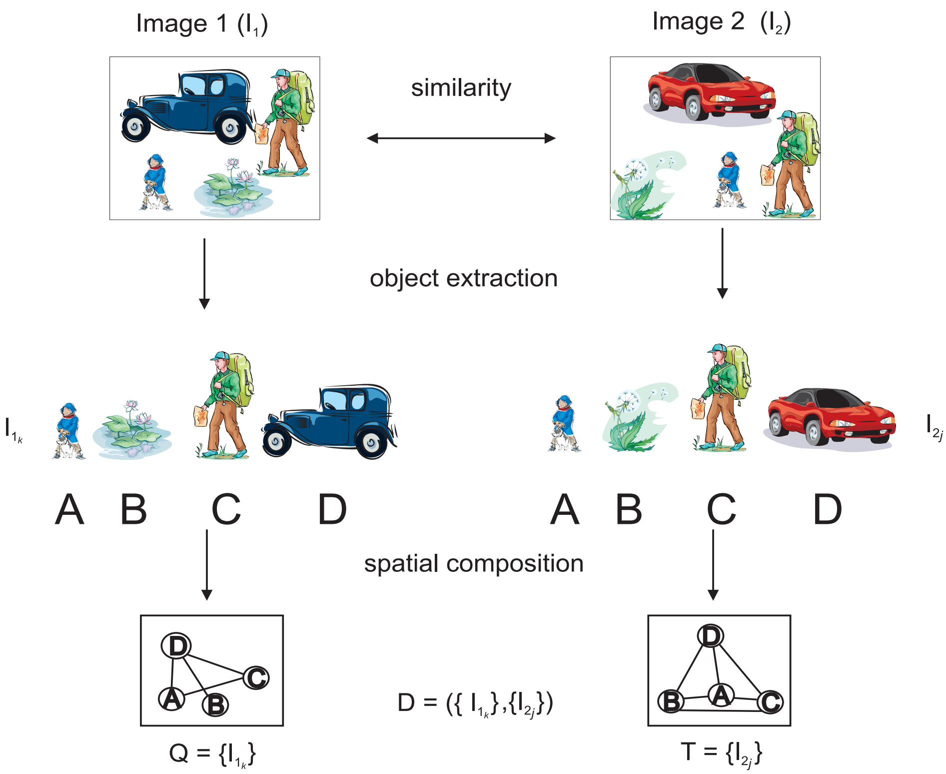


Figure 8: Integrated spatial query. Images are compared by the spatial arrangements of objects

This information is used in spatial image query [25]. In CBIR researchers have developed techniques for spatial image query based on the locations of objects [17], [24]. Two types of spatial image similarities can be recognize: relative and absolute. In a relative spatial image similarity, the objects are matched based upon the relative locations of each other. In absolute spatial image similarity the objects are matched based upon fixed positions in the images [25].

In our CBIR system objects and their features, as well as spatial attributes are extracted from the images. The comparison of the images then considers the similarity of the feature and spatial attributes of the regions, as depicted in Fig. 8.

The overall match score between images is computed by summing the weighted distances between the best matching objects in terms of spatial locations, sizes and features. The relative spatial locations of objects in the target images ($I_2$) are also compared to those in the query image ($I_1$) to determine the image matches. The image matching process is carried out on the matching objects tables. A query image Q = { $I_{1_0},...,I_{1_k}$ } consists of *k* objects. Each object has spatial attributes which are stored on a table.

To match the images, the CBIR system compares the objects from $I_1$ to the objects of the target images $I_2$ in the data base. At first, the system determines the sets of target regions that sufficiently match each query object $I_{1_k}$. After identifying the candidate target regions, the lists are combined.

Each target image and configuration of objects is assigned an overall match score to the query image. At the final stage, any specifications by the user of relative spatial constraints are checked for each target image.

## 4 Conclusions

This brief overview of similarity and dissimilarity measures indicates not only their variability, but also their different origins and underlying principles. The use of dissimilarity is popular in CBIR systems as a natural means for comparison of objects. Most of the dissimilarities are defined for the problem at hand, but, intuitively, it is clear that a good dissimilarity measure should be small for similar objects and large for different ones. Ideally, the measure should be developed such a way that it is invariant to rotation, translation and scaling and also to small aberrations and changes in the images.

The feature survey has been carried out with reference to the above presented dissimilarities. It has only focused on the most difficult features to determine. We have omitted some features which are easy to count such as area, moments of gravity, compactness, etc. because of their popularity in the literature.

## References


[1] Amsaleg L., Gros P., Berrani S-A. (2004). Robust Object Recognition in Images and the Related Database Problems. *Multimedia Tools and Applications*, 23, 1, Springer Netherlands, 221-235.

[2] Balmelli L., Mojsilović A. (2001). Wavelet domain features for texture/pattern description, classification and replicability analysis. In: A. Petrosian and F. Meyer, ed., *Wavelets in Signal and Image Analysis*. Kluwer Academic Publishers, London, Chap. 7, 193-212.

[3] Bunke H., Sanfeliu A., eds.(1990). *Syntactic and Structural Pattern Recognition. Theory and Applications*, World Scientific.

[4] Chaudhuri B., Rosenfeld A. (1996). On a metric distance between fuzzy sets. *Patter Recognition Letters*, **17**, 1157-1160.

[5] Chaudhuri B., Rosenfeld A. (1999). On a metric distance between fuzzy sets. *Information Sciences*, **118**, 159-171.

[6] Daubechies I. (1992). *Ten lectures on wavelets.* Society for Industrial and Applied Math., Filadelfia.

[7] Dubuison M., Jain A. (1994). Modified Hausdorff distance for object matching. *In International Conference on Pattern Recognition,* 1, 566-568.

[8] Esposito F., et al. (2000). *Analysis of Symbolic Data*, chapter Similarity and Dissimilarity, Springer-Verlag.

[9] Everitt B., Rabe-Hesketh S. (1997). *The Analysis of Proximity Data*, Arnold, London.

[10] Fauqueur J., Boujemaa N. (2006). *Mental image search by Boolean composition of region categories*, Multimedia Tools and Applications, Vol. 31, No. 1, Springer Netherlands, October, 95-117.

[11] Fauzi M., Lewis P. (2006). Automatic texture segmentation for content-based image retrieval application*, Pattern Analysis and Applications*, Springer-Verlag, London.

[12] Freeman H., Glass J. (1961). On the encoding of arbitrary geometric configurations. *IRE Transactions*, **EC-10**(2), 260-268.

[13] Gao Y. Y., Zhang Y. J., Yu F. (2000). *Semantic-based image descrip-tion model and its implementation for image retrieval*. Proceeding of the First International Conference on Image and Graphics, 657-660.

[14] Głomb P.: *Metody indeksowanie ruchomych obrazów wideo*, Ph.D. Dissertation, Institute of Theoretical and Applied Informatics PAS, Gliwice, 2006.

[15] Gowda K., Diday E. (1991). Unsupervised learning through symbolic clustering, *Pattern Recognition Letter*, **12**, 259-564.

[16] Gower J. (1986). Metric and Euclidean Properties of Dissimilarity Coefficients, *Journal of Classification*, **3**, 5-48.

[17] Gudivada V. N., Raghavan V. V. (1995). Design and evaluation of algorithms for image retrieval by spatial similarity, ACM Trans Inf Sys, 13(2), 115-144.

[18] Jain A., Duin R., Mao J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**(1), 4-37.

[19] Jaworska T., Partyka A, (2005). Research: *Content-based image retrieval system* [in Polish], Report RB/37/2005, Systems Research Institute, PAS.

[20] Jaworska T. (2007). Low-Level Analysis as a Basic Process for Content-Based Image Retrieval (CBIR) System, *Proceedings of the 10th IFAC/IFIP /IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems*, Seoul, Korea, September.

[21] Mallat S. (1989). Multiresolution approximations and wavelet orthonormal basis of $\mathbf{L}^2(\mathbf{R})$. In: Trans. Amer. Math. Soc., 315, 69-87.

[22] Mallat S. (1998). *A Wavelet Tour of Signal Processing.* Academic Press, London.

[23] Nadler M., Smith E. (1993). *Pattern Recognition Engineering*. John Willey & Sons Inc., New York.

[24] Petrakis E. (1993). Image representation, *Indexing and Retrieval Based on Spatial Relationships and Properties of Objects*, PhD thesis, University of Crete, Crete, Greece.

[25] Smith J. R., Chang Sh-F. (1999). Integrated spatial and feature image query. *Multimedia Systems*, **7**, 129-140.