A SURVEY ON IMAGE PROCESSING TECHNIQUES IN THE CONTEXT OF CONTENT BASED IMAGE RETRIEVAL

Ilie Mihai Bogdan[1]

[1]*"Dunarea de Jos" University of Galati, Faculty of Automatic Control, Computers, Electrical and Electronics Engineering, Stiintei Street 2nd, Galati, Romania*
mihai.ilie@ugal.ro

ABSTRACT

This paper presents an overview of the current image processing techniques used to solve various problems, especially in the field of image indexing/retrieval. We aim to provide a solid, comprehensive reference source for other researchers involved in the image processing field. We will be describe the search spaces, the most common techniques, their advantages and drawbacks and the alternatives proposed by various researchers in the community.

KEYWORDS: CBIR, image processing, image descriptors

1. INTRODUCTION

In the domains of electrical engineering and computer science, the image processing problem is related to any signal processing technique, where the input is actually an image/a picture and the output is represented by a different image or a set of characteristics.

The image classification according to their content is known as Content-Based Image Retrieval (CBIR), Query By Image Content (QBIC) or Content-Based Visual Information Retrieval (CBVIR) and represents a solution to the problem of finding similar digital images in a large data set. The "Content-Based" term means that the CBIR techniques use the real image content, as opposed to data stored in metadata, keywords, tags or other types of description associated to the analyzed image.

In order to analyze a specific image, the image processing algorithms must extract a certain number of characteristics, or features. These are split in two major categories:
- textual;
- visual.

The textual ones are usually stored in annotations and in the metadata associated to the image - labels, keywords, recording date, file name, exposure, flash etc.

The visual features are extracted through pixel analysis and contain information related to colour, texture, shape and spatial position. These characteristics are again split in two major categories:

- global;
- local.

The global ones describe the whole image and contain mediated information obtained after analyzing all the pixels (average light intensity, average presence of a certain colour etc.) The local features describe a specific part of the image, obtained through a segmentation process. Before determining these characteristics, the algorithms establish the image key points, which will be used as centres for the feature extraction techniques.

All the image processing applications are classified according to two criteria:
- computing time;
- precision.

Precision is something easy to understand, described by computing the result correctness, after applying a certain algorithm on a image set.

The computing time can be affected by a lot of factors. The sum of all the characteristics determines a vector which describes the image. This vector has very high dimensions, in what regards its length and the data contained in it. As a solution, researchers are trying to reshape the data, because it is not very dense. Also there is the possible solution of multidimensional indexing through:
- KD trees;
- VP trees;
- hashing.

Image processing is used mainly in two distinct areas:

This paper was recommended for publication by Sergiu Caraman

- enhancing the image quality to make it easier to be analyzed by the human observer - in this category we mention image printing and transmission;
- extracting structures and characteristics in order to identify edges, unique light intensity, texture, colour, or a combination of these. Shape is a characteristic which is harder to describe; nevertheless there are possibilities to describe it mathematically. All the above are usually steps taken in order to subsequently describe, identify or classify certain objects.

This paper is targeting the second category, especially in the context of CBIR. In this area, the image processing techniques have been combined and extended in order to cover a large application area, starting with those who target the common user (image collections, duplicate detection, identifying images that match a certain colour setup, detecting images similar with a certain one, provided as input etc.), up to very complex systems used in the industrial area (detecting products that do not match a specific standard), copyright, security (face and print recognition), military (tracking, aiming, digital reconstruction of a certain area) or for weather forecast. The medical applications represent a particular CBIR area, where the classification process is usually doubled by the expert human intervention.

The approaches vary a lot, depending on the problem that needs to be addressed. Some researchers have chosen to use a single type of descriptor (texture based - (Mauricio Correa, 2011), colour based - (Thomas Breuer, 2012), wavelets - (X. Cao, 2012)) in order to recognize the human body and the human skin. Other approaches are very complicated and use multiple sets of descriptors, combined with various artificial intelligence techniques in order to achieve their goal. As an example, (Luis Caro, 2012) and (Kashif Iqbal, 2012) use a combination of descriptors originating in different search spaces (colour, texture, local descriptors) in order to control a robot or to ensure biometrical security. Sometimes the authors needs to cover multiple areas of image processing in order implement the solution. As an example, a group of South Korean researchers (Young-Beom Lee, 2012) use various types of descriptors and an OCR (optical character recognition) module in order to correctly identify and classify the pills being transported on a conveyor belt.

In section 2 we will present the influence of the human factor in the image processing area, as well as the most recent approaches in extracting descriptors from all the involved spaces (sections 3, 4, 5). Subsequently we will present the most common segmentation techniques (section 6) and the extraction of local descriptors (section 7). We will conclude by presenting our conclusions and the future research directions.

2. THE HUMAN FACTOR

"A picture is worth a thousand words."

One of the factors that determined the necessity of the image processing field was the human one. Initially, the image classification was done by using labels. Besides the obvious disadvantages (imprecise, time consuming), one of the major problems risen by this approach was the subjectivity of those who were assigning the labels. Biology is still struggling to understand the optical model for the human eye, as well as the way the brain interprets and associates the data with previously collected information.

One of the most common problems in this area is establishing the meaning of a certain image, even if the descriptors are very similar. As an example, in the below image the shape descriptors are basically the same, but a human observer will easily distinguish the different objects depicted.
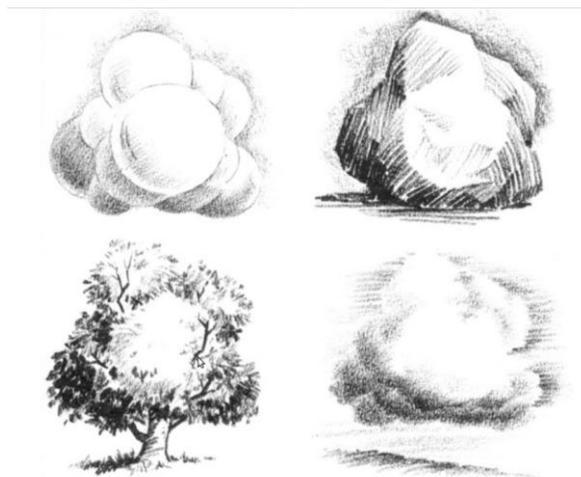
Figure 1: Establishing the meaning of images

Another common problem is rebuilding the information, in the absence of certain features. Actually, in the field of image processing there is still no solution for problems like occlusion and clutter.

Figure 2: Rebuilding information (Koenderink, 2012)

As we mentioned before, most of the state of the art algorithms require a segmentation stage in order to split the image in representative areas, which will later be analyzed.

This is another major discrepancy, as the human brain is able to correctly identify and classify the objects in a certain image even if the object boundaries are not clear or even completely missing.

Figure 3: Lack of object boundaries (Koenderink, 2012)

More than that, the human brain can easily establish additional information, like the object state. As an example, in the image below a human observer can detect the dynamical state of the objects, even if the pictures depict static statues.

Figure 4: Dynamical state (Koenderink, 2012)

However, there are many situations that may confuse the observer; most of these errors are caused by:

- the lack of the third dimension can influence the way that we perceive the perspective and the size of certain objects;
- since the brain works according to its previous training, its results are affected by associating the subject image with previously recorded information - sometimes this process can produce errors, as seen in the image below;

Figure 5: Erroneous classification (Koenderink, 2012)

- another important problem encountered when implementing image processing systems is the cultural gap:
  - the face recognition engines trained on Caucasian people have very bad results when used on Asian faces;
  - people who write from right to left have a different understanding of direction than people who write from left to right;
- the above mentioned object reconstruction process can malfunction;
- sometimes the brain works better on images affected by noise, as shown in the image below.
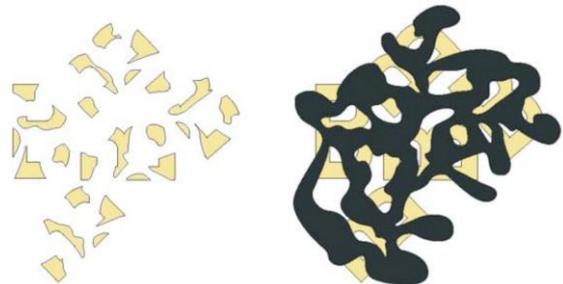
Figure 6: Image affected by noise (Koenderink, 2012)

In this context, the CBIR engines are often exposed to the public, in order to collect human relevant data and correctly weigh the results produced by a certain image processing algorithm.

3. COLOUR SPACE TECHNIQUES

The most basic features are extracted from the colour space. In what follows, we will describe the correlation between colour, as perceived by the human observer and as interpreted and stored by the machine. Also, we will describe the most common colour spaces used in the current research activity and the techniques used to extract the descriptors.

There are many approaches in what regards storing the colour information in a pixel, but all the authors agree on the below tuple:

- light (angle, colour, etc.);
- object (shape, surface, colour etc.);
- sensor (exposure, resolution etc.);

Therefore, the same object, photographed in the different circumstances can be erroneously classified, due to several factors, like:
- the camera quality;
- the light source;
- background colour;
- the surface type (coloured, shiny, mated etc.);
- the observer angle, etc.

Because of the different absorption curves of the cone cells in the human eye, the colours are perceived as different combinations of the basic colours *red-green-blue*. If mixed, we can obtain the secondary colours - *magenta* (M=R+B), *cyan* (C=G+B) and *yellow* (Y=R+G).

The RGB quantities needed to create a certain colour are named tri-stimulus values and are represented in the Cartesian system below:

$$x = \frac{X}{X+Y+Z}$$
$$y = \frac{Y}{X+Y+Z}$$
$$z = \frac{Z}{X+Y+Z}$$
$$x + y + z = 1$$

Equation 1: Tri-stimulus values

In 1931, CIE (The International Commission of Illumination) introduced the chromatic diagram, according to which blue is defined as z=1-(x+y).
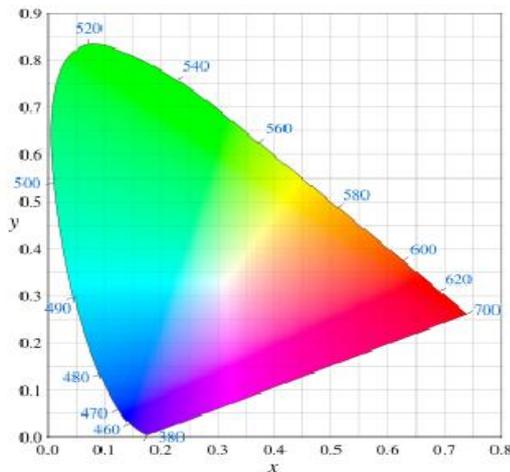
Figure 7: CIE chromatic diagram

The purpose of a certain colour model/system is to facilitate the process of identifying a colour in a standardized way. This is why the researchers around the world reached the conclusion that such a colour model must obey the below equation:

$$q_k = \int_\lambda e(\lambda)\rho(\lambda)^y f_k(\lambda) d\lambda,$$

where $e(\lambda)$ represents the illumination spectre,
$\rho(\lambda)^y$ represents the surface reflexion in the point p,

$f_k(\lambda)$ represents the spectral response
q_k represents the camera output
Equation 2: Colour model

The most common colour spaces are:
- RGB - for cameras and monitors;
- CMY/CMYL - for printers;
- HSI/HSV/HSL/HSB - image processing;
- CIE Lab - image processing.

A specific area in the process of calculating colour features is the extraction of colour invariant characteristics. In order to facilitate the calculus, most of the existing approaches start from a certain series of suppositions (Sebe N. , 2007):
- the light is always white;
- the absence of colour leads to grey;
- in some cases the assumption is that certain objects are mate.

For mated objects, the researchers use different normalisations of the RGB space, according to the below formula.

$$c = \frac{\sum R^p G^q B^r}{\sum R^s G^t B^u},$$

where p+q+r = s+t+u
Equation 3: RGB invariance

According to the targeted area, there are some colour spaces which filter out the shadows and the highlight areas. Two of the most famous implementations are c1c2c3 and l1l2l3.

c1c2c3 has the property of eliminating both shadows and the highlight areas. It is described by the equations below:

$$c_1 = arctg \frac{R}{\max(G,B)}$$
$$c_2 = arctg \frac{G}{\max(R,B)}$$
$$c_3 = arctg \frac{B}{\max(G,R)}$$

Equation 4: The c1c2c3 colour space

l1l2l3 eliminates the shadow areas as well, but preserves the highlight ones. It is described by the equations below:

$$l_1(R,G,B) = \frac{(R-G)^2}{(R-G)^2 + (R-B)^2 + (B-G)^2}$$
$$l_2(R,G,B) = \frac{(R-B)^2}{(R-G)^2 + (R-B)^2 + (B-G)^2}$$
$$l_3(R,G,B) = \frac{(B-G)^2}{(R-G)^2 + (R-B)^2 + (B-G)^2}$$

Equation 5: The l1l2l3 colour space

Probably the most common and well known method of extracting colour features from an image is the colour histogram. The histograms actually represent ca collection of statistics related to colour and intensity, obtained through

processing the values of each pixel. According to the targeted area, the histograms may have different quantisation values, but the most common approach is to use 256 colour zones.

In 2001 the researchers (Wang, 2001) introduced the notions of GCH (Global Colour Histogram) and LCH (Local Colour Histogram), in order to have more precision when comparing two images.

Another popular method to extract colour features is to use the colour moments. These techniques consider the pixel context in the whole image; mainly they calculate the average and standard deviation of each pixel. These descriptors are invariant to scaling and rotation. Even so, there are several attempts to improve the quality of these descriptors through fuzzy techniques on some pre-defined areas (Onur Küçüktunç, 2010) or by increasing the number of descriptors according to the colour space (Hui Yu, 2002).

The comparative studies (Vassilieva, 2012) provide better results for the colour moments techniques, in the context of CBIR engines. It is considered to be so because of the metric distance used to compute the similarity score. As a solution, the researchers are currently attempting to use cumulative histograms (Wikipedia, Cumulative histogram, 2012), which introduce an additional matrix, which contains similarity coeficients between colours. Another approach is to use fuzzy histograms.

4. TEXTURE SPACE TECHNIQUES

The texture is the intrinsic property of a certain surface to describe repetitive visual structures, with the same homogeneity properties. The texture may contain important information on the surface and can describe the relation between it and the surrounding environment (Siddique, 2002). The textural descriptors can describe:

- granulation;
- contrast;
- direction;
- linearity;
- regularity;
- roughness.

There are four major categories for the algorithms used to determine texture:

- statistical - determine the textures according to the pixel grey levels;
  - general statistics parameters (Siddique, 2002);
  - Haralick co-occurence matrices (Hall-Beyer, 2007);
  - Tamura features (Siddique, 2002) (Techasith, 2002) (Johansson, 2002);
- geometrical - characterize the texture according to some primitive units - texels;
  - Voronoi tesselation features;
  - structural methods;
- spectral - based on Fourier spectre analysis in order to describe the global periodicity of the grey levels;

- wavelets (G. D. Magoulas, 1999) (Wikipedia, Wavelets, 2012);
  - Gabor filters (Wikipedia, Gabor Filter, 2012);
  - ICA filters (H. Borgne, 2004).
- model based - use statistical distribution approaches, starting from random values assigned to a certain pixel (Markov Random Fields - MRF), or from fractals.

The most recent CBIR approaches target MRF, wavelets, Gabor filters and Local Binary Patterns (LBP).

4.1. Wavelets

The textures can be modelled as quasi-repetitive structures in the space-time domain. This thing allows image analysis with a smaller computational effort. By using this transformation, a function which describes a curve line, a signal etc. can be described in terms of dispersion or through other characteristics which vary from narrow to large scales.

While analyzing a signal f(x), it is decomposed according to the equation below:

$$f(x) = \sum_{j,k} \alpha_k \psi_{j,k}(x)$$

where the scaling function and the mother wavelet function are described by

$$\psi_{j,k} = 2^{j/2}\varphi(2^j x - k$$

$$j, k \in \mathbb{Z}, (x)$$

Equation 6: Wavelet decomposition

The most common functions used in this approach are Haar and Daubechies.

4.2. Gabor filters

The Gabor filter is actually a linear filter, used to detect edges (Wikipedia, Gabor Filter, 2012). The representations in frequency and direction domains are very similar to the ones provided by the human eye and are very useful for classifying textures. The Gabor filters are symmetrical - all of them can be generated by dilating and rotating a mother wavelet, as shown in the picture below.
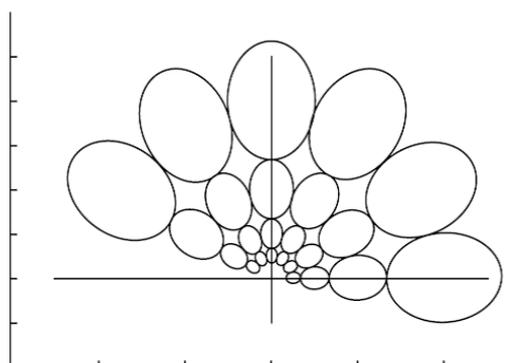
Figure 8: Gabor filters

4.3. Local Binary Patterns

This approach was introduced in 1994 (Wikipedia, Local binary patterns, 2011) by a group of Finnish researchers. The algorithm iterates through the belows steps:
- the whole window is split in NxN cells;
- in each cell, each pixel is compared to its k neighbours;
- wherever the pixel value is higher than its neighbour, the algorithms stores a 1 value - 0 otherwise; this number is usually converted to a decimal value, for better interpretation;
- the above numbers determine a histogram for each window - optionally we can normalize;
- the characteristics vector contains all the above histograms.

For the classification stage the authors usually choose Support Vector Machines (SVMs).
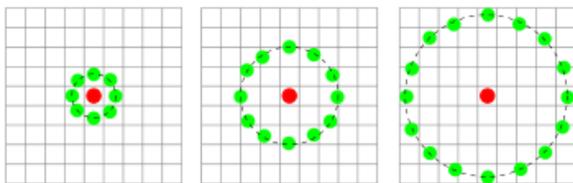
Figure 9: Local binary patterns on 9x9 pixels cells

5. SHAPE SPACE TECHNIQUES

The shape characteristics have some constraints:
- translation invariance;
- scaling invariance;
- rotation invariance;
- stability after small shape changes.

The techniques used in this area are split in two major categories:
- boundary based;
- region based.

According to (Vassilieva, 2012), there are multiple approaches, according to the image below:
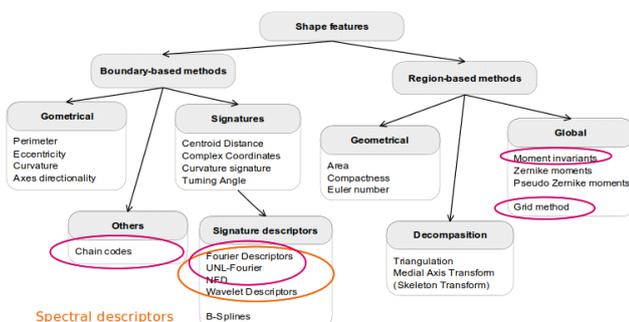
Figure 10: Shape features

The most recent approaches target techniques based on moment invariants and Zernike moments. However, as mentioned before, these techniques do not provide very good results, especially in the context of clutter and occlusion. This is why usually the authors choose to mix these descriptors

with others, like RGB histograms, Hu moment invariants and fuzzy techniques (Kashif Iqbal, 2012), (Chandan Singh, 2012).

A comparative study realised in 1997 established the below ranking:

Table 1: Shape descriptor techniques comparison

Method	T=5	T=10	T=15	T=20
Reduced chain code	55.1%	47.6%	50.0%	60.65
Fourier descriptors	72.2%	76.9%	75.9%	74.9%
UNL features	81.3%	79.9%	83.7%	89.3%
Moment invariants	84.7%	86.3%	86.8%	87.7%
Zernike moments	66.9%	66.5%	70.4%	78.2%
Pseudo Zernike moments	66.9%	66.5%	70.4%	78.2%
MI and FD	93.8%	87.3%	87.1%	89.6%
MI and UNL	93.3%	89.2%	89.3%	91.1%

6. SEGMENTATION TECHNIQUES

There are two major types of approaches:
- based on edge detection;
- based on region analysis.

The solutions can be found based on analysing other types of descriptors, extracted from colour, texture, movement etc.

The main problems tackled by segmentation are related to discontinuities:
- points;
- lines;
- edges.

Their detection is usually realised through applying masks.

One of the mostly used segmentation techniques is the one introduced by Noboyuki Otsu (Wikipedia, Otsu's method, 2012). This method separates the pixels in a certain area in two classes (foreground and background), so that the variance sum is minimal.

Another widely embraced approach is the *watershed* segmentation, which is based on establishing the margin between two areas by virtually increasing them until they intersect. The name comes from the resemblance between the phenomenon of two nearby pools flooding each other, thus the necessity of building a shed between them.
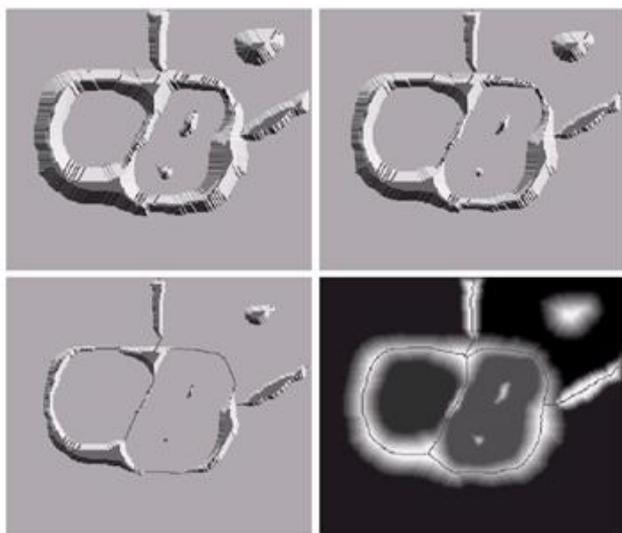
Figure 11: Watershed segmentation

Some researchers use radiating gradient vector flow (RGVF) snakes, in order to improve the results obtained by applying Sobel convolution masks, as those were affected by nearby similar areas.

7. LOCAL DESCRIPTORS TECHNIQUES

After going through the segmentation process we can obtain areas which can be computed in order to extract characteristics, like:

- edges;
- corners;
- regions of interest (RoIs);
- ridges.

The key points are those points that are not affected by translation, scaling, rotation and are minimally affected by minor changes. The main approaches in this area are:

- SIFT (scale invariant feature transform);
- SURF (speeded up robust feature);
- GLOH (gradient location and orientation histogram);
- HOG (histogram of oriented gradients);
- LESH (local energy based shape histogram).

7.1. SIFT

This approach is the first one in this area, proposed by David Lowe in 1999 (Scholarpedia, 2012). It is considered to be the parent of all the subsequent algorithms and it is especially important in the CBIR area. The approach is described in the table below (Wikipedia, Scale-invariant feature transform , 2012):

Table 2: SIFT

Problem	Technique	Advantage
Key localization / scale / rotation	DoG / scale-space pyramid / orientation assignment	accuracy, stability, scale & rotational invarianc
geometric	blurring /	affine

	resampling of local image orientation planes	invariance
distortion		
indexing and matching	nearest neighbour / Best Bin First search	Efficiency / speed
Cluster identification	Hough Transform voting	reliable pose models
Model verification / outlier detection	Linear least squares	better error tolerance with fewer matches
Hypothesis acceptance	Bayesian Probability analysis	reliability

The results can be observed in the image below:
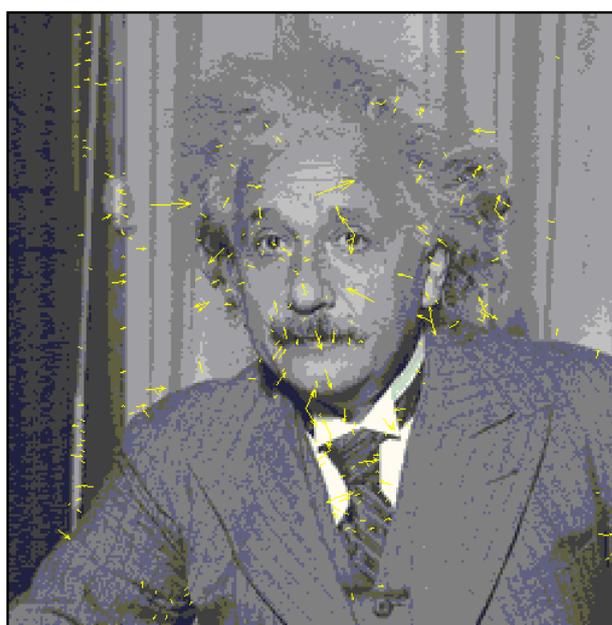
Figure 12: SIFT descriptors

7.2. SURF

This approach was first introduced by Herbert Bay in 2006. It is inspired by SIFT but according to the authors it is much faster and more resistant to noise than its parent (Wikipedia, SURF, 2012). As sub-techniques, it uses a sum of 2D Haar wavelets and aproximations through by using the integral image technique (Wikipedia, Summed area table , 2012).

It rapidly increased in popularity therefore the researchers can find many open source implementations, where they can vary the accuracy as opposed to the computing time (BoofCV, 2012).

7.3. GLOH

The GLOH approach is also a SIFT spin off. Its main characteristics are the precision consistency on different image sources and the dimensionality reduction through PCA (principal component analysis). Even though applying PCA

usually decreases the performance, a comparative study (Krystian Mikolajczyk, 2005) places this approach among the best. Even so, PCA increases the computing time by much, according to (Herbert Bay T. T., 2008).

7.4. HOG

This approach was first introduced by two INRIA (French National Institute for Research in Computer Science) researchers Navneet Dalal and Bill Triggs in 2005 (Wikipedia, Histogram of oriented gradients , 2012). Initially, the algorithm was used for pedestrial detection in static images, but since then it evolved a lot and it is used in different areas, like pedestrial detection in video streams, animal recognition or vehicle recognition.

The main idea behind this approach is that an object can be described through the gradient distribuiton or by the edge orientation. Combining multiple such histograms results in a descriptor. For a better accuracy, the results are normed according to the contrast on larger areas (named blocks) which makes the results to be more stable when confrunted with photometric changes. As the implementation uses cells, the results are stable when dealing with scaling as well.
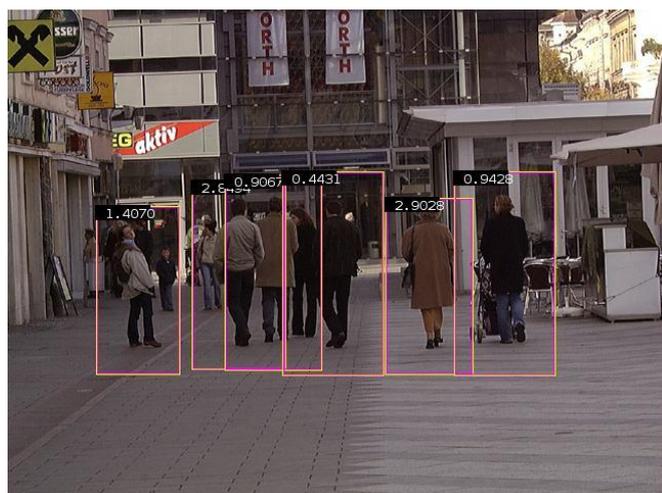
Figure 13: HOG classification

7.5. LESH

This technique is the most recent in this area. Initially it was used in the face recognition area but it can be used in solving other problems as well, like shape recognition, object recognition or posture recognition (Wikipedia, LESH, 2010). As opposed to GLOH, LESH limits the dimensionality to 128, but uses more types of histogram, which leads to increasing complexity.

The histograms used as a basis for LESH have first been introduced in 1997 by Robins and Owen, but those had bad results when identifying the object in the image and when dealing with noise. Later on, this approach was extended by Kovesi in 2000, by adding some weighing mechanisms, which lead to minimizing the noise and increasing the robustness when varying the light intensity (Saquib Sarfraz, 2008).

Unfortunately, since this is a very recent approach, there are no comparative studies to rank its results.

7.6. CONCLUSIONS

One of the most complete studies conducted in this area have been realised by Krystian Mikolajczyk and Cordelia Schmid. These researchers compared and classified ten techniques targeting the extraction of local features, according to different classification methods, on different image sets, compressed or not, affected or not by noise. The results are described in the table below (to be noted that SURF and LESH are missing):

Table 3: Local descriptors comparative study

Descriptor	recall	1-precision	#nearest neighbor correct matches
GLOH	0.25	0.52	192
SIFT	0.24	0.56	177
Shape context	0.22	0.59	166
PCA-SIFT	0.19	0.65	139
Moments	0.18	0.67	133
Cross correlation	0.15	0.72	113
Steerable filters	0.12	0.78	90
Spin images	0.09	0.84	64
Differential invariants	0.07	0.87	54
Complex filters	0.06	0.89	44

Another comparative study in the SURF area conducted by Herbert Bay, Tinne Tuytelaars and Luc Van Gool, resulted in two reports published in 2006 and 2008. The first report (Herbert Bay A. E., 2006) includes three main approaches - SURF (with two accuracy values), SIFT and GLOH (with two accuracy values). The results place the SURF implementations on the first position, followed by SIFT and GLOH, even if the authors used the same test data as Mikolajczyk si Schmid.

The second report (Herbert Bay T. T., 2008) uses the same test data, but introduces a new SURF approach (U-SURF). Some new test scenarios have been added to the existing ones (angle changes, chromatic differences, scaling and light intensity variations). Under these conditions, in the context of recognizing the objects stored in an art museum, the collected results have been ranked as it follows:
1. SURF-128 - 85.7%;
2. U-SURF - 83.8%;
3. SURF - 82.6%;
4. GLOH - 78.3%;
5. SIFT - 78.1%;
6. PCA-SIFT - 72.3%.

8. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

The CBIR problem is very complex at all its levels and the implementations vary according to the targeted area. As an

example, in a field related to robotics, the response time may be more important than the precision. This is why the researchers use RGB histograms and try to approximate the trigonometric functions through others, of a lower complexity. As opposite, in the medical field, the precision is very important, therefore the histograms are computed in the CIE colour space. In this field, the images are stored in a specific format, where the header includes information related to the capturing angle, rotation degree, translation etc. Therefore, the CBIR modules that were traditionally processing these characteristics have been considered to be redundant and eliminated.

There are very few implementations which consider the semantic level when querying the CBIR module. Most of the current systems use an image and not plain text as input. Of course, at this level the problem becomes much more complicated, as the author needs to consider other issues as well, like the geographical position, ethnical details, trust factor etc. and weighing all these inputs. As a result, the research seems to be switching to other directions as well (like developing a more descriptive UI), in order to facilitate the user to better describe his/her query. As an example, when confronted with a graphical interface, an user may find it difficult to depict a cat.

In the classification stage, the authors choose to use artificial intelligence modules (like neural networks, SVMs or fuzzy logic). There are other approaches which use AI in other stages as well, like swarm technologies, genetic algorithms or just plain unsupervised machine learning.

In the **colour space,** most of the authors choose to use histograms. Very few researchers use colour moments in order to extract colour features.

In the **texture space**, the authors seem to embrace the Gabor filters. However, the most recent approaches use more and more LBP. We need to mention that all the implementations are considering only grey gradients, which are very sensitive to conversions.

In the **shape space**, the preferred approach is *moment invariants.* The only other alternatives are based on Zernike moments or Fourier descriptors. As mentioned before, currently there are no solutions to clutter and occlusion, which is leading the researchers towards combining more algorithms in order to get better results.

In what regards **segmentation**, the situation is very similar to the one in the shape space. This is why the community recently introduced the notion of *weak* and *strong* segmentation, in order to better describe the results. Weak segmentation describes a result which is surely part of the targeted object, but does not guarantee to include all the object. There are studies which show that strong segmentation can be reached; however all the tests have only been conducted in laboratories. In its simplest form, at this level the researchers use the point segmentation as a basis for their algorithms to extract local descriptors.

As a common element for all the **CBIR** implementations we can notice the attempt to reduce the computing complexity, through different methods:

- histograms are created in the RGB space, for low precision systems;
- the presence of fuzzy modules in the decision module;
- swarm logic when dealing with distributed resources;
- the *bag-of-words* approaches, which are trying to reduce the dimensionality, just as hashing does;
- approximating trigonometric functions;
- weighing the local characteristics according to their importance;
- splitting the search area in smaller cells (the pyramid approach).

ACKNOWLEDGEMENTS

REFERENCES

BoofCV Performance: SURF [Online] // BoofCV. - 2012. - http://boofcv.org/index.php?title=Performance:SURF.

Chandan Singh Pooja Local and global features based image retrieval system using orthogonal radial moments [Journal] // Optics and Lasers in Engineering. - 2012.

Coelho Luis Pedro Mahotas: Image Processing in Python [Online] // www.luispedro.org. - 2010. - 2012. - http://luispedro.org/software/mahotas.

G. D. Magoulas S. A. Karkanis, D. A. Karras and M. N. Vrahatis Comparison Study of Textural Descriptors for Training Neural Network Classifiers [Conference]. - [s.l.] : Proceedings of the 3rd IEEE-IMACS World Multi-conference on Circuits, Systems, Communications and Computers, Athens, Greece, 1999.

H. Borgne A. Guerin-Dugue, A. Antoniadis Representation of images for classification with independent features [Journal] // Pattern Recognition Letters. - 2004.

Hall-Beyer Mryka The GLCM Tutorial Home Page [Online] // GLCM Texture Tutorial. - 2007. - http://www.fp.ucalgary.ca/mhallbey/tutorial.htm.

Herbert Bay Andreas Ess, Tinne Tuytelaars, Luc Van Gool Speeded-Up Robust Features (SURF) [Report]. - 2006.

Herbert Bay Tinne Tuytelaars, Luc Van Gool Speeded-Up Robust Features (SURF) [Report]. - Zurich, Leuven, Belgia : [s.n.], 2008.

Hui Yu Mingjing Li, Hong-Jiang Zhang, Jufu Feng Color Texture Moments for Content-Based Image Retrieval [Book]. - [s.l.] : Proc. IEEE Intl Conf. on Image Processing, 2002. - pp. 929-932.

Johansson Bjorn QBIC (Query By Image Content) [Online] // QBIC (Query By Image Content). - 2002. - http://www.isy.liu.se/cvl/Projects/VISIT-bjojo/survey/surveyonCBIR/node26.html.

Kashif Iqbal Michael O. Odetayo, Anne James Content-based image retrieval approach for biometric security using colour, texture and shape features controlled by fuzzy

heuristics [Journal] // Journal of Computer and System Sciences. - 2012.

Koenderink Jan Human Visual Awareness - Space, Time, Shape // 3R's of Computer Vision: Recognition, Registration, Reconstruction. - Sicily : International Computer Vision Summer School, 2012.

Krystian Mikolajczyk Cordelia Schmid A Performance Evaluation of Local Descriptors [Article] // IEEE transactions on pattern analysis and machine intelligence. - 2005.

Luis Caro Javier Correa, Pablo Espinace, Daniel Langdon, Daniel Maturana, Ruben Mitnik, Sebastian Montabone, Stefan Pszczółkowski, Anita Araneda, Domingo Mery, Miguel Torres, Alvaro Soto Indoor Mobile Robotics at Grima, PUC [Journal] // Journal of Intelligent & Robotic Systems. - 2012.

Mauricio Correa Gabriel Hermosilla, Rodrigo Verschae, Javier Ruiz-del-Solar Human Detection and Identification by Robots Using Thermal and Visual Information in Domestic Environments [Journal] // Springer Science+Business Media. - 2011.

Mehmet Sezgin Bulent Sankur Survey over image thresholding techniques and quantitative performance evaluation [Journal] // Journal of Electronic Imaging. - January 2004. - pp. 146–165.

Onur Küçüktunç Ugur Güdükbay, Özgür Ulusoy Fuzzy color histogram-based video segmentation [Journal] // Computer Vision and Image Understanding . - 2010.

Saquib Sarfraz Olaf Hellwich Head Pose Estimation in Face Recognition across Pose Scenarios [Conference] // Proceedings of VISAPP 2008, Int. conference on Computer Vision Theory and Applications. - Madeira, Portugal : [s.n.], 2008. - pp. 235-242.

Sarfraz Saquib Head Pose Estimation in Face Recognition across Pose Scenarios [Conference] // Proceedings of VISAPP 2008, Int. conference on Computer Vision Theory and Applications. - Madeira, Portugal : [s.n.], 2008. - pp. 235-242.

Scholarpedia Scale Invariant Feature Transform [Online] // Scholarpedia. - 2012. - http://www.scholarpedia.org/article/Scale_Invariant_Feature_Transform.

Sebe N. Feature extraction & content description - DELOS - MUSCLE Summer School on Multimedia digital libraries, Machine learning and cross-modal technologies for access and retrieval. [Online] // www.videolectures.net. - 25 02 2007. - 2012. - www.videolectures.net/dmss06_sebe_fecd/.

Sebe Nicu Feature extraction & content description [Online] // www.videolectures.net. - 25 02 2007. - 2012. - http://videolectures.net/dmss06_sebe_fecd/.

Sebe Nicu Feature extraction & content description - DELOS - MUSCLE Summer School on Multimedia digital libraries, Machine learning and cross-modal technologies for access and retrieval [Online] // www.videolectures.net. - 25 02 2007. - 2012. - http://videolectures.net/dmss06_sebe_fecd/.

Siddique Sharmin A Wavelet Based Technique for Analysis and Classification of Texture Images [Book]. - 2002.

Techasith Pravi Image Search Engine [Report]. - Imperial College, London, UK : [s.n.], 2002.

Th. Gevers A.W.M. Smeulders Image Search Engines - An Overview [Book Section] // Emerging Topics in Computer Vision / book auth. G. Medioni S. B. Kang. - [s.l.] : Prentice Hall, 2004.

Thomas Breuer Geovanny R. Giorgana Macedo, Ronny Hartanto, Nico Hochgeschwender, Dirk Holz, Frederik Hegger, Zha Jin, Christian Müller, Jan Paulus, Michael Reckhaus, José Antonio Álvarez Ruiz, Paul G. Plöger, Gerhard K. Kraetzschmar Johnny: An Autonomous Service Robot for Domestic Environments [Journal] // Springer Science+Business Media B.V.. - 2012.

Vassilieva Natalia RuSSIR - Russian Summer School in Information Retrieval [Online]. - 2012. - http://videolectures.net/russir08_vassilieva_cbir/.

Wang Shengjiu A Robust CBIR Approach Using Local Color Histograms [Report]. - Department of Computer Science, University of Alberta, Edmonton, Alberta, Canada : [s.n.], 2001.

Wikipedia Cumulative histogram [Online]. - 2012. - http://en.wikipedia.org/wiki/Histogram#Cumulative_histogram.

Wikipedia Gabor Filter [Online] // Wikipedia. - 2012. - http://en.wikipedia.org/wiki/Gabor_filter.

Wikipedia Histogram of oriented gradients [Online] // Wikipedia. - 2013. - 2013. - http://en.wikipedia.org/wiki/Histogram_of_oriented_gradients.

Wikipedia Histogram of oriented gradients [Online] // Wikipedia. - 2012. - http://en.wikipedia.org/wiki/Histogram_of_oriented_gradients.

Wikipedia Independent component analysis [Online] // Wikipedia. - 2012. - http://en.wikipedia.org/wiki/Independent_component_analysis.

Wikipedia LESH [Online] // Wikipedia. - 2010. - http://en.wikipedia.org/wiki/LESH.

Wikipedia Local binary patterns [Online] // www.wikipedia.org. - 08 11 2011. - 2012. - http://en.wikipedia.org/wiki/Local_binary_patterns.

Wikipedia Otsu's method [Online] // Wikipedia. - 2012. - http://en.wikipedia.org/wiki/Otsu%27s_Method.

Wikipedia Scale invariant feature transform [Online] // Wikipedia. - 2013. - 2013. - http://en.wikipedia.org/wiki/Scale-invariant_feature_transform.

Wikipedia Scale-invariant feature transform [Online] // Wikipedia. - 2012. - http://en.wikipedia.org/wiki/Scale-invariant_feature_transform#David_Lowe.27s_method.

Wikipedia Summed area table [Online] // Wikipedia. - 2012. - http://en.wikipedia.org/wiki/Integral_image.

Wikipedia SURF [Online] // Wikipedia. - 2012. - http://en.wikipedia.org/wiki/SURF.

Wikipedia Wavelets [Online] // Wikipedia. - 2012. - http://en.wikipedia.org/wiki/Wavelet.

X. Cao W. Shen, L.G. Yu, Y.L. Wang, J.Y. Yang, Z.W. Zhang Illumination invariant extraction for face recognition using neighboring wavelet coefficients [Journal] // Pattern Recognition. - 2012.

Young-Beom Lee Unsang Park, Anil K. Jain, Seong-Whan Lee Pill-ID: Matching and retrieval of drug pill images [Journal] // Pattern Recognition Letters. - 2012.