Raport Badawczy

Research Report

RB/43/2016

Content-based image retrieval tools and techniques

T. Jaworska

Instytut Badań Systemowych
Polska Akademia Nauk

Systems Research Institute
Polish Academy of Sciences

POLSKA AKADEMIA NAUK

Instytut Badań Systemowych

ul. Newelska 6

01-447 Warszawa

tel.: (+48) (22) 3810100

fax: (+48) (22) 3810105

Kierownik Zakładu zgłaszający pracę:
Prof. dr hab. inż. Janusz Kacprzyk

Warszawa 2016

Tatiana Jaworska

Content-Based Image Retrieval Tools and Techniques

In the beginning was an image.

To my mother
who inspired me
to develop intellectually

Contents

9 Search Engines – Retrieval Techniques

9.1 Introduction

Below, we analyse the most common search engines beginning from the simplest one based on low-level features, through engines including annotations and ending with engines attempting to use semantic matching. In each case we describe the search method, we emphasize goals to which the engine is dedicated and we conclude with a presentation of pros and cons.

Search engines are constructed to fulfil particular criteria which we have described in sect. 2.2. The discussion of these issues will determine which matching mechanism listed below is recommended as more efficient than the others. For instance, if the user wants to find one object in many pictures, e.g. a face in an airport video, they will need a different mechanism than the user who only orders their collection of holidays photos, etc.

Hence, the currently predominant engine categories listed below are based on [213]:

- low-level features and local similarities;
- search by metadata;
- global similarities;
- using object ontology to define high-level concepts,
- bag-of-visual-words (BoVW), stemming from text analysis,
- object retrieval using SIFT and its modification methods,
- relevance feedback (RF) into a retrieval loop for continuous learning about users' intentions,
- generating a semantic template (ST) to support high-level image retrieval,
- making use of both the visual content of images and the textual information obtained from the Web for WWW (the Web) image retrieval,
- combining visual properties of selected objects (or a set of relevant visual features), spatial or temporal relationships of graphical objects [155], [214], with semantic properties [215], [213],
- convolutional neural network (CNN) and deep learning.

All the below-described search engines had to be tested, as a result of which many classified, indexed or annotated reference image databases were developed (see sect. 7.2).

9.2 Visualization and Browsing of Image Databases

Image browsing systems [216] attempt to provide the user with an intuitive interface, displaying at once many images as thumbnails in order to harness the cognitive power of the human mind to recognize and comprehend an image in a second. Interaction with a traditional QBE system can often lead to confusion and frustration on the part of the users, which was confirmed in the study by Rodden and Wood [217].

Fig. 9.1 DB browsing based on visual similarity [218].

Browsing systems give a useful alternative to QBE, providing an overview of the database to the user, which allows for intuitive navigation throughout the system. This is particularly the case when images are arranged according to mutual similarity, as has been shown in [219], where a random arrangement of images was compared with a visualisation which positioned images according to their visual similarities, i.e. where images that are visually similar to each other are located close to one another in the visualisation space [220]. The user can then focus on regions of the visualisations that they are attracted to or believe will harbour a particular concept they have in mind. Browsing such visualisations can increase the rate of retrieval.

For image database browsing, the mapping-based visualization is a typical mechanism which shows the potential relationships within the DB. In order to visualize these high-dimensional features, we have to map them down to 2D on a computer screen.

A variety of methods have been devised in order to visualise images:

- Principal Component Analysis (PCA) which is the simplest dimensionality reduction approach, working in a linear manner (cf. sect. 5.5).
- Multi-Dimensional Scaling (MDS) in turn preserves the original distances in a high dimensional space, calculating a similarity matrix which describes all pair-wise distances between objects in the original space and next it projects them to the low-dimensional space. Based on the similarity matrix, the 'stress' measure can be formulated as follows [220]:

$$\mathrm{st} = \frac{\sum_{i,j}\left(\hat{\delta}_{ij} - \delta_{ij}\right)^2}{\sum_{i,j}\delta_{ij}^2} \tag{9.1}$$

  where δ_{ij} is the original distance between objects i and j, and $\hat{\delta}_{ij}$ is the distance in the low-dimensional space. Rubner et al. [12] who employed MDS based on colour signatures of images and the earth mover's distance (EMD) was able to create a representation of the high-dimensional feature space using MDS, placing image thumbnails at the co-ordinates derived by the algorithm, see Fig. **9.1**.
- Fast Map is an alternative dimensionality reduction technique devised by Faloutsos and Lin [221]. Fast Map reduces high-dimensional spaces down to a linear 2D or 3D space. The algorithm, having a linear complexity $O(kn)$, selects two pivot objects, an arbitrary image and its furthest possible neighbour. All points are mapped to the line that connects the two pivots.
  Later Santos et al. [222] applied the Fast Map to their CBIR system and introduced/introducing the user's modifications of control point positions. This aimed to reduce the semantic gap by pointing out the similarity and diversity among images in human understanding. Their system stored the user's search space modification in the standard CBIR structure.
- Clustering-based visualization. Content-based clustering uses extracted feature vectors in order to group perceptually similar images together (see Fig. 9.2). The advantage of this approach is that no metadata or prior annotation is required in order to arrange images in this manner, although image features or

similarity measures which do not model human perception well, can create groupings that may potentially make it difficult for a user to intuitively browse an image database.

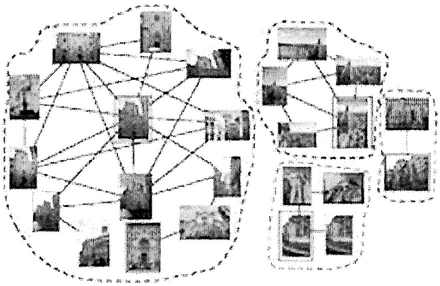


Fig. 9.2 A content-based image clustering method for public image repositories [223].

- Graph-based visualization utilizes links between images to construct a graph where the nodes of the graph are the images, and the edges form the links between similar images. Links can be established through a variety of means including visual similarity between images, or shared keyword annotations, for instance the Pathfinder network [224], see Fig. 9.3. The graph-based visualization appears to be less common because it is typically quadratic in complexity, and therefore can only be computed off-line in order to allow for real-time browsing.

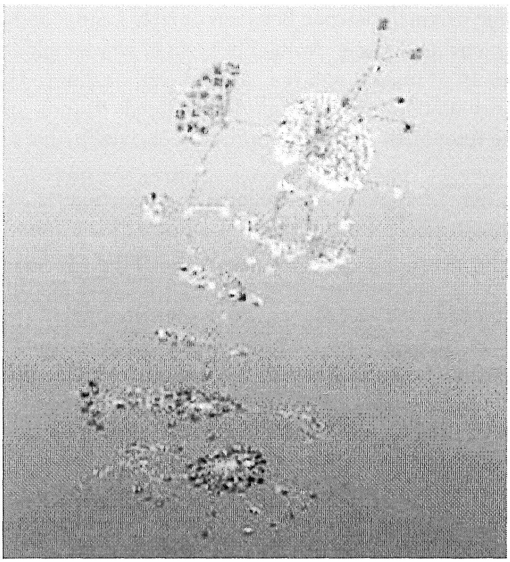


Fig. 9.3 Pathfinder networks of images organized by colour histogram [224].

- Self-Organizing Maps (SOM) [216], [225] is a specific kind of artificial neural network (ANN) which is trained to perform feature extraction and visualization without any supervised signal, simply from the input of raw data. Using an input layer of neurons (see Fig. 9.4), the feature vectors are computed and assigned to best matching units (BMUs). Each unit has the same dimension and is associated with the feature vector computed from the samples in the dataset. A learning rule is typically defined as a process in which the new value $w_i(s+1)$ is computed iteratively from the old one $w_i(s)$ and the new data item $x(s)$, which looks as follows:

$$w_i(s+1) = w_i(s) + \alpha(s)\, h(i^*, i, s)\, [x(s) - w_i(s)] \tag{9.2}$$

where: s is the current iteration, $x(s)$ is a set of input vectors, i^* is the index of the winning neuron, $w_i(s)$ is a weight vector of node u, $h(i^*,i,s)$ is a neighbourhood function modifying the weights around the BMU in the 2D map, i.e. around the winner neuron in step s, $\alpha(s) \in (0,1)$ the monotonically decreasing learning rate.

A spectacular example can be observed in astronomy [226] where solar radio spectrograms coming from the records of three solar radio spectrometers are investigated in Zurich. Even though the information they contain - the radio spectrum between 0.1 and 4 GHz collected over two decades - is not spatial, they are visualized as images, showing the intensity of the radio emission as a function of time and frequency.

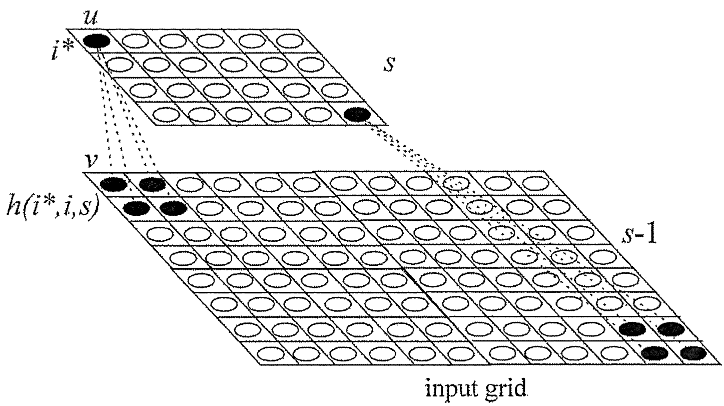


Fig. 9.4 Schematic representation of the SOM ANN architecture.

The SOM was applied to generate local indexing features from images. The input space of the SOM was the region space. A sample of randomly selected regions was taken from all the available images used by a map to learn the point distribution in the region space. Each region was associated with a reaction of a single cell of the map during training. The reactions were summed up into a 'total map' that showed all the reactions associated with a specific image. Then the indexing features were defined as the cells of the map, and their values corresponded to the number of times a cell reacted.

The production of indexing features with SOMs was attractive because:

- The SOM classified the image regions depending on their shape and colour.
- The SOM used actual data distribution to determine classification.
- The learning of the distribution in the region space by the SOM can rely on a large number of regions.

The SOM package used in the Csillaghy's work was developed by Kohonen's group [225]. The tuning parameters of the SOM have been determined experimentally [226].

9.3 Information Retrieval Based on Low-level Features

Image I can be modelled as a function O of the raw image file D, its features F, and representations R. The image model is described below and is also shown in Fig. 9.5:

$$I = O(D,F,R) \tag{9.3}$$

where D is the raw image data, for instance, an image file, $F = \{f_j\}, j = 1,\ldots, J$ is a set of low-level image features, such as colour, shape, texture, etc, $R_j = \{r_{jk}\}$, $k = 1,\ldots, K_j$ is a set of representations for a given feature f_j, e.g. the colour histogram and colour moments are representations of the colour feature. Each representation r_{jk} is a vector consisting of multiple components, i.e.:

$$r_{jk} = [r_{jk1},\ldots,r_{jkl},\ldots,r_{jkL_{jk}}] \tag{9.4}$$

where L_{jk} is the length of the vector r_{jk}.

This image model has three abstract information levels (data, feature, representation), increasing informative granularity. Furthermore, different weights (U at the data level, V_j at the feature level and W_{jk} at the representation level) exist to reflect a particular entity's importance of its level.

In order to compare the distance between two images, we need to define the retrieval model. The image model $O(D,F,R)$, together with a set of distance measures, specifies the retrieval model. Hence, we measure the distance at three levels: image – query $\Phi()$, features $\Theta()$ and representations $\Psi()$. Let r_{mjk} be the jk^{th} representation vector for the m^{th} image in the database, where $m = 1,\ldots, M$ and M is the total number of images in the DB. Let q_{jk}, $j = 1,\ldots, J$, $k = 1,\ldots, K_j$ be the query vector for the jk^{th} representation. The retrieval process is illustrated in Fig. 9.5 and can be described as follows.

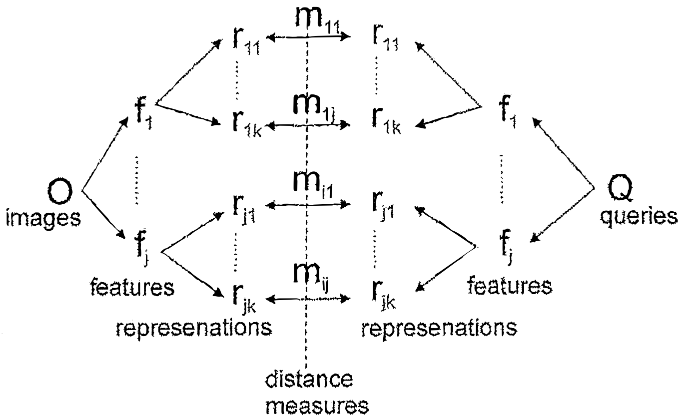


Fig. 9.5 Retrieval process based on feature object representation [227].

First, we initialize the values of the weights U, V_j and W_{jk}. The distance between image and a query in terms of the jk^{th} representation is:

$$d_m(r_{jk}) = \Psi_{jk}(r_{mjk}, q_{jk}, W_{jk}), \tag{9.5}$$

$$m = 1,\ldots, M, j = 1,\ldots, J, k = 1, \ldots, K_j$$

where $d_m(r_{jk})$ denotes the distance between the m^{th} image and the query in terms of representation jk. Then, the distance between the image and the query in terms of feature j is:

$$d_m(f_j) = \Theta_j(d_m(r_{jk}), V_j) = \Theta_j(\Psi_{jk}(r_{mjk}, q_{jk}, W_{jk}), V_j) \tag{9.6}$$

Then, the overall distance is:

$$d_m = \Phi(d_m(f_j), U) = \Phi(\Theta_j(\Psi_{jk}(r_{mjk}, q_{jk}, W_{jk}), V_j), U) \tag{9.7}$$

The images in the DB are ordered by their overall distances to the query (d_m). The N most similar ones are returned to the user, where N is the number of images the user wants to retrieve.

According to the user's preferences, the system dynamically updates the weights U, V_j and W_{jk}. For the Euclidean distance among the feature vector Y. Rui and Th. Hhuang [227] suggested that the computed weight should be $w_{jk} = \frac{1}{\sigma_{jk}}$ which is one over standard deviation.

9.3.1 Scale-Invariant Feature Transform SIFT

The detection and description of local image features can help in object recognition. The SIFT features are local and based on the appearance of the object at particular interest points, and are invariant to image scale and rotation. They are also resistant to changes in illumination, noise, and minor changes in viewpoint. In addition to these properties, they are highly distinctive, relatively easy to extract and allow for correct object identification with a low probability of mismatch. They are relatively easy to match against a (large) database of local features but high dimensionality can be an issue, and generally probabilistic algorithms such as *k-d* trees with best bin first search are used. Object description by a set of SIFT features is also robust to partial occlusion; as few as three SIFT features from an object are enough to compute its location and pose.

An object is recognized in a new image by individually comparing each feature from the query to an image from a database and finding candidate matching features based on the Euclidean distance of their feature vectors. From the full set of matches, subsets of key points that agree on the object and its location, scale, and orientation in a query are identified to filter out good matches. Consistent clusters are determined by using an efficient hash table implementation of the generalized Hough transform. Each cluster of 3 or more features that agree on an object and its pose is then subject to further detailed model verification and subsequently outliers are discarded. Finally, the probability that a particular set of features indicates the presence of an object is computed through the Bayesian probability analysis, given the accuracy of the fit and number of probable false matches. Object matches that pass all these tests can be identified as correct with high confidence.

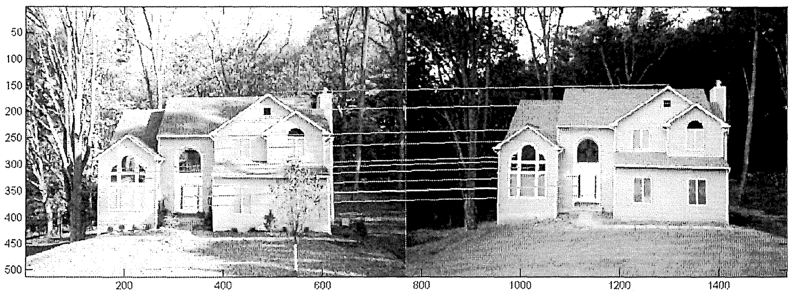


Fig. 9.6 Point-to-point correspondence found by the SIFT descriptors.

This property suggested that this method retrieves all images containing a specific object, even in a large scale image dataset, when that object is given as a query by example (QBE).

Hence, SIFT needs the query-by-example, but in some situations it may be difficult to provide, for instance, when we have an image in our mind but it is difficult to find it as QBE and additionally, we do not need the whole collection of similar images.

SIFT's additional advantage is the fact that it solved the problem of searching for disparity, independently of the issue of epipolar lines in stereovision. The example of point-to-point correspondence is presented in Fig. 9.6.

9.4 Object Ontology to Define High-level Concepts

Generally speaking, ontologies define the concepts and relationships used to describe and represent an area of knowledge. Ontology gives the ability to model the semantics contained in images, such as objects or events. It provides, in a formal way, mutual understanding in a specific domain between humans and computers. Hence, ontology represents knowledge in a hierarchical structure which is used to describe and organize an image collection and it also shows the relation between these images.

In the early approaches high-level concepts were described using the intermediate-level descriptors of the object's ontology. These descriptors were automatically mapped from the low-level features calculated for each region in the database, thus allowing the association of high-level concepts and potentially relevant image regions [228]. Later, ontology was employed to spatial relationships in images such as connectivity, disjoint, meet, adjacency, overlap, cover, or inside. But the image was divided into 3x3, 5x5 or 9x9 windows instead of separate objects [229].

For ontological DBs the Web Ontology Languages (OWL), as a family of knowledge representation languages, have been constructed for authoring ontologies characterized by formal semantics.

An example of a search engine for multimedia has been proposed by Doulaverakis [230] and the system architecture is illustrated in Fig. 9.7. Here the user initiates a query by providing a QBE. This is depicted as case A in Fig. 9.7 and comprises three steps. In the first step (1A) the content-based search is completed by analysing the provided multimedia content (i.e. performing the segmentation, extracting the low-level MPEG-7 descriptors and evaluating the distance between the prototype and the other figures stored in the multimedia database). The second step (2A) takes into account the metadata (which are mapped to the relevant ontologies) of the highest ranked results. For instance, the system may detect the highest ranked results in terms of visual similarity. Based on this information, an ontology-based query is formulated internally in the search engine, which links the knowledge base and enriches the result set with multimedia content that is close semantically to the initial content-based results (3A).

Eventually, the response returned to the user covers a wider range of items of interest, thus facilitating the browsing through the collection and shifting the burden of composing queries to the system instead of the user. The reverse process is equally interesting (case B in Fig. 9.7). Here, the initial query is a combination of terms defined in the ontology, e.g. 'Artefacts from the 1st century BC'. The knowledge base storing the ontology returns the items that fall into that category, as the first step (1B). The second step (2B) involves the extraction and clustering of the low-level multimedia features of this initial set, which is followed by multimedia retrieval, leading to the final step (3B).

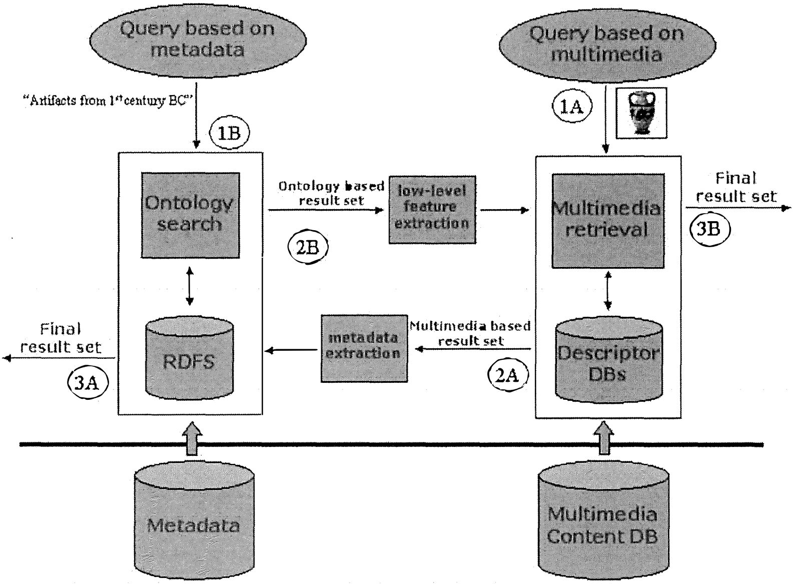


Fig. 9.7 A hybrid ontology and content-based search engine architecture follows [230].

At present, applications use some separate ontologies. For example, Allani et al. [231] defined an image content ontology O_c with a set of image concepts, a metadata ontology O_m addressed surrounding textual information about an image and a visual feature ontology O_F (see Fig. 9.8) with a set of low-level image features. When a query image is introduced, image annotation is processed in order to extract concepts and use them to select relevant features to apply during the retrieval process. Query images are classified given their content into 6 classes. On each class of query images 7 retrieval strategies are performed given feature categories.

Ontology is also a method for organizing extra large-scale image collections, like the ImageNet dataset, created at Stanford University [232].

There are some advantages of ontology:

- its application bridges the semantic gap;
- there is a special language for the user to ask a question;

- ontology-based algorithms are easy to design and are suitable for applications with simple semantic features.

The disadvantage is the necessity of preparing a special DB and annotating the introduction.

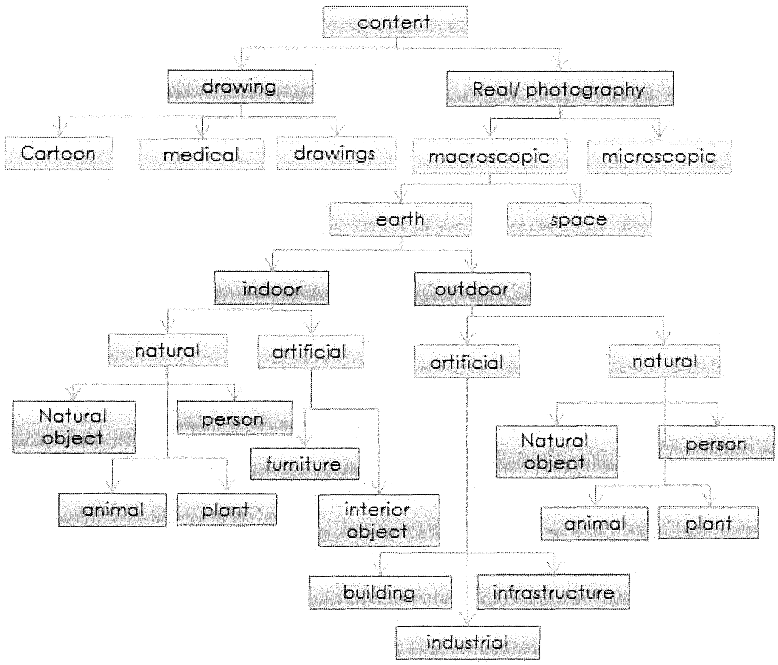


Fig. 9.8 Visual feature ontology **[231]**.

9.5 Bag of Visual Words (BoVW)

A simple approach to classifying images is to treat them as a collection of regions, describing only their appearance and ignoring their spatial structure which is very important in image representation. Similar models have been successfully used in the text community to analyse documents and are known as 'bag-of-words' models, since each document is represented by a distribution over fixed vocabulary. Using such a representation, methods, such as the probabilistic latent semantic analysis (pLSA) [233] and the latent Dirichlet allocation (LDA) [234] extract coherent topics within document collections in an unsupervised manner.

Some time ago, Fei-Fei and Perona [235] and Sivic et al. [236] applied such methods to the visual domain using [233] and [234] in their algorithm.

They modelled an image as a collection of local patches which are detected by a sliding grid and random sampling of scales. Each patch was represented by a code-word from a large vocabulary of code-words sorted in descending order according to the size of their membership and representing simple orientations and illumination patterns. By learning they achieved a model that best represents the distribution of these code-words in each category of scenes. In the recognition process they identified all the code-words in the unknown image. The training and testing process is presented in Fig. 9.9 in a symbolic way.

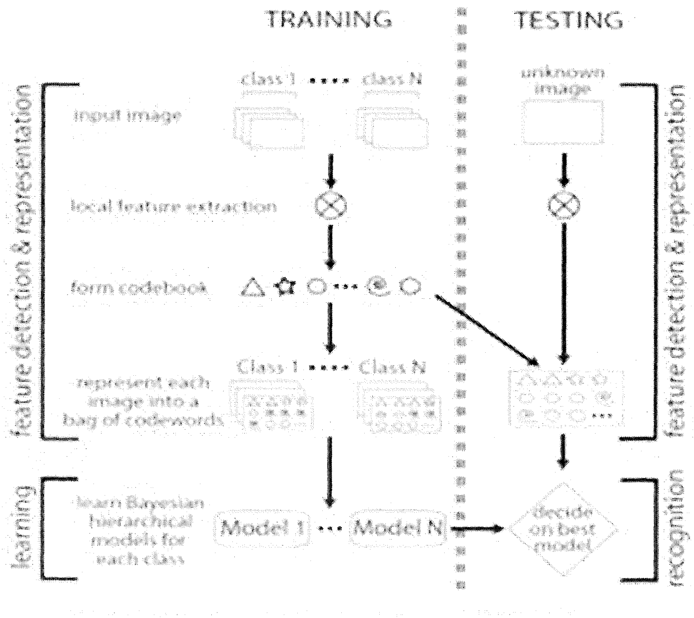


Fig. 9.9 Flow chart of the algorithm follows [235].

They found the category model that matched best the distribution of the code-words of the particular image. Their model was based on a principled probabilistic approach to learn automatically the distribution of code-words and the intermediate-level themes treated as texture descriptions.

It is a method used not only for image retrieval but also for video analysis in order to recognize human actions. Bautista-Ballester et al. [237] applied a BoVW together with a multichannel SVM to the recognition of contextual information. The main goal of this method is to introduce object information relevant to the action into the BoVW-based representation of action. Each video contains one action, and one object per action is detected. The method selected one example image of each object per video and used this image to find the object along the whole video by matching a set of points previously extracted from the frame and the example image.

Concerning the combination of features, six different descriptors are combined for three different pieces of contextual information, namely, 'people' (the histogram of optical flow (HOF) and the histogram of oriented gradient (HOG3D)), 'objects' (HOF and HOG), and 'scene' (GIST and colour histograms). Their combination is accomplished using a multiple MIL approach, which is a concatenation of bag representations and classified with an L_2-Regularized Linear SVM. Information describing the object involved in an action uses a BoVW-based action recognition approach. At first, a set of points belonging to the object is detected by matching these points to an instance of the object. This process also labels the bounding boxes, which are later used to compute a new codebook – the dictionary employed to compute the relative frequencies in a BoVW description–, and the information about the objects in the actions is preserved in consequence. Afterwards, such a codebook is employed to encode the video frames computing a BoVW description. Finally, both sources of information, motion and context are combined by means of a multikernel SVM.

An advantage of the BoVW model is that it is applicable in the case of complex indoor and outdoor images [238]. One of the notorious disadvantages of BoW is that it ignores the spatial relationships among the patches, which are very important in image representation. Additionally, the system needs the preparation of codebooks, classes and Bayesian hierarchical models or an SVM classifier for each class.
[2016_62]

9.6 Relevance Feedback (RF)

Relevance feedback [239], [29], [32] is an interactive technique based on feedback information between a user and a search engine by requiring the user to label semantically similar or dissimilar images with the query image, which are treat as positive and negative samples, respectively. During the last decade, various RF techniques have been proposed to involve the user in the loop to enhance the performance of CBIR [29], [32], [31].

Large modern DBs actively employ user's interaction for relevance feedback (RF). This is an interactive technique based on feedback information between the user and a search engine in which the user labels semantically similar or dissimilar images with a query image, which is treated as positive and negative samples, respectively. Images labelled in this way are incorporated into a training set. The general architecture of such systems is presented in Fig. 9.10.

A more precisely labelled training set boosts algorithms to build a wider boundary between cluster features. For this purpose either Support Vector Machine (SVM) is applied to estimate the density of positive feedbacks or regarding the RF as a strict two-class on-line classification problem or discriminant analysis is used to find a low dimensional subspace of the feature space, so that positive feedbacks and negative feedbacks (which we can see in a relevance feedback in Fig. 9.10) are well separated after projecting onto subspace.

During the last decade, various RF techniques have been proposed to involve the user in the loop to enhance the performance of CBIR [240], [241]. For example, Rui and Huang [227] suggest that for each of the retrieved images, the user provides a degree-of-relevance score, according to the user's feedback, such that the adjusted query q_{jk} and the weights U, V_j and W_{jk} (cf. (9.7)) better match the user's information needs. The user may use a special scroll bars to interactively introduce values of weights which is a more effective mechanism that only binary distinction (as it is illustrated in Fig. 9.10).

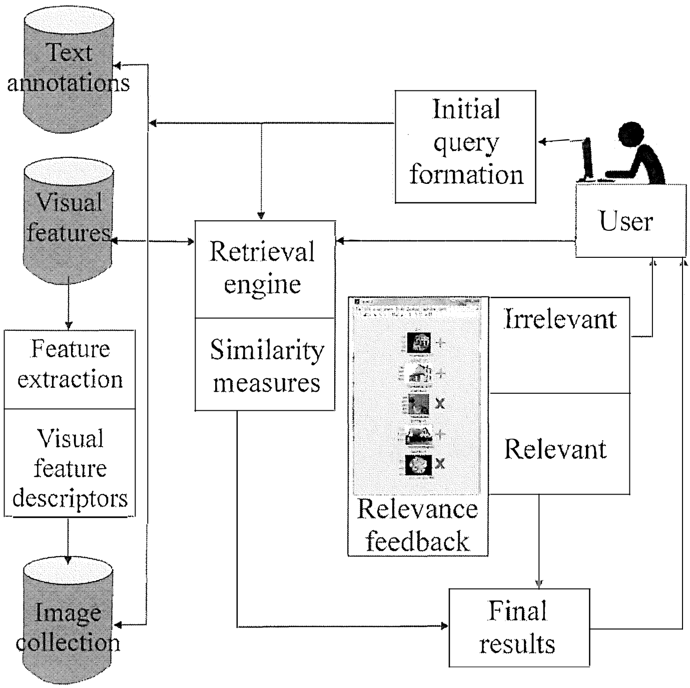


Fig. 9.10 CBIR architecture with the relevance feedback (RF) mechanism.

Whether a retrieval model can update its weights it can better distinguish the *interactive approach* from the *isolated approach* in which all the weights are fixed. Because of the fixed parameters, this approach models the user's information needs and perception subjectivity less effectively. For the interactive approach weights and query vectors are dynamically updated via relevance feedback which improves the efficiency of the system.

Whereas, for instance, L. Zhang et al [29] propose a framework of subspace learning when the training images are associated with only similar and dissimilar pairwise constraints, i.e., Conjunctive Patches Subspace Learning (CPSL) with side information, to explicitly exploit the user's historical feedback log data. It means that they minimize the distances between samples with similar pairwise constraints and to maximize the distances between samples with dissimilar pairwise constraints simultaneously. Samples are whole images for which neighbourhood is calculated as locally linear embedding (LLE) [242].

An option of RF is the adaptive technique based on the ostensive model of developing information needs proposed by J. Urban [239].

Generally, an advantage of RF approach is the fact that the system can start with a limited number of samples because the user will provide next labelled samples. RF has been proved to be effective in boosting image retrieval accuracy. The disadvantage is that most current systems requires about several iterations before it converges to a stable performance level, but users are usually impatient and may give up after two or three tries.

9.7 Semantic Template

In [243] Chang et al. introduced the idea of the semantic visual template (SVT) to link low-level image features to high-level concepts for video retrieval. A visual template is a set of icons or example scenes or objects denoting a personalized view of concepts such as meetings, sunsets, etc. The feature vectors of these example scenes or objects are extracted for the query process. To generate SVTs, the user first defines the template for a specific concept by specifying the objects and their spatial and temporal constraints, the weights assigned to each feature of each object. This initial query scenario is put to the system. Through the interaction with users, the system finally converges to a small set of exemplar queries that 'best' match (maximize the recall) the concept in the user's mind.

Firstly, the user submits a query image with a concept representing the image. After several iterations, the system returns some relevant images to the user. The feature centroids of these images are calculated and used as the representation of the query concept. Then the ST is defined as $ST = \{C, F, W\}$ with C the query concept, F the centroid feature obtained, and W being the weight applied to feature vectors [244]. During the retrieval process, once the user submits a query concept, the system can find a corresponding ST, and use the corresponding F and W to find similar images.

A disadvantage of this system is the necessity of possessing a big lexical database [245].

9.8 WWW Image Retrieval

Image search is based on comparison of metadata associated with the image as keywords, text, etc. and it is obtained a set of images sorted by relevance. The metadata associated with each image can reference the title of the image, format, colour, etc. and can be generated manually or automatically. This metadata generation process is called audio-visual indexing.

WWW search engines exploit the evidence from both the HTML text and visual features of images and develop two independent classifiers based on text and visual image features, respectively. The URL of an image file often has a clear hierarchical structure, including some information about the image, such as image category. In addition, the HTML document also contains some useful information in the image title, ALT-tag, the descriptive text surrounding the image, hyperlinks, etc.

However, the disadvantage is the fact that the retrieval precision is poor and as a result the user has to go through the entire list to find the desired images. This is a time-consuming process which always contains multiple topics which are mixed together. To improve the Web image retrieval performance, researchers are making an effort to fuse the evidence from textual information and visual image contents.

For example, Rasiwasia at al. proposed a combination of a query-by-visual-example (QBVE) with a query-by-semantic-example (QBSE) based on the probability of existance of a visual level represented as a set of feature vectors and the probability of a semantic concept by which an image is annotated. By using the Bayes rule and a similarity function based on methods measuring the distance between two probability distributions (such as the Kullback-Leibler Divergence, Jensen-Shannon Divergence, correlation, etc), they retrieve images most similar to the semantic signature [186].

On the other hand Wang et al. combine the visual features of images with the signatures received from the visual semantic space. For each relevant keyword, a semantic signature of the image is extracted by computing the visual similarities between the image and the reference classes of the keyword using the earlier trained classifiers. The reference classes form the basis of the semantic space of the keyword. If an image has N relevant keywords, then it has N semantic signatures to be computed and stored offline [187].

An advantage of the Web image retrieval is that some additional information on the Web is available to facilitate semantic-based image retrieval.

9.9 Hybrid Semantic Strategy

In this section, we address the information flow and the search engine in the hybrid sematic system (HSS). Fig. 9.11 presents a complex approach to our CBIR system whose particular elements have been described in detail in previous chapters. Here, we can analyse the information flow from introducing a new image up to the results displayed to the user. As it has been mentioned in the concept of the HSS (sect. 2.4), the system consists of several blocks. The separation of particular functions among applications, as shown in Fig. 9.11, is not self-evident, which is why the information flow in our CBIR system is explained below. In a graphical way this flow is illustrated in Fig. 9.12.

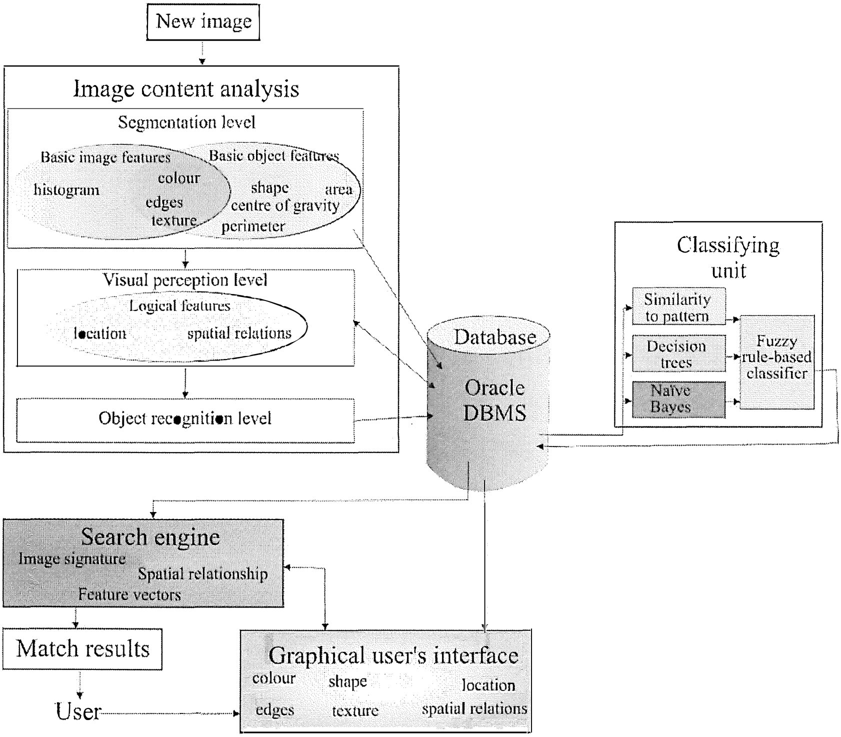


Fig. 9.11 The full structure of our hybrid semantic CBIR system.

All the image content analysis is carried out by Matlab, but it is not a sequential process. Firstly, a new image is segmented (compare sect. 4.2.4 and sect. 3.3.1) and output parameters of this segmentation are sent to Oracle and stored in the database (compare sect. 7.4). This procedure is implemented with the support of the following Matlab Toolboxes: Image Processing, Statistics and Wavelet. Data Base Toolbox supports the communication between Matlab and Oracle. The low-level feature vectors are stored in the OBIEKT table in the DB.

Secondly, the stored parameters pertaining to the DB are transferred to the classifying unit (compare sect. 5.3) for object classification and later identification.

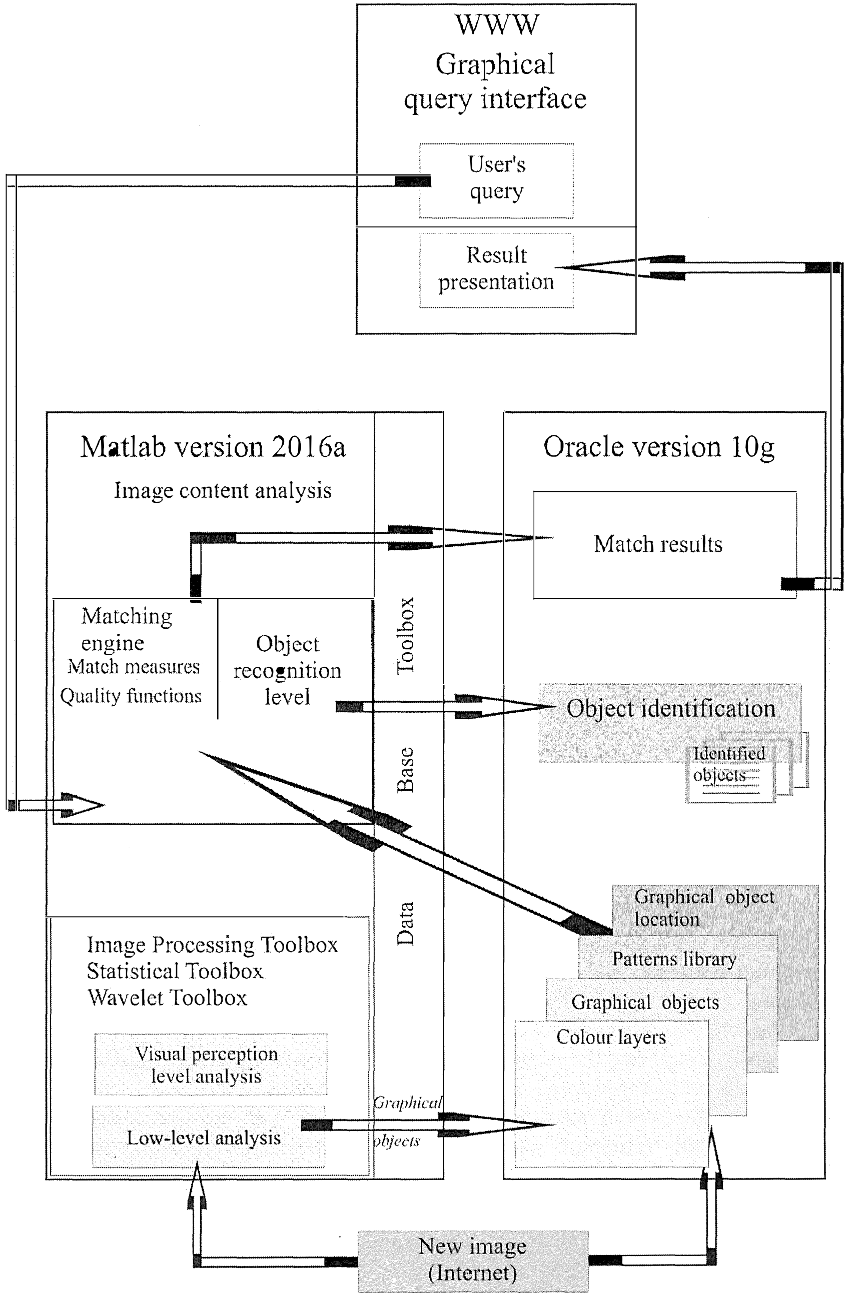


Fig. 9.12 Information flow in our hybrid semantic CBIR system.

In order to find the similarity of all the above-mentioned parameters describing the object, its feature vector has to be compared with those stored in the pattern library (cf. (5.17)). In the classification unit the list of classes is prepared in a semantic way, which means that everything is designed for the user to operate the query-answer process in the most natural and evident way. People tend to attribute a designation to the objects they can see. If we can see a triangular object, we shall more frequently classify it as a roof than other parts of a house, whereas a dark rectangular object will be usually recognized as a window. The part of the identification process which attributes the objects is based on artificial intelligence algorithms and soft computing. This process is implemented in the classification module of the system (see sect. 5.3).

The next stage involves the system for asking-answering user's queries. For this purpose the user's interface (compare sect. 8.2) cooperates with the matching engine.

The matching engine for the HSS carries out three different kinds of comparisons. First of all, the comparison concerns the asymmetric signature as it has been described in detail in sect. 6.4 (cf. (6.9) and (6.10)).

If the maximum component of (6.10) is bigger than a given threshold (a parameter of the search engine), then image I_b is rejected, i.e., not considered further in the process of answering query I_q. Otherwise, we proceed to the next step and we find the spatial similarity $\mathrm{sim}_{\mathrm{PCV}}$ (9.8) of images I_q and I_b, based on the Euclidean, City block or Mahalanobis distance between their PCVs as:

$$\mathrm{sim}_{\mathrm{PCV}}(I_q, I_b) = 1 - \sqrt{\sum_{i=1}^{3} (PCV_{bi} - PCV_{qi})^2} \tag{9.8}$$

This comparison takes into account the spatial information which is very important and rather rarely considered in other systems. The most evident example is the comparison of two mirror images – they have equal numbers of objects and exactly the same objects. Only the spatial information provides the distinction between them.

If the similarity (9.8) is smaller than the threshold (a parameter of the query), then image I_b is rejected. The order of steps (6.10) and (9.8) can be reversed because they are the global parameters and hence can be selected by the user.

Next, we proceed to the final step, namely, we compare the similarity of the objects representing both images I_q and I_b. For each object o_{qi} present in the representation of the query I_q, we find the most similar object o_{bj} of the same class, i.e. $L_{qi} = L_{bj}$. If there is no object o_{bj} of the class L_{qi}, then $\mathrm{sim}_{\mathrm{ob}}(o_{qi}, o_b) = 0$. Otherwise, similarity $\mathrm{sim}_{\mathrm{ob}}(o_{qi}, o_b)$ between objects of the same class is computed as follows:

$$\mathrm{sim}_{\mathrm{ob}}(o_{qi}, o_{bj}) = 1 - \sqrt{\sum_{l} (Fo_{qil} - Fo_{bjl})^2} \tag{9.9}$$

where l is the index of feature vectors F_O used to represent an object. In order to find this similarity, we have to eliminate recursively the pairs of the most similar objects from the process of further comparison. This elimination protects us against matching two or more objects from one image with only one object from the other. The idea is shown in Fig. 9.13, where without this elimination objects O_{q1} and O_{q2} would be matched to the object O_{b2}. This process, described by Mucha and Sankowski [246], is realized according to the Hungarian algorithm for the assignment problem implemented by Munkres.

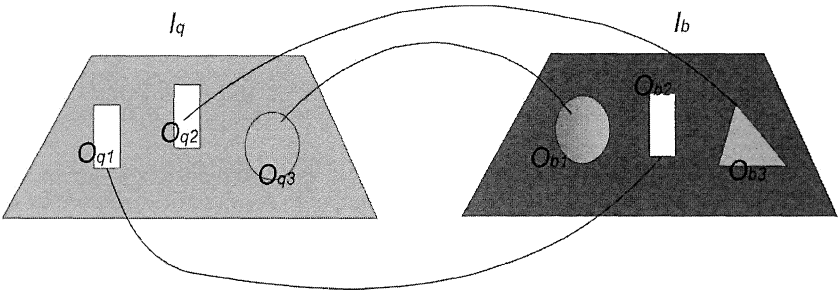


Fig. 9.13 The method for object comparison, where I_q – query and I_b – an image from the DB.

Thus, we obtain the vector of similarities between query I_q and image I_b:

$$\mathrm{sim}(I_q, I_b) = \begin{bmatrix} \mathrm{sim}_{\mathrm{ob}}(o_{q1}, o_{b1}) \\ \vdots \\ \mathrm{sim}_{\mathrm{ob}}(o_{qn}, o_{bn}) \end{bmatrix} \tag{9.10}$$

where n is the number of objects present in the representation of I_q. In order to compare images I_b with the query I_q, we compute the sum of $\mathrm{sim}_{ob}\,(o_{qi}, o_b)$ and then use the natural order of the numbers. Therefore, the image I_b is listed as the first in the answer to the query I_q, for which the sum of similarities is the highest.

Fig. 9.14 presents the main elements of the search engine interface with reference images which are present in the CBIR system. The main (middle) window displays the query signature and PCV, and below it the user is able to set threshold values for the signature, PCV and object similarity. At this stage of system verification it is useful to have these thresholds and metrics at hand. In the final internet version these parameters will be invisible to the user, or limited to the best ranges. The lower half of the window is dedicated to matching results. In the top left of the figure we can see a user designed query comprising elements whose numbers are listed in the signature line. Below the query there is a box with a query miniature, a graph showing the centroids of query components and, further below, there is a 3D plot with PCV components. In the bottom centre windows there are two elements of the same class (e.g. a roof) and we calculate their similarity. On the right side there is a box which is an example of PCA for an image from the DB. The user introduces thresholds to calculate each kind of similarity. For the optimal assigned thresholds a maximum of 11 best matched images from our DB are presented by the search engine.

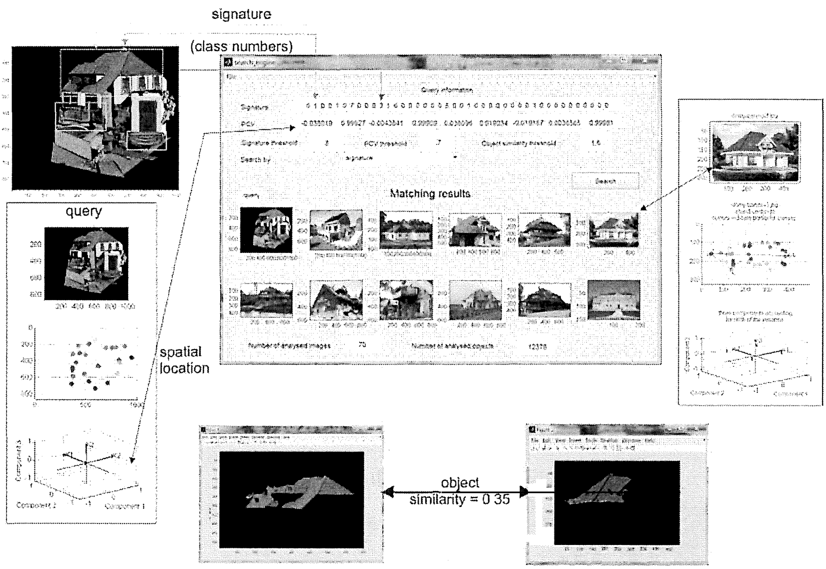


Fig. 9.14 A main concept of the hybrid search engine.

The strong point of our system, as the results will show below, is its semantic context which limits the semantic gap by taking into account middle-level features, such as objects, their numbers and spatial locations in an image. Additionally, we offer the user the GUI to compose their query by which we eliminate the necessity of looking for a QBE.

9.9.1 Retrieval Results

In this section, we conduct experiments on the colour images generated by the user-designed query (UDQ), full images taken from our DB and we will compare our results with another academic CBIR system and the Google image search engine. All images are in the JPG format but in different sizes. Only in order to roughly compare our system's answer to the query, we used SSIM (Universal image similarity index) proposed by Wang and Bovik [247], being aware that it is not fully adequate to present our search engine ranking. SSIM is based on the computation of three constituents, namely the luminance, contrast and structural component, which are relatively independent. In the case of a big difference of images the components can be negative which may result in a negative index.

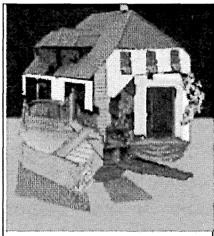
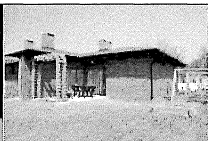
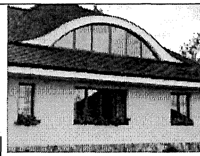
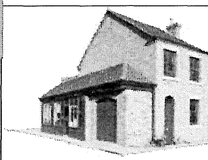
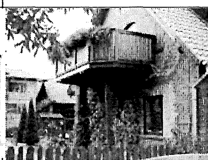
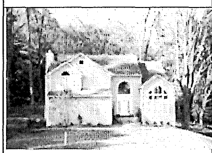
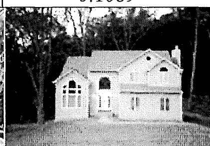
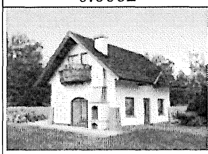
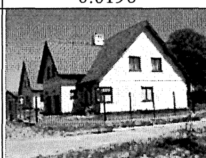
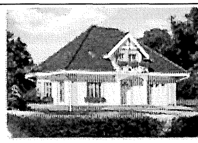
Although there are different sizes of matched images, all of them are resized to the query resolution.

Even though we mentioned two most frequently used measurements evaluated the performance of the system, namely, recall (cf. (2.2)) and precision (cf. (2.1)), below we present the results in the form of images not as graphs, because for the three-stepped search engine there is not one similarity measure. It means that there is not a unique answer if that particular image belongs to the positive condition or negative condition set in a confusion matrix.

9.9.1.1 Results for User Designed Query

A query is generated by the UDQ interface and its size depends on the user's decision, as well as the number of elements (patches). The search engine displays a maximum of 11 best matched images from the DB. Although the user designed few details, the search results are quite acceptable (see Table 9.1 and Table 9.2).

Table 9.1 The retrieval results obtained for two different PCV similarities calculated based on: (1) the Euclidean distance, (2) the City block distance (for thresholds: signature = 17, PCV = 3.5, object = 0.9) attributed to the universal image similarity index.

 Query (1)	 0.1571	 0.1099	 0.0237
	 0.1525	 0.1089	 0.0541
	 0.0062	 0.0196	
	 0.1149	 0,1496	 0.1154

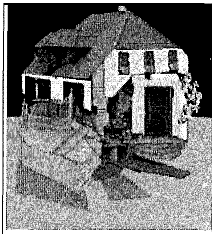
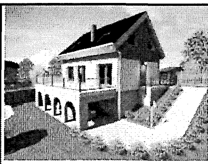
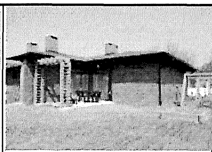
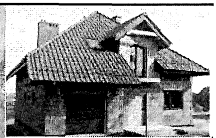
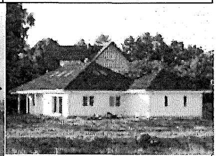
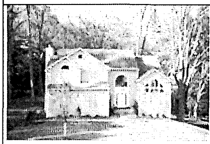
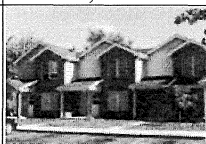
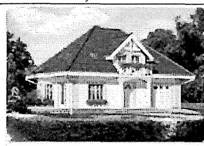
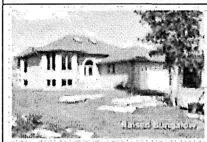
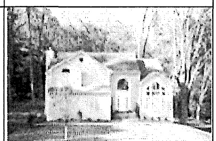
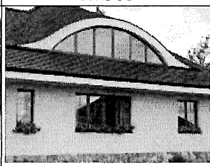
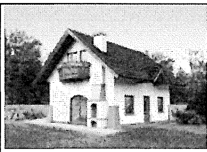
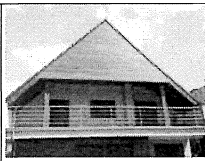
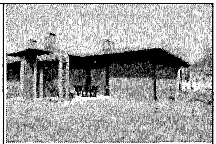
 Query (2)	 0,1492	 0,1571	 0,1099
	 0,1525	 0,1346	 0,0542
	 0,0062	 0,0419	
	 0,1497	 0,1154	

Table 9.2 The retrieval results obtained for PCV similarity calculated based on the City block distance (for thresholds: signature = 20, PCV = 4, object = 0.9) attributed to the universal similarity image index.

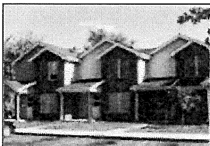
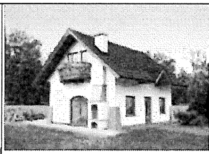
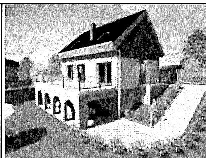
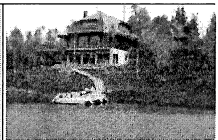
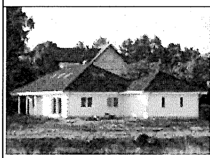
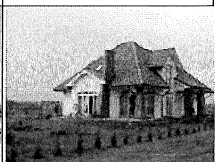
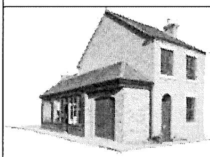
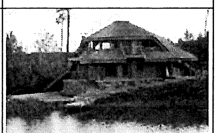
 query	 0.1745	 0.1399	 0.0571
	 0.1443	 0.1505	 0.0012
	 -0.0378	 0.0642	

	 0.2009	 0.0833	 0.2221

9.9.1.2 Results for Full Image

Applying the UDQ is not obligatory. The user can choose their QBE from among the images of the DB if they find an image suitable for their aim. Then the matching results are presented in Table 9.3.

Table 9.3 The retrieval results obtained for PCV similarity calculated based on the Euclidean distance attributed to the universal similarity image index when.

 query	 0.2519	 0.2175	 -0.0255
	 0.1276	 0.3129	 0.2908
	 0.1002	 0.2888	
	 0.2366	 0.0151	 0.0738

9.9.1.3 Comparison to Another Academic CBIR System

We decided to compare our results with the Curvelet Lab system which is based on the Fast Discrete Curvelet Transform (FDCT), developed at Caltech and Stanford University [248] as a specific transform based on the FFT. The FDCT is, among others, dedicated to post-processing applications, such as extracting patterns from large digital images, detecting features embedded in very noisy images. The Curvelet Lab system additionally offers image retrieval, based on such transforms as: DCT (Discrete Cosine Transform), LBP (Local Binary Pattern), colour and combine. Fig. 9. 15 and Fig. 9.16 present the results obtained for a joint set of images, meaning ours and Curvelet Lab system's.

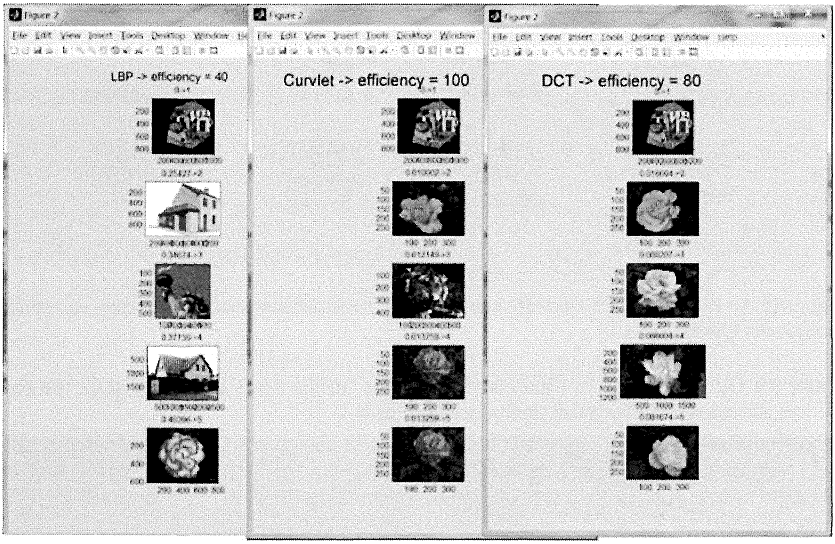


Fig. 9. 15 An example of the Curvelet Lab system retrieval for our query. (Efficiency according to Curvelet Lab system).

9.9.1.4 Comparison with the Google Image Search Engine

We also decided to compare our results with the Google image search engine. We have opted for this comparison because these systems match images without annotations, which has been the most important condition. Systems using annotations belong to quite a different category while our focus is on pure image matching. Results are presented in Table 9.4:

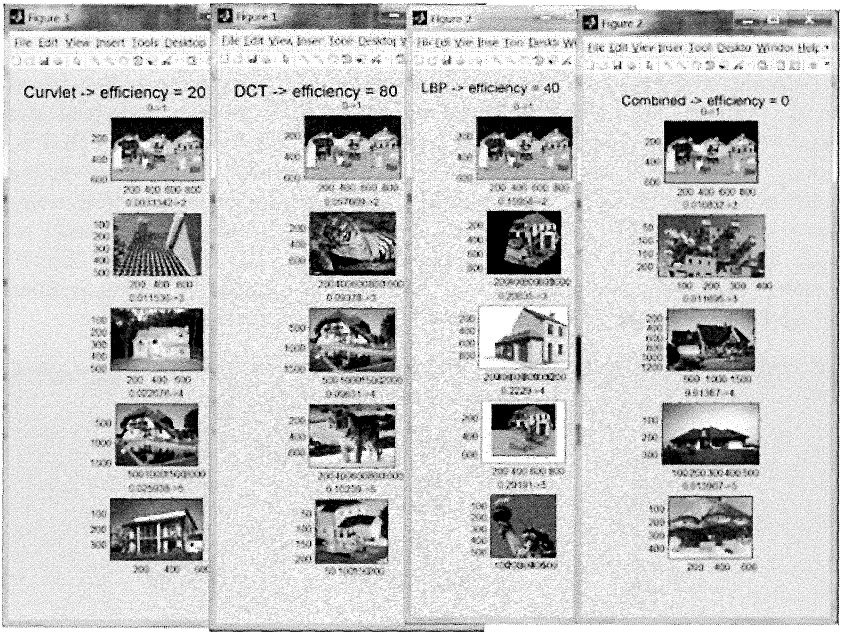
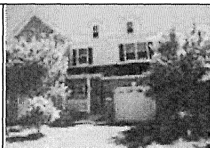
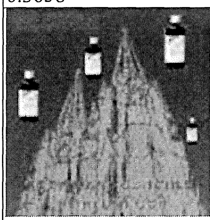
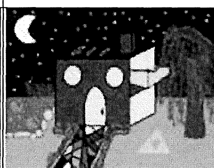
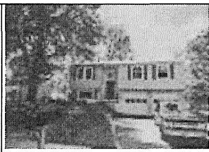
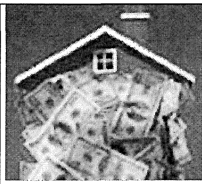
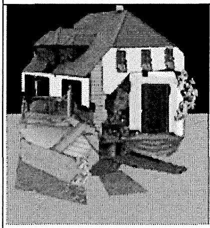
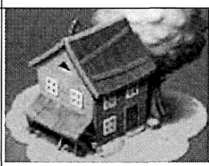
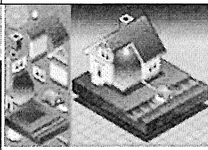
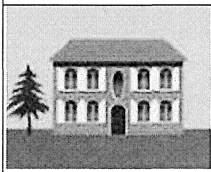
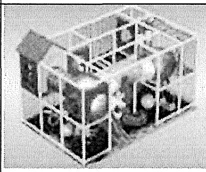


Fig. 9.16 An example of the Curvelet Lab system retrieval for our query. (Efficiency according to Curvelet Lab system).

Table 9.4 The retrieval results obtained with using the Google image search engine for two our queries attributed to the universal similarity image index.

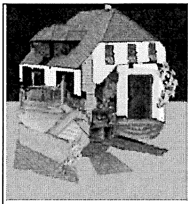
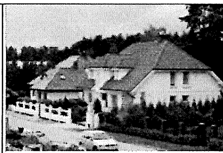
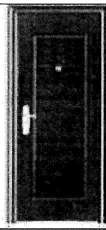
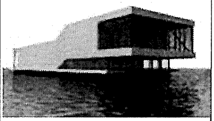
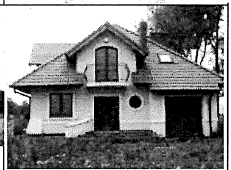
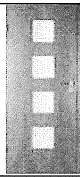
 query	 0.3658	 0.0939	 0.0232
	 0.0240	 0.3174	 0.2056

	 0.1095	 0.1807	
 query	 0.0821	 0.1054	 0.1765
	 0.2666	 0.1076	 0.0876
	 0.2089	 0.1267	

9.9.1.5 Results for SIFT Method

We have opted for the comparison images retrieved by our search engine (presented in Table 9.1) with images retrieved by the SIFT method (presented in Table 9.5) because both systems match images without annotations, which has been the most important condition. Systems using annotations belong to quite a different category while our focus is on pure image matching.

Table 9.5 The retrieval results received based on the SIFT method.

 query			
			
			

9.10 Deep Learning (DL)

Deep learning (DL) is a set of algorithms that attempt to model high level abstractions in data, for instance, images. Some data representations are better than others at simplifying the learning task (e.g., face recognition). One of the promises of deep learning is replacing handicrafted features with efficient algorithms for unsupervised or semi-supervised feature learning and hierarchical feature extraction. Research in this area attempts to create models from large-scale unlabelled data. These works are inspired by advances in neuroscience, especially in the functioning of the brain. Various DL architectures, such as deep neural networks, convolutional deep neural networks, deep belief networks and recurrent neural networks have been applied in image processing. DL in the context of artificial neural networks was introduced by Igor Aizenberg and colleagues in 2000 [249]. In sect. 5.4 we have mentioned how the convolution neural network is built.

DL algorithms are based on distributed representations and exploit the idea of hierarchical explanatory factors where higher level, more abstract concepts are learned from the lower level ones. These architectures are often constructed with a greedy layer-by-layer method.

Wan et al. [40] proposed a deep learning framework for CBIR, which consists of two stages: (i) training a deep learning model in an architecture of CNNs from a large collection of training data (ILSVRC-2012); and (ii) applying the trained deep model, based on the basics of CNNs, to learning feature representations of CBIR tasks in a new domain. For feature representation they used three schemes: (i) direct, (ii) refining by similarity learning, and (iii) refining by model retraining.

The great advantage of deep learning is its capability to deal with large scale image retrieval tasks and it is considered one of the most powerful techniques in AI. It seems to have a great potential when we deal with a big classification or retrieval task comprising of even over a million images or video scenes. Certainly, DL will dynamically develop in the nearest future.

The disadvantage of deep CNNs is the fact that they require a huge amount of data samples to train networks efficiently. Although many benchmarks manage to create abundant samples to be used for training, they lack efficiency when trying to train CNNs to their full potential [249]. Additionally, today's models push the limits of hardware capacity, can take weeks to train, and are carefully fine-tuned for that last push to achieve state-of-the-art results. While deep models distinguish themselves by being able to learn high level abstract representations from data alone, they are prone to having many minute detail parameters. Those parameters can be manually set with reasonable effort for decent results, but must be carefully considered to push the model to its limits [250].

Recently, research has been made into the stability of CNN through different techniques [251]. The point is how far a CNN trained on noisy images might incorrectly classify the next images.

References

[1] Y. Yao, Y. Zeng, N. Zhong and X. Huang, "Knowledge Retrieval," in *Proceedings of the 2007 IEEE/WIC/ACM International Conference on Web Intelligence*, Silicon Valley, USA, 2007.

[2] "http archive," 2016. [Online]. Available: http://httparchive.org/trends.php?s=Top1000&minlabel=Jan+20+2011&maxlabel=Oct+15+2014#bytesImg&reqImg.

[3] S. Nandagopalan, B. S. Adiga and N. Deepak, "A Universal Model for Content-Based Image Retrieval," *World Academy of Science, Engineering and Technology,* vol. 46, pp. 644-647, 2008.

[4] M. Yasmin, S. Mohsin, I. Irum and M. Sharif, "Content Based Image Retrieval by Shape, Color and Relevance Feedback," *Life Science Journal,* vol. 10, no. 4s, pp. 593-598, 2013.

[5] M. Rehman, M. Iqbal, M. Sharif and M. Raza, "Content Based Image Retrieval: Survey," *World Applied Sciences Journal,* vol. 19, no. 3, pp. 404-412, 2012.

[6] Y. J. Lee, . L. C. Zitnick and M. F. Cohen, "ShadowDraw: Real-time User Guidance for Freehand Drawing.," *ACM Transactions on Graphics (TOG),,* vol. 30, no. 4, pp. 1-27, July 2011.

[7] T. M. Lehmann, M. O. Güld, C. Thies, B. Fischer, D. Keysers, K. Spitzer, H. Ney, M. Kohnen, H. Schubert and B. B. Wein, "Content-Based Image Retrieval in Medical Applications," *Methods on Imformatic in Medicine,* vol. 43, pp. 354-361, 2004.

[8] S. Antani, J. Cheng, J. Long, R. L. Long and G. R. Thoma, "Medical Validation and CBIR of Spine X-ray Images over the Internet," in *Proceedings of IS&T/SPIE Electronic Imaging. Internet Imaging VII*, San Jose, C, 2006.

[9] R. K. Srihari, "Automatic Indexing and Content-Based Retrieval of Captioned Images," *IEEE Computer,* vol. 28, no. 9, pp. 49-56, September 1995.

[10] V. Khanaa, M. Rajani, K. Ashok and A. Raj, "Efficient Use of Semantic Annotation in Content Based Image Retrieval (CBIR)," *International Journal of Computer Science Issues,* vol. 9, no. 2, pp. 273-279, March 2012.

[11] C. Carson, S. Belongie, H. Greenspan and J. Malik, "Blobworld: Image Segmentation Using Expectation-Maximization and Its Application to Image Querying," *IEEE Transaction on Pattern Analysis and Machine Intellignece,* vol. 24, no. 8, pp. 1026-1038, Aug. 2002.

[12] Y. Rubner, C. Tomasi and L. J. Guibas , "The Earth Mover's Distance as a Metric for Image Retrieval," *International Journal of Computer Vision,* vol. 40, no. 2, pp. 99-121, 2000.

[13] B. Xiao , X. Gao, D. Tao i X. Li, „Recognition of Sketches in Photos," w *Multimedia Analysis, Processing and Communications*, tom 346, W. Lin, D. Tao, J. Kacprzyk , Z. Li, E. Izquierdo i H. Wang , Redaktorzy, Berlin, Springer-Verlag, 2011, pp. 239-262.

[14] T. Kato, "Database architecture for content-based image retrieval," in *Proceedings of SPIE Image Storage and Retrieval System*, San Jose, CA, USA, 1992, April,.

[15] V. N. Gudivada and V. V. Raghavan , "Content-Based Image Retrieval Systems," *IEEE Computer,* vol. 28, no. 9, pp. 18-22, Sep. 1995.

[16] M. Flickner, H. Sawhney, W. Niblack , J. Ashley, Q. Huang, B. Dom, M. Gorkani , J. Hafner, D. Lee, D. Petkovic, D. Steele and P. Yanker , "Query by Image and Video Content: The QBIC System," *IEEE Computer,* vol. 28, no. 9, pp. 23-32, September 1995.

[17] V. E. Ogle and M. Stonebraker, "CHABOT: Retrieval from a Relational Database of Images," *IEEE Computer,* vol. 28, no. 9, pp. 40-48, September 1995.

[18] R. Mehrotra and J. E. Gary, "Similar-Shape Retrieval in Shape Data Management," *IEEE Computer,* vol. 28, no. 9, pp. 57-62, Sep. 1995.

[19] M. Nakazato i T. S. Huang, „3D MARS: Immersive Virtual Reality for Content-Based Image Retrieval," w *IEEE International Conference on Multimedia and Expo*, Tokyo, August 22-25, 2001.

[20] S. Saurin, „Saurin Shah Portfolio," 2014. [Online]. Available: http://www.shahsaurin.com/projects_demo/threejs-webgl/ .

[21] G. Chang, M. J. Healey , J. A. M. McHugh i J. T. L. Wang, Mining the World Wide Web: An Information Search Approach., Norwell: Kluwer Academic, 2001.

[22] T. Jaworska, „Object extraction as a basic process for content-based image retrieval (CBIR) system.," *Opto-Electronics Review,* tom 15, nr 4, pp. 184-195, Dec. 2007.

[23] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *Internationa Journal of Computer Vision,* vol. 60, no. 2, pp. 91-110, 2004.

[24] D. G. Lowe, "Object Recognition from local scale-invariant features," in *International Conferences on Computer Vision*, Corfu, Greece, 1999.

[25] . C. Leininger , „Fusion d'images : des outils au service des neurochirurgiens," June 2006. [Online]. Available: https://interstices.info/jcms/c_16870/fusion-d-images-des-outils-au-service-des-neurochirurgiens.

[26] M. R. Azimi-Sadjadi, J. Salazar and S. Srinivasan, "An Adaptable Image Retrieval System With Relevance Feedback Using Kernel Machines and Selective Sampling," *IEEE Transactions on Image Processing,* vol. 18, no. 7, p. 1645 1659, 2009.

[27] J. Urban, J. M. Jose and C. J. van Rijsbergen, "An adaptive technique for content-based image retrieval," *Multimedial Tools Applied,* no. 31, pp. 1-28, July 2006.

[28] X. S. Zhou and T. S. Huang, "Relevance Feedback in Image Retrieval: A Comprehensive Review," *ACM Multimedia Systems,* vol. 8, no. 6, pp. 536-544, 2003.

[29] L. Zhang, L. Wang and W. Lin, "Conjunctive patches subspace learning with side information for collaborative image retrieval," *IEEE Transactions on Image Processing,* vol. 21, no. 8, pp. 3707-3720, 2012.

[30] M. M. Rahman, S. K. Antani and G. R. Thoma, "A query expansion framework in image retrieval domain based on local and global analysis," *Information Processing and Management,* vol. 47, pp. 676-691, 2011.

[31] L. Zhang, L. Wang and W. Lin, "Generalized biased discriminant analysis for content-based image retrieval," *IEEE Transactions on System, Man, Cybernetics, Part B - Cybernetics,* vol. 42, no. 1, pp. 282-290, 2012.

[32] L. Zhang, L. Wang and W. Lin, "Semi-supervised biased maximum margin analysis for interactive image retrieval," *IEEE Transactions on Image Processing,* vol. 21, no. 4, pp. 2294-2308, 2012.

[33] L. Wang, W. Lin and L. Zhang, "Geometric Optimum Experimental Design for Collaborative Image Retrieval," *IEEE Transactions on Circuits and System for Video Technology,* vol. 24, pp. 346-359, 2014.

[34] F. Long, H. Zhang and D. D. Feng, "Fundamentals of content-based image retrieval," in *Multimedia Information Retrieval and Management Technological Fundamentals and Applications.*, New York, Sprainger-Verlag, 2003, pp. 1-26.

[35] S. Gould and X. He, "Scene Understanding by labellilng Pixels," *Communications of the ACM,* vol. 57, no. 11, pp. 68-77, November 2014.

[36] J. Yao, S. Fidler and R. Urtasun, "Describing the Scene as a Whole: Joint Object Detection, Scene Classification and Semantic Segmentation," in *The 26th IEEE Conference on Computer Vision and Pattern Recognition*, Providence, Rhode Island, 2012.

[37] L.-J. Li, H. Su, . E. P. Xing and L. Fei-Fei, "Object Bank: A High-Level Image Representation for Scene Classification and Semantic Feature Sparsification," in *24th Annual Conference on Neural Information Processing Systems* , Vancouver, Canada, 2010.

[38] D. M. Wells, A. P. French, A. Naeem, O. Ishaq and R. Traini, "Recovering the dynamics of root growth and development using novel image acquisition and analysis methods," *Phisiological Transactions of The Royal Society B,* no. 367, p. 1517–1524, 2012.

[39] C. Steger, M. Ulrich and C. Wiedemann, Machine Vision Algorithms and Applications, Weinheim: Wiley-VCH, 2008.

[40] J. Wan, D. Wang, S. C. Hoi, P. Wu, J. Zhu, Y. Zhang and J. Li, "Deep Learning for Content-Based Image Retrieval: A Comprehensive Study," in *Proceedings of the ACM International Conference on Multimedia*, Orlando, Florida, 3-7 Nov. 2014.

[41] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain, "Content-Based Image Retrieval at the End of the Early Years," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 22, no. 12, pp. 1349-1380, Dec 2000.

[42] T. Jaworska, „A Search-Engine Concept Based on Multi-Feature Vectors and Spatial Relationship," w *Flexible Query Answering Systems*, tom 7022, H. Christiansen, G. De Tré, A. Yazici, S. Zadrożny i H. L. Larsen, Redaktorzy, Ghent, Springer, 2011, pp. 137-148.

[43] C.-R. Su, J.-J. Chen and K.-L. Chang, "Content-Based Image Retrieval on Reconfigurable Peer-to-Peer Networks," in *International Symposium on Biometrics and Security Technologies*, 2013.

[44] "List of CBIR engines," 2015. [Online]. Available: http://en.wikipedia.org/wiki/List_of_CBIR_engines.

[45] L.-J. Li, C. Wang, Y. Lim, D. M. Blei and L. Fei-Fei, "Building and Using a Semantivisual Image Hierarchy," in *IEEE Conference on Computer Vision and Pattern Recognition* , June, 2010.

[46] F. Wu, Advances in Visual Data Compression and Communication: Meeting the Requirements of New Applications, CRC Press, 2014, p. 513.

[47] J. G. Kolo, K. P. Seng, L.-M. Ang and S. R. S. Prabaharan, "Data Compression Algorithms for Visual Information," in *Informatics Engineering and Information Science*, vol. 253, A. A. Manaf, . S. Sahibuddin, . R. Ahmad , . S. M. Daud and . E. El-Qawasmeh , Eds., Berlin, Springer-Verlag, 2011, pp. 484-497.

[48] N. Sharda, "Multimedia Transmission ober Wireless Sensor Networks," in *Visual Information Processing in Wireless Sensor Networks: Technology, Trends and Applications*, L. Ang, Ed., 2011.

[49] T. Jaworska, „Object extraction as a basic process for content-based image retrieval (CBIR) system.," *Opto-Electronics Review,* tom 15, nr 4, pp. 184-195, December 2007.

[50] T. Jaworska, "Database as a Crucial Element for CBIR Systems," in *Proceedings of the 2nd International Symposium on Test Automation and Instrumentation*, Beijing, China, 16-20 Nov., 2008.

[51] T. Jaworska, "Application of Fuzzy Rule-Based Classifier to CBIR in comparison with other classifiers," in *11th International Conference on Fuzzy Systems and Knowledge Discovery*, Xiamen, China, 19-21.08.2014.

[52] T. Jaworska, "Spatial representation of object location for image matching in CBIR," in *New Research in Multimedia and Internet Systems*, vol. 314, A. Zgrzywa, K. Choroś and A. Siemiński, Eds., Wrocław, Springer, 2014, pp. 25-34.

[53] T. Jaworska, "Query techniques for CBIR," in *Flexible Query Answering Systems*, vol. 400, T. Andreasen, H. Christiansen, J. Kacprzyk, H. Larsen, G. Pasi, O. Pivert, G. De Tre, M. A. Vila, A. Yazici and S. Zadrożny, Eds., Cracow, Springer, 2015, pp. 403-416.

[54] Y.-J. Zhang, Y. Gao and Y. Luo, "Object-Based Techniques for Image Retrieval," in *Multimedia Systems and Content-Based Image Retrieval*, S. Deb, Ed., Hershey, London, IDEA Group Publishing, 2004, pp. 156-181.

[55] T. Tuytelaars and K. Mikolajczyk, "Local Invariant Feature Detectors: A Survey," *Computer Graphics and Vision*, vol. 3, no. 3, p. 177–280, 2007.

[56] W. Niblack, M. Flickner, D. Petkovic, P. Yanker, R. Barber, W. Equitz, E. Glasman, C. Faloutsos and G. Taubin, "The QBIC Project: Querying Images by Content Using Colour, Texture and Shape," *SPIE*, vol. 1908, pp. 173-187, 1993.

[57] G. Pass and R. Zabith, "Histogram refinement for content-based image retrieval," *IEEE Workshop on Applications of Computer Vision*, pp. 96-102, 1996.

[58] M. Pietikäinen, Ed., Texture Analysis in Machine Vision, vol. 40, World Scientific, 2000.

[59] N. Sebe and M. S. Lew, "Texture Features for Content-Based Retrieval," in *Principles of Visual Information Retrieval*, M. S. Lew, Ed., London , Springer Science & Business Media, 2013, pp. 50-81.

[60] M. Tuceryan and A. K. Jain, "Texture Analysis," in *The Handbook of Pattern Recognition and Computer Vision*, 2 ed., C. H. Chen, L. F. Pau and P. S. P. Wang, Eds., World Scientific Publishing Co., 1998, pp. 207-248.

[61] S. W. Zucker, "Toward a Model of Texture," *Computer Graphics and Image Processing*, vol. 5, pp. 190-202, 1976.

[62] N. Ahuja, "Dot Pattern Processing Using Voronoi Neighborhoods," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, no. 4, pp. 336-343, May 1982.

[63] R. M. Haralick, "Statistical and Structural Approaches to Texture," *Proceedings of the IEEE*, vol. 67, pp. 786-804, 1979.

[64] . M. Pietikäinen, T. Ojala and D. Harwood, "A Comparative Study of Texture Measures with Classification Based on Feature Distributions.," *Pattern Recognition*, vol. 29, no. 1, pp. 51-59, January 1996.

[65] T. Ojala, M. Pietikäinen and T. Mäenpää, "Multiresolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns.," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, 2002.

[66] M. Pietikäinen, A. Hadid, G. Zhao and T. Ahonen, Computer Vision Using Local Binary Patterns, vol. 40 in Computational Imaging and Vision, Springer Science & Business Media, 2007.

[67] H. Tamura, S. Mori i T. Yamawaki, „Texture features corresponding to visual perception," *IEEE Transactions On Systems, Man and Cybernetics*, tom 8, pp. 460-473, 1978.

[68] R. Sriram , J. M. Francos and W. A. Pearlman, "Texture coding using a Wold decomposition model.," *IEEE Transactions of Image Processing* , vol. 5, no. 9, pp. 1382-1386, 1996.

[69] G. L. Gimel'farb and A. K. Jain, "On retrieving textured images from an image data base.," *Pattern Recognition*, vol. 29, no. 9, pp. 1461-1483, 1996.

[70] A. P. Pentland, "Fractal-based description of natural scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 661-674., June 1984.

[71] B. B. Mandelbrot, Fractal Geometry of Nature, New York: Freeman, 1982.

[72] H. E. Hurst, "Long-term storage capacity of reservoirs," *Transactions of the American Society of Civil Engineers,* vol. 116, no. 1, pp. 770-799, 1951.

[73] S. Ezekiel and J. A. Cross, "Fractal-based Texture Analysis," in *APCC/OECC'99, Joint Conference of 5th Asia-Pacific Conference on Communications (APCC) and 4th Opto-Electronics and Communications Conference (OECC),* 1999.

[74] J. Millard, P. Augat, T. M. Link, M. Kothari, D. C. Newitt, H. K. Genant, and S. Majumdar, "Power Spectral Analysis of Vertebral Trabecular Bone Structure from Radiographs: Orientation Dependence and Correlation with Bone Mineral Density and Mechanical Properties," *Calcified Tissue International,* vol. 63, pp. 482-489, 1998.

[75] S. Selvarajah and S. R. Kodituwakku, "Analysis and Comparison of Texture Features for Content Based Image Retrieval," *International Journal of Latest Trends in Computing,* vol. 2, no. 1, pp. 108-113, March 2011.

[76] G. M. Haley and B. S. Manjunath, "Rotation-Invariant Texture Classification Using a Complete Space-Frequency Model," *IEEE Transactions on Image Processing,* vol. 8, no. 2, Feb. 1999.

[77] D. Gabor, "Theory of communication," *Journal of the Institution of Electrical Engineers ,* pp. 445 - 457, 1946.

[78] T. S. Lee, "Image Representation Using 2D Gabor Wavelets," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE,* vol. 18, no. 10, October 1996.

[79] T. Jaworska, "Point-to-point correspondence into stereo pair of images," Silesian University of Technology, Gliwice, Poland, 2001.

[80] N. Sebe and M. S. Lew, "Wavelet Based Texture Classification," in *Proceedings. 15th International Conference on Pattern Recognition,* 2000.

[81] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE TRANSACTIONS ON COMMUNICATIONS,* Vols. COM-31, no. 4, pp. 532-540, April 1983.

[82] J. L. Crowley, "A representation for visual information," 1987.

[83] I. Daubechies, Ten lectures on wavelets, Philadephia: Society for Industrial and Applied Mathematics, 1992.

[84] S. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 11, no. 7, pp. 674-693, 1989.

[85] S. Mallat, "Multiresolution Approximation and Wavelet Orthonormal Bases of L2(R)," *Transactions American Mathematical Society,* vol. 315, no. 1, pp. 69-87, 1989.

[86] Y. Meyer, Les ondelettes. Algorithmes et applications, Paris: Armand Colin, 1992.

[87] P. Wojtaszczyk, Wavelet Theory (in Polish), Warsaw: PWN, 2000.

[88] S. Mallat, A wavelet tour of signal processing, Academic Press, 1998.

[89] M. Faizal, A. Fauzi and P. H. Lewis, "Automatic texture segmentation for content-based image retrieval application," *Pattern Analysis and Applications,* vol. 9, p. 307–323, 2006.

[90] R. A. Kirsch , "Computer determination of the constituent structure of biological images," *Computers and Biomedical Research,* vol. 4, no. 3, p. 315–328, July 1971.

[91] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion simulations," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 13, no. 6, p. 583–598, 1991.

[92] O. Basir, H. Zhu and F. Karray, "Fuzzy Based Image Segmentation," in *Fuzzy Filters foe Image processing,* vol. 122, Berlin, Springer, 2003, pp. 101-128.

[93] H. M. Sobel, Multivariate Observations, Wiley, 1984.

[94] J. M. S. Prewitt, "Object Enhancement and Extraction," in *Picture Processing and Psychopictorics,*, B. S. B. S. Lipkin and A. Rosenfeld, Eds., NY, Academic Press, 1970.

[95] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vols. PAMI-8, no. 6, pp. 679-698, 1986.

[96] C. Xu and J. L. Prince, "Snakes, Shapes, and Gradient Vector Flow," *IEEE TRANSACTIONS ON IMAGE PROCESSING,* vol. 7, no. 3, pp. 359-369, March 1998.

[97] R. O. Duda and P. E. Hart, "Use of the HOUGH Transformation to Detect Lines and Curves in Pictires," 1971.

[98] Q. Zhu, . L. Wang, Y. Wu and J. Shi,, "Contour Context Selection for Object Detection: A Set-to-Set Contour Matching Approach,," in *The 10th European Conference on Computer Vision (ECCV)*, Marseille, France, 2008.

[99] D. Zhang and G. Lu, "Review of shape representation andd escription techniques," *Pattern Recognition,* vol. 37, p. 1 – 19, 2004.

[100] S. Abbasi, F. Mokhtarian and J. Kittler, "Curvature scale space image in shape similarity retrieval," *Multimedia Systems,* no. 7, p. 467–476, 1999.

[101] C.-J. Sze, H.-R. Tyan, H.-Y. M. Liao, C.-S. Lu and S.-K. Huang, "Shape-based Retrieval on a Fish Database of Taiwan," *Tamkang Journal of Science and Engineering,* vol. 2, no. 3, pp. 63-173 , 1999.

[102] T. B. Sebasian and B. B. Kimia, "Curves vs Skeltons in Object Recognition," in *Proceedings of International Conference on Image Processing,* Thessaloniki, 7-10 Oct. 2001 .

[103] L. Kotoulas i I. Andreadis, „Image analysis using moments," w *Proceedings of 5th International Conference on Technology and Automation*, Thessaloniki, Greece, 2005.

[104] M. R. Teague, "Image analysis via the general theory of moments," *Journal of the Optical Society of America,* vol. 70, no. 8, pp. 920-930, 1980.

[105] R. Arandjelović and A. Zisserman, "Three things everyone should know to improve object retrieval," in *IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, USA, 2012.

[106] K. Mikolajczyk and C. Schmid, "Scale & Affine Invariant Interest Point Detectors," *International Journal of Computer Vision,* pp. 63-86, 2004.

[107] F. Perronnin, J. Sanchez and T. Mensink, "Improving the Fisher Kernel for Large-Scale Image Classification," in *European Conference on Computer Vision, Lecture Notes in Computer Science*, Heraclion, Greece, Sep, 2010.

[108] F. Perronnin and C. Dance, "Fisher Kernels on Visual Vocabularies for Image Categorization," in *Proceeding Computer Vision and Pattern Recognition*, 2007.

[109] J. Krapac and S. Śegvić, "Weakly Supervised Object Localization with Large Fisher Vectors," in *Proceedings of the 10th International Conference on Computer Vision Theory and Applications*, Berlin, 11-14 Mar, 2015.

[110] H. Jegou, M. Douze, C. Schmid and P. Perez, "Aggregating local descriptors into a compact image representation," in *IEEE Conference on Computer Vision and Pattern Recognition*, San Francisco, 13-18 June, 2010.

[111] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," in *IEEE International Conference on Computer Vision*, 2005.

[112] E. Rosten i T. Drummond, „Machine learning for high-speed corner detection," w *European Conference on Computer Vision*, 2006.

[113] E. Rublee, V. Rabaud , K. Konolige and G. Bradski, "ORB: an efficient alternative to SIFT or SURF," in *IEEE International Conference on Computer Vision (ICCV)*, Barcelona, Spane, 6-12, Nov, 2011.

[114] M. Brown, R. Szeliski i S. Winder, „Multi-image matching using Multi-Scale Oriented Patches," *Computer Vision and Pattern Recognition,* nr 2, pp. 510-517, 2005.

[115] The Moving Picture Experts Group, "MPEG," [Online]. Available: http://mpeg.chiariglione.org/. [Accessed 2015].

[116] MPEG, "MPEG standards - Full list of standards developed or under development," 20 April 2010. . [Online]. Available: http://mpeg.chiariglione.org/standards.htm.

[117] I. JTC1/SC29/WG11, "CODING OF MOVING PICTURES AND AUDIO MPEG-7". Palma de Mallorca, Spain Patent N6828, Oct. 2004.

[118] M. J. Swain and D. H. Ballard, "Color Indexing," *International Journal of Computer Vision,* vol. 7, no. 1, pp. 11-32, 1991.

[119] V. Castelli i L. D. Bergman, Redaktorzy, Image Databases: Search and Retrieval of Digital Imagery, New York: Wiley, 2002.

[120] J.-J. Chen, C.-R. Su, W. E. L. Grimson, J.-L. Liu and D.-H. Shiue, "Object Segmentation of Database Images by Dual Multiscale Morphological Reconstructions and Retrieval Applications," *IEEE Transactions on Image Processing,* vol. 21, no. 2, pp. 828-843, Feb. 2012.

[121] P. Melin and O. Castillo, Hybrid Intelligent Systems for Pattern Recognition Using Soft Computing. An Evolutionary Approach for Neural Networks and Fuzzy Systems., Berlin: Springer, 2005, p. 272.

[122] J. C. Bezdek, Pattern Recognition with Fuzzy Objective Function Algorithms., New York: Plenum Press, 1981, p. 272.

[123] Y. Cheng , "Mean Shift Mode Seeking, and Clustering," *IEEE TRANSACTIONS on PATTERN ANALYSIS and Machine Intelligence,* vol. 17, no. 8, Aug, 1995.

[124] G. Seber, Multivariate Observations, New York: Wiley, 1984, p. 686.

[125] H. Späth, Cluster analysis algorithms for data reduction and classification of objects, vol. 4, Pensilvania University: E. Horwood, 1980, p. 226.

[126] M. Acharyya and M. K. Kundu, "An adaptive approach to unsupervised texture segmentation using M-Band wavelet transform," *Signal Processing,* no. 81, pp. 1337-1356, 2001.

[127] L. J. Latecki and R. Lakamper, "Application of planar shape comparison to object retrieval in image databases," *Pattern Recognition,* no. 35, pp. 15-29, 2002.

[128] W.-B. Goh and K.-Y. Chan, "A Shape Descriptor for Shapes with Boundary Noise and Texture," in *British Machine Vision Conference*, Norwich, 24 June, 2003.

[129] C. Xu and J. Liu, "2D Shape Matching by Contour Flexibility," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE,* vol. 31, no. 1, Jan. 2009.

[130] J. Mutch and D. G. Lowe, "Object class recognition and localization using sparse features with limited receptive fields," *International Journal of Computer Vision (IJCV),* vol. 80, no. 1, pp. 45-57, Oct 2008.

[131] T. Serre, L. Wolf and T. Poggio, "Object Recognition with Features Inspired by Visual Cortex," in *Proceedings on Computer Vision and Pattern Recognition*, Los Alamos, 2005.

[132] Y. Li and L. G. Shapiro, "Object Recognition for Content-Based Image Retrieval," Dagstuhl Seminar, Leibniz, Austria, 2002.

[133] G. Quellec, M. Lamard, G. Cazuguel, B. Cochener and C. Roux, "Fast Wavelet-Based Image Characterization for Highly Adaptive Image Retrieval," *IEEE Transactions on Image Processing,* vol. 21, no. 4, pp. 1613-1623, April 2012.

[134] B. V. Dasarathy, Ed., Nearest neighbor (NN) norms : NN pattern classification techniques, 6th ed., Los Alamitos, Callifornia: IEEE Computer Society Press, 1991.

[135] C. Cortes and V. Vapnik , "Support-Vector Networks," *Machine Learning,* vol. 20, p. 273–297, 1995.

[136] I. Rish, "An empirical study of the Naïve Bayes classifier," in *Proceedings of the IJCAI-2001 Workshop on Empirical Methods in AI,* Brussels, 2001.

[137] G. P. Zhang, "Neural Networks for Classification: A Survey," *IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and reviews,* vol. 30, no. 4, pp. 451-462, Nov 2000.

[138] J. M. Ali, "Content-Based Image Classification and Retrieval: A Rule-Based System Using Rough Sets Framework," in *Artificial Intelligence for Maximizing Content Based Image Retrieval,* Z. Ma, Ed., NY, Springer, 2009, pp. 68-82.

[139] T. Jaworska, "Towards Fuzzy Classificaton in CBIR," in *Information Systems Architecture and Technology.* , Vols. Knowledge Based Approach to the Design, Control and Decision Support, J. Świątek, L. Borzemski, A. Grzech and Z. Wilimowska, Eds., Wrocław, Oficyna Wydawnicza Politechniki Wrocławskiej, 2013, pp. 53-62.

[140] U. M. Fayyad and K. B. Irani, "The attribute selection problem in decision tree generation," in *the 10th National Conference on Artificial Intelligence, AAAI,* 1992.

[141] L. Breiman , J. Friedman , C. J. Stone and R. A. Olshen, Classification and Regression Trees, New York: Chapman and Hall, 1984, p. 368.

[142] J. R. Quinlan, "Induction of Decision Trees," *Machine Learning,* vol. 1, pp. 81-106, 1986.

[143] J. R. Quinlan, C4.5: Programs for Machine Learning, San Mateo: Morgan Kaufmann Publishers, 1993.

[144] H. Schulz, B. Waldvogel, R. Sheikh and S. Behnke, "CURFIL: Random Forests for Image Labeling on GPU," in *Proceedings of the 10th International Conference on Computer Vision Theory and Applications,* Berlin, 11-14 Mar, 2015.

[145] J. Ylioinas, J. Kannala, A. Hadid and . M. Pietikainen, "Learning Local Image Descriptors Using Binary Decision Trees," in *Proceedings of IEEE Winter Conference on Applications of Computer Vision (WACV 2014),*, Steamboat Springs, CO, USA,, 2014.

[146] B. Bouchon-Meunier and C. Marsala, "Fuzzy decision tree and databases," in *Flexible Query Answering Systems,* T. Andreasen, H. Christiansen and H. L. Larsen, Eds., Kluwer Academic Publisher, 1997, pp. 277-288.

[147] J. D. M. Rennie, L. Shih, J. Teevan and D. R. Karge, "Tackling the Poor Assumptions of Naive Bayes Text Classifiers," in *Proceedings of the 20th International Conference on Machine Learning,* Washington, DC, USA, 2003.

[148] N. M. Murty and S. V. Devi, Pattern Recognition: An Algorithmic Approach, vol. z serii Undergraduate Topics in Computer Science, Springer Science & Business Media, 2011, p. 263.

[149] L. Wang, Ed., Support Vector Machines: Theory and Applications, Berlin: Springer, 2005, p. 450.

[150] H. Ishibuchi and Y. Nojima, "Toward Quantitative Definition of Explanation Ability of Fuzzy Rule-Based Classifiers," in *IEEE International Conference on Fuzzy Systems,* Taipei, Taiwan, June 27-39, 2011.

[151] H. Ishibuchi and T. Yamamoto, "Rule weight specification in fuzzy rule-based classification systems," *IEEE Transactions on Fuzzy Systems,* vol. 13, no. 4, pp. 428-435, 2005.

[152] K. Nozaki, H. Ishibuchi and H. Tanaka , "Adaptive fuzzy rule-based classification systems," *IEEE Transactions on Fuzzy Systems,* vol. 13, no. 4, pp. 238-250, 1996.

[153] H. Ishibuchi and Y. Nojima, "Toward Quantitative Definition of Explanation Ability of Fuzzy Rule-Based Classifiers," in *IEEE International Conference on Fuzzy Systems,* Taipei, Taiwan, June 27-39, 2011.

[154] T. Jaworska, "Application of Fuzzy Rule-Based Classifier to CBIR in comparison with other classifiers," in *11th International Conference on Fuzzy Systems and Knowledge Discovery*, Xiamen, China, 2014.

[155] S. K. Candan and W.-S. Li, "On Similarity Measures for Multimedia Database Applications," *Knowledge and Information Systems,* vol. 3, pp. 30-51, 2001.

[156] A. Hamilton-Wright and D. W. Stashuk, "Constructing a Fuzzy Rule Based Classification System Using Pattern Discovery," in *Annual Meeting of the North American Fuzzy Information Processing Society*, 2005.

[157] Y. LeCun, Y. Bengio and G. Hinton, "Deep learning," *Nature,* vol. 521, pp. 436-444, 28 May 2015.

[158] C. Olah, „Conv Nets: A Modular Perspective," blog, July 2014. [Online]. Available: http://colah.github.io/posts/2014-07-Conv-Nets-Modular/.

[159] A. Krizhevsky, I. Sutskeve and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems*, 2012.

[160] MathWorks Inc., "Deep learning with MATLAB," 2016. [Online]. Available: https://www.mathworks.com/discovery/deep-learning.html.

[161] C.-C. Chang and T.-C. Wu, "An exact match retrieval scheme based upon principal component analysis," *Pattern Recognition Letters,* vol. 16, pp. 465-470, 1995.

[162] D. S. Guru and P. Punitha, "An invariant scheme for exact match retrieval of symbolic images based upon principal component analysis," *Pattern Recognition Letters* , vol. 25, p. 73–86, 2004.

[163] S. Rolewicz, Functional Analysis and Control Theory: Linear Systems, vol. Series: Mathematics and its applications, Warsaw: PWN-Polish Scientific Publishers , 1987.

[164] J. Z. Wang, J. Li and G. Wiederhold, "SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture LIbraries," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE,,* vol. 23, no. 9, pp. 947-963, Sep. 2001.

[165] C. Mallows, "A Note on Asymptotic Joint Normality," *The Annals of Mathematical Statistics,* vol. 43, no. 2, pp. 508-515., 1972.

[166] D. Zhou, J. Li and H. Zha, "A new Mallows distance based metric for comparing clusterings," in *Proceedings of the 22nd International Conference on Machine Learning,*, Bonn,m Germany, Aug. 2005.

[167] E. Pękalska and R. P. Duin, The Dissimilarity Representation for Pattern Recognition. Foundations and Applications., 1 ed., Vols. Series in Machine Perception and Artificial Intelligence - Vol. 64, New Jersey, London: World Scientific, 2005, p. 607.

[168] B. Ko and H. Byun, "Integrated Region-Based Image Retrieval Using Region's Spatial Relationships," in *Proceedings of 16th International Conference on Pattern Recognition*, 11-15 Aug. 2002.

[169] C. Beecks, M. S. Uysal and T. Seidl, "A Comparative Study of Similarity Measures for Content-Based Multimedia Retrieval," in *Multimedia and Expo (ICME)*, Suntec City, 19-23 July, 2010.

[170] T. Jaworska, "A Search-Engine Concept Based on Multi-Feature Vectors and Spatial Relationship," in *Flexible Query Answering Systems*, vol. 7022, H. Christiansen, G. De Tré, A. Yazici, S. Zadrożny and H. L. Larsen, Eds., Ghent, Springer, 2011, pp. 137-148.

[171] T. Jaworska, "An Asymmetric Approach to Signature Matching," in *Multimedia and Network Information Systems*, vol. 506, A. Zgrzywa, K. Choraś and A. Siemiński, Eds., Wrocław, Springer, 2016, pp. 27-37.

[172] G. Wu, E. Y. Chang and N. Panda, "Formulating context-dependent similarity functions," in *The 13th annual ACM international conference on Multimedia*, Singapore , Nov., 2005.

[173] A. Natsev and J. R. Smith, "A study of image retrieval by anchoring," in *IEEE International Conference on Multimedia and Expo*, Lausanne, Switzerland , Aug. 2002.

[174] C.-T. Nguyen, X. Wang, J. Liu and Z.-H. Zhou, "Labeling Complicated Objects: Multi-View Multi-Instance Multi-Label Learning," in *28th AAAI Conference on Artificial Intelligence*, Hilton Québec Canada, June, 2014.

[175] H. Mueller, W. Mueller, S. Marchand-Maillet and T. Pun, "A Framework for Benchmarking in CBIR," *Multimedia Tools and Applications,* no. 21, pp. 55-73, 2003.

[176] D. A. Narasimhalu , M. S. Kankanhalli and J. Wu, "Benchmarking Multimedia Databases," *Multimedia Tools and Applications,* vol. 4, no. 3, p. 333–356, May 1997.

[177] J. R. Smith, "Image retrieval evaluation," in *IEEE Workshop on Content-Based Access of Image and Video Libraries (CBAIVL '98)*, Santa Barbara, 1998.

[178] A. Dimai , "Assessment of effectiveness of content-based image retrieval systems," in *3rd International Conference on Visual Information Systems (VISUAL '99)*, Amsterdam, The Netherlands, 1999.

[179] E. L. van den Broek, T. Kok, T. E. Schouten and L. G. Vuurpijl, "Human-Centered Content-Based Image Retrieval," in *Proceedings of XIII Conference on Human Vision and Electronic Imaging*, Feb. 14, 2008.

[180] M. Everingham, A. S. Eslami, L. Van Gool, C. K. I. Williams, J. Winn and A. Zisserman, "The PASCAL Visual Object Classes Challenge: A Retrospective," *International Journal of Computer Vision,* no. 111, p. 98–136, 2015.

[181] Corel comp., "The COREL Database for Content based Image Retrieval".

[182] Z. Yang and C.-C. Jay Kuo, "Learning image similarities and categories from content analysys and relebance feedback," in *Proceedings of the ACM Multimedia Workshops. Multimedia00'*, Los Angeles, CA, USA , Oct 30 - Nov 03, 2000.

[183] the Eastman Kodak Company, [Online]. Available: http://r0k.us/graphics/kodak/.

[184] D.-C. He and A. Safia, "Multiband Texture Database," 2015. [Online]. Available: http://multibandtexture.recherche.usherbrooke.ca/.

[185] D.-C. He and A. Safia, "New Brodatz-based Image Databases for Grayscale Color and Multiband Texture Analysis," *ISRN Machine Vision,* vol. Article ID 876386, pp. 1-14, 2013.

[186] N. Rasiwasia, P. J. Moreno and N. Vasconcelos, "Bridging the Gap: Query by Semantic Example," *IEEE TRANSACTIONS ON MULTIMEDIA,* vol. 9, no. 5, pp. 923-938, Aug 2007.

[187] X. Wang, S. Qiu, K. Liu i X. Tang, „Web Image Re-Ranking Using Query-Specific Semantic Signatures," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE,* tom 36, nr 4, pp. 810-823, April 2014.

[188] M. Everingham, L. Van Gool, C. K. I. Williams, A. Zisserman, J. Winn, A. S. Eslami and Y. Aytar , "The PASCAL Visual Object Classes Homepage," 2015. [Online]. Available: http://host.robots.ox.ac.uk/pascal/VOC/index.html.

[189] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," in *IEEE Conference on Computer Vision and Pattern Recognition*, Miami, USA, June, 2009.

[190] L. Fei-Fei, K. Li, O. Russakovsky, J. Krause, J. Deng and A. Berg, "ImageNet," Stanford Vision Lab, Stanford University, Princeton University, 2014. [Online]. Available: http://www.image-net.org/.

[191] G. Griffin, A. D. Holub and P. Perona, "The Caltech 256," California Institute of Technology, Los Angeles, 2006.

[192] G. Griffin, "Caltech256," 2006. [Online]. Available: http://www.vision.caltech.edu/Image_Datasets/Caltech256/.

[193] J. Philbin, O. Chum and M. a. S. J. a. Z. A. Isard, "Object Retrieval with Large Vocabularies and Fast Spatial Matching," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2007.

[194] J. Philbin, R. Arandjelović and A. Zisserman, "The Oxford Buildings Dataset," Department of Engineering Science, University of Oxford, Nov 2012. [Online]. Available: http://www.robots.ox.ac.uk/~vgg/data/oxbuildings/.

[195] J. Philbin, O. Chum and M. a. S. J. a. Z. A. Isard, "Lost in Quantization: Improving Particular Object Retrieval in Large Scale Image Databases," in *IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, USA, 23-28 June, 2008.

[196] J. Philbin i A. Zisserman, „The Paris Dataset," Visual Geometry Group, Department of Engineering Science, University of Oxford , 2008. [Online]. Available: http://www.robots.ox.ac.uk/~vgg/data/parisbuildings/.

[197] B. C. Becker, "PubFig83 + LFW Dataset," 2015. [Online]. Available: http://www.briancbecker.com/blog/research/pubfig83-lfw-dataset/.

[198] B. C. Becker and E. G. Ortiz, "Evaluating Open-Universe Face Identification on the Web," in *CVPR 2013, Analysis and Modeling of Faces and Gestures Workshop.*, Portland, Oregon, USA, 23-28 June, 2013.

[199] P.-S. P. Chen, "Entity-relationships model – Toward a Unified View of Data," *ACM Transactions on Database Systems,* vol. 1, no. 1, pp. 9-36, 1976.

[200] R. Barker, Entity-Relationship Modelling. Case MethodSM, London, : Addison-Wesley, 1995.

[201] R. Barker and C. Longman , Function and Process Modelling. Case MethodSM, London: Addison-Wesley Pub. Co., 1993.

[202] K. Rodden and K. R. Wood, "How Do People Manage Their Digital Photographs?," in *SIGCHI Conference on Human Factors in Computing Systems*, Ft. Lauderdale, Florida, USA., April 5–10, 2003.

[203] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain, "Content-Based Image Retrieval at the End of the Early Years," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGEN,* vol. 22, no. 12, pp. 1349 - 1380, Dec 2000.

[204] X. Wang, K. Liu and X. Tang, "Query-Specific Visual Semantic Spaces forWeb Image Re-ranking.," in *Computer Vision and Patern Recognition Paper*, 2011.

[205] W. Niblack, M. Flickner, D. Petkovic, P. Yanker, R. Barber, W. Equitz, E. Glasman, C. Faloutsos and G. Taubin, "The QBIC Project: Querying Images by Content Using Colour, Texture and Shape," *SPIE,* vol. 1908, pp. 173-187, 1993.

[206] B. Xiao , X. Gao, D. Tao and X. Li, "Recognition of Sketches in Photos," in *Multimedia Analysis, Processing and Communications*, vol. 346, W. Lin, D. Tao, J. Kacprzyk , Z. Li, E. Izquierdo and H. Wang , Eds., Berlin, Springer-Verlag, 2011, pp. 239-262.

[207] J.-H. Lim and J. S. Jin, "A structured learning framework for content-based image indexing and visual query," *Multimedia Systems,* vol. 10, p. 317–331, 2005.

[208] J. Assfalg, A. Del Bimbo and P. Pala, "Three-Dimensional Interfaces for Querying by Example in Content-Based Image Retrieval," *IEEE Transactions on Visualization and Computer Graphics* , vol. 8, no. 4, pp. 305-318, Oct-Dec 2002.

[209] J. Fauqueur and N. Boujemaa, "Mental image search by boolean composition of region categories," *Multimed Tools and Applications,* vol. 31, p. 95–117, 2006.

[210] T. Jaworska, "Multi-criteria object indexing and graphical user query as an aspect of content-based image retrieval system.," in *Information Systems Architecture and Technology,* L. Borzemski, A. Grzech, J. Świątek and Z. Wilimowska, Eds., Wrocław, Wrocław Technical University Publisher, 2009, pp. 103-112.

[211] . B. Moghaddam, H. Biermann and D. Marg, "Regions-of-Interest and Spatial Layout for Content-Based Image Retrieval," *Multimedia Tools and Applications,* vol. 14, no. 2, pp. 201-210, June 2001.

[212] M. M. Rahman, S. K. Antani and G. R. Thoma, "A query expansion framework in image retrieval domain based on local and global analysis," *Information Processing and Management,* vol. 47, pp. 676-691, 2011.

[213] J. Fauqueur, "Instantaneous mental image search with range queries on multiple region descriptors," Cambridge, UK, Jan, 2005.

[214] Y. Liu, D. Zhang, G. Lu and W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern Recognition,* vol. 40, pp. 262-282, 2007.

[215] J. C. Cubero, N. Marín, J. M. Medina, E. Pons and A. M. Vila, "Fuzzy Object Management in an Object-Relational Framework," in *Proceedings of the 10th International Conference IPMU,* Perugia, Italy, 4-9 July, 2004.

[216] F. Berzal, J. C. Cubero, J. Kacprzyk, N. Marín, A. M. Vila and S. Zadrożny, "A General Framework for Computing with Words in Object-Oriented Programming.," in *International Journal of Uncertainty. Fuzziness and Knowledge-Based Systems.*, vol. 15 (Suppl), Singapore, World Scientific Publishing Company, 2007, pp. 111 131,.

[217] W. Plant and G. Schaefer, "Visualization and Browsing of Image Databases," in *Multimedia Analysis, Processing and Communications,* vol. 346, W. Lin, D. Tao, J. Kacprzyk , Z. Li, E. Izquierdo and H. Wang, Eds., Berlin, Springer, 2011, pp. 3-57.

[218] K. Rodden, „Evaluating similarity-based visualisations as interfaces for image browsing," University of Cambridge, Cambridge, 2002.

[219] K. Rodden, K. R. Wood, W. Basalaj and D. Sinclair, "Evaluating a Visualisation of Image Similarity as a Tool for Image Browsing," in *IEEE Symposium on Information Visualisation,* 1999.

[220] W. Basalaj, "Proximity visualisation of abstract data," University of Cambridge, Cambridge, 2001.

[221] C. Faloutsos and K. Lin, "Fast Map: A Fast Algorithms for Indexing, Data-Mining and Visualization of Traditional and Multimedia Datasets," in *ACM SIGMOD international conference on Management of data* , New York, USA, May, 1995.

[222] L. F. D. Santos, R. L. Dias and M. X. Ribeiro, "Combining Diversity Queries and Visual Mining to Improve Content-Based Image Retrieval Systems: The DiVI Method," in *IEEE International Symposium on Multimedia,* Miami, Dec. 2015.

[223] A. Bursuc and T. Zaharia, "ARTEMIS@ MediaEval 2013: A Content-Based Image Clustering Method for Public Image Repositories," *ACM Multimedia,* pp. 18-19, Oct. 2013.

[224] C. Chen, G. Gagaudakis and P. Rosin, "Similarity-Based Image Browsing," in *Proceedings of the 16th IFIP World Computer Congress, International Conference on Intelligent Information Processing,* Beijing, China, 2000.

[225] T. Kohonen, "The Self_Organizing Map," *Proceedings of IEEE,* vol. 78, no. 9, pp. 1464-1480, Sep. 1990.

[226] A. Csillaghy , H. Hinterberger and A. B. Benz, "Content-Based Image Retrieval in Astronomy," *Information Retrieval Journal,* vol. 3, no. 3, pp. 229-241, 2000.

[227] Y. Rui and T. S. Huang, "Relevance Feedback Techniques in Image Retrieval," in *Principal of Visual Information Retrieval,* M. S. Lew, Ed., London, Springer, 2001, pp. 219-258.

[228] V. Mezaris, I. Kompatsiaris and M. G. Strintzis, "An ontology approach to object-based image retrieval," in *Proceedings of International Conference on Image Processing ICIP 2003.*, 2003.

[229] A. D. Gudewar and L. R. Ragha, "Ontology to Improve CBIR System," *International Journal of Computer Applications,* vol. 52, no. 21, pp. 23-30, 2012.

[230] C. Doulaverakis, E. Nidelkou, A. Gounaris and Y. Kompatsiaris, "A Hybrid Ontology and Content-Based Search Engine For Multimedia Retrieval," in *Workshop Proceedings in Advances in Databases and Information Systems ADBIS '2006,* Thessaloniki, 2006.

[231] O. Allani, N. Mellouli, H. B. Zghal, H. Akdag and H. B. Ghzala, "A Relevant Visual Feature Selection Approach for Image Retrieval," in *Proceedings of the 10th International Conference on Computer Vision Theory and Applications*, Berline, 11-14 Mar, 2015.

[232] O. Russakovsky and L. Fei-Fei, "Attribute Learning in Large-scale Datasets," in *Proceedings of the 12th European Conference of Computer Vision (ECCV), 1st International Workshop on Parts and Attributes.*, Crete, Greece, 2010.

[233] T. Hofmann, "Probabilistic latent semantic analysis," in *Proceedings of the15th Conference on Uncertainty in Artificial Intelligence*, Stockholm, 1999.

[234] D. M. Blei, A. Y. Ng and M. I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research,* vol. 3, pp. 993-1022, 2003.

[235] L. Fei-Fei and P. Perona, "A Bayesian Heirarcical Model for Learning Natural Scene Categories," in *Computer Vision & Pattern Recognition CVPR*, 2005.

[236] J. Sivic, B. C. Russell, A. A. Efros, A. Zisserman and W. T. Freeman, "Discovering objects and their location in images," in *Proceedings of Internationa Conference of Computer Vision*, Beijing, 2005.

[237] J. Bautista-Ballester, J. Verges-Llahi and D. Puig, "Using Action Objects Contextual Information for a Multichannel SVM in an Action Recognition Approach based on Bag of VisualWords," in *Proceedings of the 10th International Conference on Computer Vision Theory and Applications*, Berlin, 11-14 Mar, 2015.

[238] T. Kinnunen, J.-K. Kamarainen, L. Lensu and H. Kälviäinen, "Unsupervised object discovery via self-organisation," *Pattern Recognition Letters,* no. 33, p. 2102–2112, Aug 2012.

[239] J. Urban, J. M. Jose and C. J. van Rijsbergen, "An adaptive technique for content-based image retrieval," *Multimedial Tools Applied,* no. 31, pp. 1-28, July 2006.

[240] L. Zhang, L. Wang and W. Lin, "Generalized biased discriminant analysis for content-based image retrieval," *IEEE Transactions on System, Man, Cybernetics, Part B - Cybernetics,* vol. 42, no. 1, pp. 282-290, 2012.

[241] L. Zhang, L. Wang and W. Lin, "Semi-supervised biased maximum margin analysis for interactive image retrieval," *IEEE Transactions on Image Processing,* vol. 21, no. 4, pp. 2294-2308, 2012.

[242] S. T. Roweis and L. K. Saul, "Nonlinear Dimensionality Reduction by Locally Linear Embedding," *Science,* vol. 290, no. 5500, pp. 2323-2326, Dec. 2000.

[243] S.-F. Chang, W. Chen and H. Sundaram, "Semantic Visual Templates: Linking Visual Features to Semantics," in *International Conference on Image Processing, 1998. ICIP 98.*, Chicago, 1998.

[244] Y. Zhuang, X. Liu and Y. Pan, "Apply Semantic Template to Support Content-based Image Retrieval," in *the Proceeding of IS&T and SPIE Storage and Retrieval for Media Databases 2000*, San Jose, California, USA, Jan, 2000.

[245] G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross and K. Miller, "Introduction to WordNet: An On-line Lexical Database," *Communications of the ACM,* vol. 38, no. 11, pp. 39-41 , Nov. 1995.

[246] M. Mucha and P. Sankowski, "Maximum Matchings via Gaussian Elimination," in *Proceedings of the 45th Annual Symposium on Foundations of Computer Science (FOCS'04)*, 2004.

[247] Z. Wang , A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image Qualifty Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing,* vol. 13, no. 4, p. 600–612, April 2004.

[248] E. Candes, L. Demanet, D. Donoho and L. Ying, "Fast Discrete Curvelet Transforms," 2006.

[249] I. Aizenberg, N. N. Aizenberg and J. P. Vandewalle, Multi-Valued and Universal Binary Neurons, Springer US, Springer Science+Business Media Dordrecht, 2000, p. 276.

[250] T. Yamashita, T. Watasue, Y. Yamauchi and H. Fujiyoshi, "Improving Quality of Training Samples Through Exhaustless Generation and Effective Selection for Deep Convolutional Neural Networks," in *Proceedings of the 10th International Conference on Computer Vision Theory and Applications,* Berlin, 11-14 Mar, 2015.

[251] F. Jurišić, I. Filković and Z. Kalafatić, "Evaluating the Effects of Convolutional Neural Network Committees," in *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2016),* Rome, Italy, 27-29 Feb, 2016.

[252] H. H. Aghdam, E. J. Heravi and D. Puig, "Analyzing the Stability of Convolutional Neural Networks against Image Degradation," in *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2016),* Rome, Italy, 27-29 Feb, 2016.

[253] S. Srinivasulu and P. Sakthivel , "Extracting Spatial Semantics in Association Rules for Weather Forecasting Image," in *Trendz in Information Sciences & Computing(TISC2010),* Chennai, 17-19 Dec. 2010.

[254] A. Moumtzidou, V. Epitropou, S. Vrochidis, K. Karatzas, S. Voth, A. Bassoukos, J. Moßgraber, A. Karppinen, J. Kukkone and I. Kompatsiaris, "A model for environmental data extraction from multimedia and its evaluation against various chemical weather forecasting datasets.," *Ecological Informatics,* no. 23, p. 69–82, Sep. 2014.

[255] K. Choroś, "False and Miss Detections in Temporal Segmentation of TV Sports News Videos - Causes and Remedies," in *New Research in Multimedia and Internet Systems,* Advances in Intellignet Systems and Computing ed., vol. 314, A. Zgrzywa, . K. Choroś and A. Siemiński, Eds., Wrocław, Springer, 2015, pp. 35-46.

[256] J. Li, „The application of CBIR-based system for the product in electronic retailing," w *2010 IEEE 11th International Conference on Computer-Aided Industrial Design & Conceptual Design (CAIDCD),* Yiwy, China, 17-19 Nov. 2010.

[257] G. De Tre, D. Vandermeulen, J. Hermans, P. Claes, J. Nielandt and A. Bronselaer, "Bipolar Comparison of 3D Ear Models," in *Information Processing and Management of Uncertainty in Knowledge-Based Systems - 15th International Conference - IPMU,* Montpellier, France, 2014.

[258] A. E. Carpenter, "Extracting Rich Information from Images," in *Cell-Based Assays for High-Throughput Screening,* P. A. Clemons, N. J. Tolliday and B. K. Wagner , Eds., Springer, 2009, pp. 193-211.

[259] M. Mansourvar and M. A. Ismail, "Content-Based Image Retrieval in Medical Systems," *International Journal of Information Technology,* vol. 20, no. 2, pp. 1-9, 2014.

[260] A. Obero and M. Singh, "Content Based Image Retrieval System for Medical Databases (CBIR-MD) - Lucratively tested on Endoscopy, Dental and Skull Images," *IJCSI International Journal of Computer Science Issues,* vol. 9, no. Issue 3, No 1, May 2012.

[261] M. S. Chaibou and K. Kalti, "A New Labeled Quadtree-based Distance for Medical Image Retrieval," in *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2016),* Rome, Italy, 27-29 Feb., 2016.

[262] H.-s. Kim, H.-W. Chang, H. Liu, J. Lee and D. Lee, "BIM: IMAGE MATCHING USING BIOLOGICAL GENE SEQUENCE ALIGNMENT," 2010. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5414214.

[263] A. T. Inc., "image pattern recognition using vector quantization - uszczegółowić". the United States Patent and Trademark Office Patent 7,502,519, 2009.

[264] J. Mallik, A. Samal and S. L. Gardnerb, "A content based image retrieval system for a biological specimen collection," *Computer Vision and Image Understanding,* vol. 114, no. 7, p. 745–757, July 2010.

[265] G. Csurka, J. Ah-Pine and S. Clinchant, "Unsupervised Visual and Textual Information Fusion in CBMIR Using Graph-Based Methods," *ACM Transactions on Information Systems,* vol. 33, no. 2, pp. 9:1--9:31, Feb, 2015.

[266] L. Anselin and S. J. Rey, Eds., Perspectives on Spatial Data Analysis, Berlin: Springer, 2010, p. 290.

[267] C. Hahne, A. Aggoun, S. Haxha, V. Velisavljevic and J. C. J. Fernández, "Light field geometry of a standard plenoptic camera," *Optics Express,* vol. 22, no. 22, pp. 26659-26673, Nov. 2014.

[268] S. Cloix, T. Pun and D. Hasler, "Real-time Scale-invariant Object Recognition from Light Field Imaging," in *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2016),* Rome, Italy, 27-29 Feb., 2016.

[269] *IEEE Transactions on Image Processing,* vol. 13, no. 3, p. all, March 1994.

[270] S. Lyu, D. Rockmore i H. Farid, „A digital technique for art authentication," *Proceedings of the National Academy of Sciences of the United States of America,* tom 101, nr 49, p. 17006–17010, 7 Dec. 2004.

[271] M. Aubry, B. C. Russell and J. Sivic, "Painting-to-3D Model Alignment Via Discimanative Visual Elements," *ACM Transactions on Graphics,* vol. 28, no. 4, pp. 1-14, Article No. 106 , Aug. 2009.

[272] J. K. Gilbert, Ed., Visualization in Science Education, Springer Science & Business Media, 2006, p. 346.

[273] E. Alepis and M. Virvou, Object-Orianted User Interfaces fro Personalized Mobile Learning, vol. 64, J. Kacprzyk and J. C. Lakhimi, Eds., Heidelberg: Springer, 2014, p. 129.

[274] G. Ghiani, M. Manca and F. Paternò, "Authoring Context-dependent Cross-device User Interfaces based on Trigger/Action Rules," in *The 14th International Conference on Mobile and Ubiquitous Multimedia,* Linz, Austria, 30 Nov. - 2nd Dec. 2015.

[275] Z. Raisi, F. Mohanna and M. Rezaei, "Applying Content-Based Image Retrieval Techniques to Provide New Services for Tourism Indusry," *International Journal of Advanced Networking and Applications,* vol. 6, no. 2, pp. 2222-2232, Oct. 2014.

[276] W. Premchaiswadi, "An Image Search for Tourist Information Using a Mobile Phone," *WSEAS Transactions on Information Science and Applications,* vol. 4, no. 7, pp. 532-541, Apr 2010.

[277] M. Markkula and E. Sormunen, "Searching for Photos - Journalists' Practices in Pictorial IR," in *Electronic Workshops in Computing – Challenge of Image Retrieval,* Newcastle, UK,, Feb. 1998.

[278] D. Gurari, S. D. Jain, M. Betke and K. Grauman, "Pull the Plug? Predicting If Computers or Humans Should Segment Images," in *the IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* Las Vegas, June, 2016.

[279] R. Datta, T. Joshi, J. Li and J. Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of the New Age," *ACM Computing Surveys,* vol. 40, no. 2, pp. 5:1-5:60, Apr. 2008.

[280] B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian Motions, Fractional Noises and Applications," *SIAM Review,* vol. 10, no. 4, pp. 422-437, October 1968.

[281] A. Kundu and J.-L. Chen, "Texture classification using QMF bank-based subband decomposition," *CVGIP: Graphical Models and Image Processing,* vol. 54, no. 5, p. 369–384, 1992.

[282] C. Xu and J. L. Prince, "Snakes, Shapes, and Gradient Vector Flow," *IEEE TRANSACTIONS ON IMAGE PROCESSING,* vol. 7, no. 3, pp. 359-369, March 1998.

[283] "Fast Wavelet-Based Image Characterization for Highly Adaptive Image Retrieval," *IEEE Transactions on Image Processing,* 2012.

[284] D. Eads, D. Helmbold and E. Rosten, "Boosting in Location Space," Santa Cruz, 2013.

[285] C. Faloutsos, R. Barber, M. Flickner, J. Hafner, W. Niblack and D. Petkovic, "Efficient and Effective Querying by Image Content.," *Journal of Intelligent Information Systems,* vol. 3, pp. 231-262, 1994.

[286] M. Koyuncu and B. Cetinkaya, "A Component-Based Object Detection Method Extended with a Fuzzy Inference Engine," in *Proceedings of the International Conference on Fuzzy Systems Fuzz-IEEE2015,* Istambul, 2015.

[287] J. Philbin, O. Chum and M. a. S. J. a. Z. A. Isard, "Object Retrieval with Large Vocabularies and Fast Spatial Matching," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* 2007.

[288] K. Chen, "Deep and Modular Neural Networks," in *Handbook Computational Intelligence,* 1 ed., J. Kacprzyk and W. Pedrycz, Eds., Berlin, Springer, 2015, pp. 473-494.

[289] A. Huneiti and M. Daoud, "Content-Based Image Retrieval Using SOM and DWT," *Journal of Software Engineering and Applications,* no. 8, pp. 51-61, Feb 2015.

[290] L. Deng and D. Yu, "Deep Learning Methods and Applications," in *Foundations and Trends in Signal Processing,* Vols. 7, nos. 3–4, Now the essence of knowledge, 2014, p. 197–387.

[291] J. Bautista-Ballester, J. Verges-Llahi and D. Puig, "Using Action Objects Contextual Information for a Multichannel SVM in an Action Recognition Approach based on Bag of VisualWords," in *Proceedings of the 10th International Conference on Computer Vision Theory and Applications,* Berlin, 11-14 Mar, 2015.

[292] O. Allani, N. Mellouli, H. B. Zghal, H. Akdag and H. B. Ghzala, "A Relevant Visual Feature Selection Approach for Image Retrieval," in *VISAPP 2015 - International Conference on Computer Vision Theory and Applications,* Berlin, 2015.

[293] R. K. Srihari , "Automatic indexing and content-based retrieval of captioned images," *IEEE Computer,* pp. 49 - 56, Sep. 1995.

[294] Y. Liu, D. Zhang, G. Lu and W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern Recognition,* vol. 40, pp. 262-282, 2007.

[295] S. K. Pal and P. Mitra, Pattern Recognition Algorithms for Data Mining. scalability, Knowledge Discovery and Soft Granular Computing., London, New York: Chapman and Hall CRC Press Company, 2004, p. 244.

[296] C. Beecks, M. S. Uysal and T. Seidl, "Signature Quadratic Form Distances fer Content-Based Similarity," in *ACM Multimedia,* Beijing, China, Oct. 19-24, 2009.

[297] H. E. Hurst, „Long-term storage capacity of reservoirs," *Transactions of the American Society of Civil Engineers,* pp. 770-808, 1951.

[298] N. Sebe and M. S. Lew, "Texture Features for Content-Based Retrieval," in *Principles of Visual Information Retrieval,* M. S. Lew, Ed., Springer Science & Business Media, 2013, pp. 50-81.

[299] I. Rish, "An empirical study of the naive Bayes classifier," in *IJCAI-2001 workshop on Empirical Methods in AI,* 2001.

[300] R. Datta, J. Li and J. Z. Wang, "Content-Based Image Retrieval - Approaches and Trends of the New Age," in *Multimedia Information Retrieval (MIR '05)*, Singapour, 2005.

[301] T. Jaworska, "The Concept of a Multi-Step Search-Engine for the Content-Based Image Retrieval Systems," in *Information Systems Architecture and Technology. Web Information Systems Engineering, Knowledge Discovery and Hybrid Computing*, Wrocław, 2011.

[302] Z. Wang , A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, "Image Qualifty Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, p. 600–612, April 2004.

Index

List of Figures