Harmonic methods for solid noise filtering from the image periphery [(*)]

Dimo Dimov and Nadezhda Zlateva

***Abstract**: This paper presents a method for removing solid/coarse "artifact" noise from the periphery of trademark images. The idea is to find a closed curve that separates optimally the area comprised of trademark elements from that, containing the artifact-noise. The proposed approach uses one-dimensional harmonic filters to find a series of "segmentation points" which approximate the curve. The experiments, carried out with real images, show 50% improvement.*

***Key words**: Image processing, Convolution, Fast Fourier Transform, Gaussian, Laplacian, Artifact-noise.*

1. INTRODUCTION

A great number of fundamental tasks in image recognition are related to identifying the key elements in an image and filtering out the insignificant ones, which may also be interpreted as noise. Noise isolation and/or suppression is also a standard problem for systems that manage image databases (IDB), where the so-called CBIR (Content Based Image Retrieval) approaches are applied to access data. Given an input image, the CBIR access methods need to extract the key content, like color, texture, shape, spectrum and use it to organize a search for similar images in the IDB. The CBIR method needs to provide a certain level of noise tolerance, at least in respect to regular noise which is typical for the process of image acquisition.

An example of a CBIR system is EFIRS (Effective and Fast Image Retrieval System [1, 2]), which is developed by IIT-BAS and operates with trademark images. It has been established that for certain images the level of accuracy of the EFIRS decreases. These images contain noise, referred to as "artifact-noise," which is mostly located in the periphery of the image. In particular, the artifact-noise overlaps the object (mark) to be recognized to such an extent that the usual approaches for noise isolation fail.

A method for isolating artifact-noise in the periphery of trademark images is proposed herein. The heart of the method is in converting the filtration problem from two-dimensional (2D) image space into a series of one-dimensional (1D) filtrations, each of which along a polar direction of the centralized image object hereinafter referred to as *object*.

The use case we consider is from the practice of the Patent Office of Republic of Bulgaria (PORB), with their large IDBs of trademark images. Nevertheless, the proposed method is also applicable for a wide class of images whose significant CBIR information is located mostly around the image frame center. The reasoning behind is that photographers and artists like positioning the key, emphasis objects in the center of the lens or the picture.

2. BACKGROUND

2.1. Area of application and restricting conditions

The analysis of an excerpt of images that EFIRS operates on, established that the artifact-noise therein is often introduced by operator negligence. For example: the trademark is copied from a newspaper or an envelope, whose content (texts and/or graphics) surround the trademark of interest; the trademark is scanned from a stained paper; the trademark background differs in color from the accidentally caught pad, etc.

To improve the targeted EFIRS resistance against noise of this kind, we had to choose an approach that conforms to the following assumptions (i.e. a priori information):

[(*)] This work was supported by following grants of the Institute of Information Technologies (IIT) at Bulgarian Academy of Sciences (BAS): Grant # I-1306/2003 of the National Science Fund at Bulgarian Ministry of Education & Science, Grant # RC6/2004 of the Bulgarian Agency for ITC Development, and Grant #010056/2003 of BAS.

- the object content is located mainly in the center of the image;
- the solid/coarse artifact-noise (if any) is scattered through the image periphery;
- the object components (of a sophisticated object) overlay the background that is considered dominant, of constant intensity and of frame proximity;
- the areas of the object components and those of the artifact noise do not overlap;
- the area of the artifacts is considerably smaller than that of the object components;
- the (averaged) intensity of the artifacts may be close to that of the object components.

2.2. Idea of the proposed approach

A *segmenting curve* is being sought, i.e. a simple, closed, and maximally smooth curve, that would separate the area of the centrally-located object elements from that of the artifact-noise at the periphery. In a certain sense, we may even look at the segmenting curve as at a contour [6]. Yet, instead of solving a 2D problem, we decompose it to a set of 1D problems by looking for a series of *segmenting points*, which approximate the segmenting curve. Fig.1 summarizes the process of filtering artifact-noise from an input image.

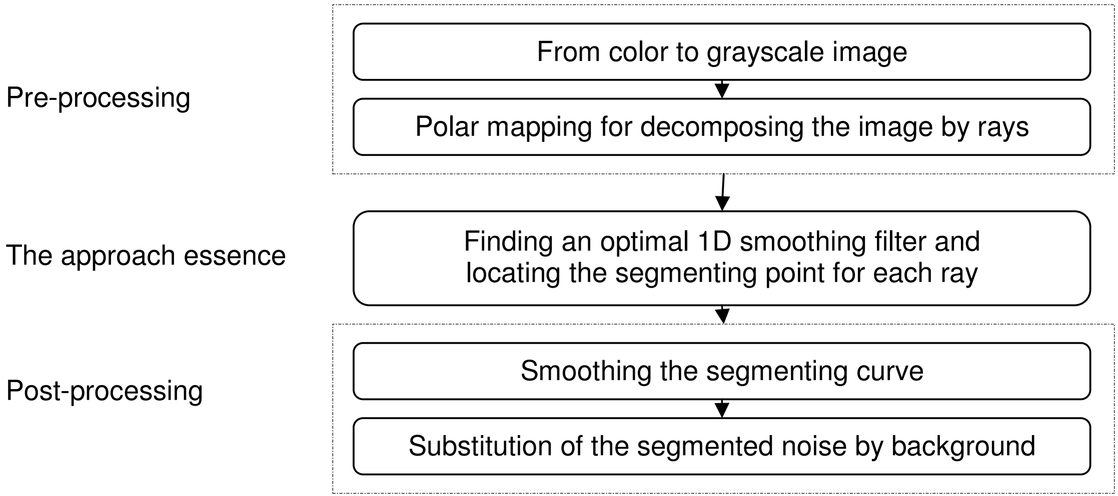


Fig.1. Stages of the image processing for artifact-noise isolation

3. PROBLEM ANALYSIS

For exposition accuracy, we will be working with grayscale images already converted into "positive", i.e. whose average intensity exceeds half of permissible range (see § 4.1).

The chosen approach and tools for solving the problem are based on the a priori information (see §2.1) and result from the analysis that follows.

We expect that the explored filter will separate both entities - the object components and the noise-artifacts, in such a way that the boundary between their regions will appear like some closed curve. This stems from the assumption of a centralized trademark content against peripheral artifact-noise. The segmenting curve will be positioned over a background, and as of the above assumptions, it should be unambiguously defined towards some center in the area of the trademark object (see Fig.2a).

On the other hand, the assumptions of central orientation and compactness of the object components allow us to interpret the trademark image as inscribed in some "circle," its center being "the center of gravity" of the image. Thus, instead of looking for a unique 2D filter that will optimally respond to the sporadic location of object components within the image, we can search for an optimal (and also 2D) filter for each sector of the image within the circle. We should interpret the image as decomposed into sectors at intervals of $\Delta\theta$, considered from some basic ray $\theta=0$, whose origin is the center of gravity of the image.

Thus, we could optimize such a filter for each sector over a series of filters of regularly increasing "length" (see Fig.2a).

The disadvantage of this approach is that the other dimension, the "width" of the optimal filter, needs to vary proportionally to the current sector's width in the filtering position, i.e. the optimization will be parameterized by the current position in the sector of filtering (see Fig.2a). To reduce this complication, we can average the filter width and consider it as fixed, while the corresponding sector – as approximated by a band of the same width. Thus, instead of using 2D filters over the sector, we can optimize by using 2D filters along each axis (*ray*) of the sector (band approximated). If we choose a small enough $\Delta\theta$, then we can reach an approximating band of 1 pixel width, and thus, instead of 2D - we can only look for 1D filters along the corresponding radial ray, from the chosen center (see Fig.2a).

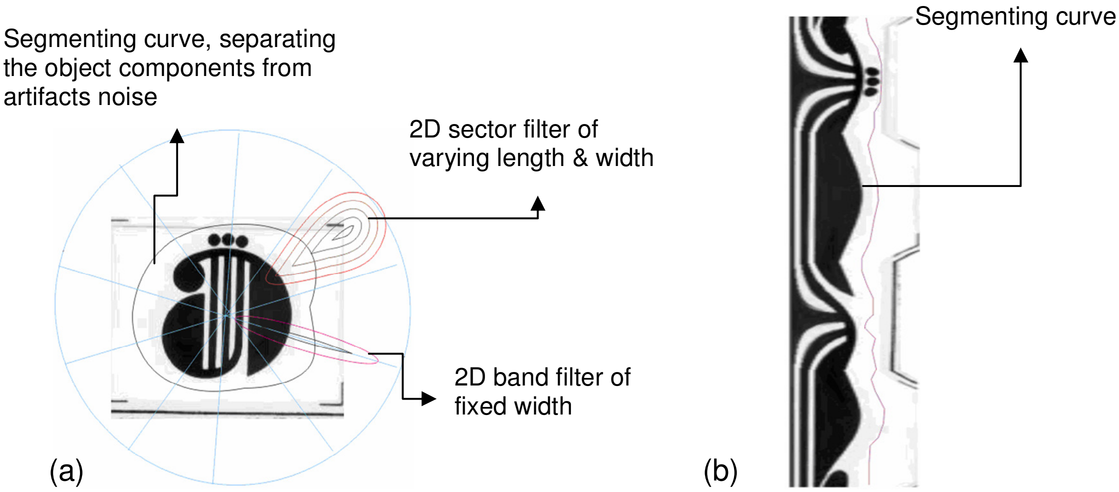


Fig.2. An example trademark image: (a) with appropriately added enclosing circle; and (b) the image, in polar mapping towards the center of gravity

3.1. Polar Mapping Application

The analysis above will come naturally, if besides looking at the original (input) image we also consider its transformation via the so called *Polar Mapping* (PM). Formally, the PM is defined as conformal (i.e. angles preserving) transformation [5]:

$$P : Z \to Z,\ \ P(x,y) = (\rho,\theta),\quad (x,y) \in Z,\quad (\rho,\theta) \in Z \qquad (1)$$
$$\rho = \sqrt{x^2 + y^2},\ \ \theta = arctg(y/x),$$

In practice, PM may be described as 'snipping a circular image along a given ray to its center, and stretching it, without bending its radial rays, to form a rectangle'. Compare with Fig.2b, where the PM center is chosen at the center of gravity of the image.

For programming convenience we position the distances *ρ* (from the PM center to the periphery) horizontally, while the angles *θ* (from ray to ray) – vertically (see Fig.2b). Our smoothing filter for the 1D signal, defined on each horizontal line of the PM, is sought at an optimal length, such that would only give us a single point, separating the object- and artifact areas along the ray. We expect that the aggregation of all *segmenting points*, found in the above fashion, will approximate the aimed 2D *segmenting curve*. The latter follows from our intuitive belief that the correlation of the intensity along the rays is much smaller than that along the polar circles.

In an idealized case, the aimed segmenting curve of a given image may turn out to be a circle with a center – the PM centre. Its shape in the PM will be a straight line. In the common case, however, the segmenting curve will be identified as some closed and unambiguously defined curve (towards the PM centre), while its transform in the PM space will be an almost linear poly-line (a vertical coast line), i.e. a function $\rho = \rho(\theta)$, $\theta \in [0,2\pi)$, see also Fig.2b.

Actually, the polar mapping of an image may also be considered periodic in (-∞, +∞), which is very handy when applying Fourier transforms (FT) by θ, i.e. vertically at Fig.2b. However, we will need a FT by ρ, and therefore will have to expand the PM-image to a mirror-symmetrical one along the axis ρ=0 (not illustrated).

3.2. Definition of the 1D filters

Let's look at the image across a given ray (respectively *a horizontal* in PM). If we initiate its successive smoothing with some averaging filter, with progressively increasing effective length *m, m*=1,2…, then we can expect the following tendency: at some minimal, yet big enough $m = m_{opt}$ the result will be described by 2 maxima. The left maximum corresponds to the object (a slice of it), and the right one – to the artifact-noise (see Fig.3).

The segmenting point can be chosen somewhere in the minimum (that is, in the background area) between these two maxima. Hence, the criterion for defining the one-dimensional filter is in finding the optimal parameter m_{opt}.

The parameter m_{opt} exists, since we intuitively expect that the function for the number of maxima *N*:

$$N = N(m),\ m > 0, \tag{2}$$

will be monotonically decreasing (see Fig.3b). Therefore, for programming efficiency, m_{opt} may be estimated through a binary search in the signal (the intensity) definition domain $s=s(\rho)$, $\rho \in [0, \rho_{max}]$, θ=cte.

Actually, the statement that $N = N(m),\ m > 0$ is absolutely monotonous, is not entirely correct. However, a tolerance ΔN for *N* might be estimated for a chosen filter type, in the frames of which monotonicity might be considered. The corresponding theoretical explanation, though, exceeds the goals of this paper.

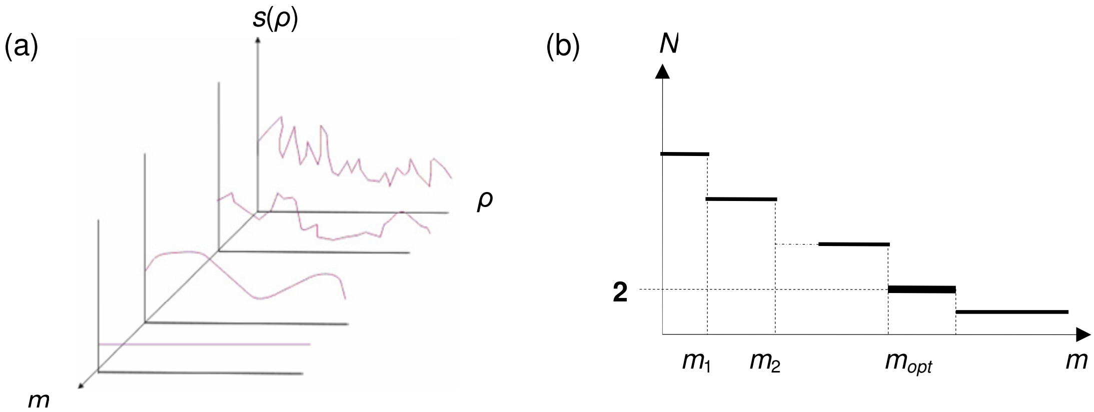


Fig.3. Illustration of the monotonicity of *N* (number of maxima) on the smoothing level *m* (filter width): (a) the signal $s = s(\rho)$ after smoothing with increasing *m*, and (b) an optimal value m_{opt} exists for *m, m≥1*.

For definiteness, let's consider smoothing with the simplest filter *M*(*x,m*) of the "moving average" type, of *length m*, *m*>0:

$$M(x;m) = \begin{cases} 1, & x \in (-m/2, +m/2) \\ 0, & x \notin (-m/2, +m/2) \end{cases} \tag{3}$$

Six possible situations result from this smoothing, see also Fig.4.

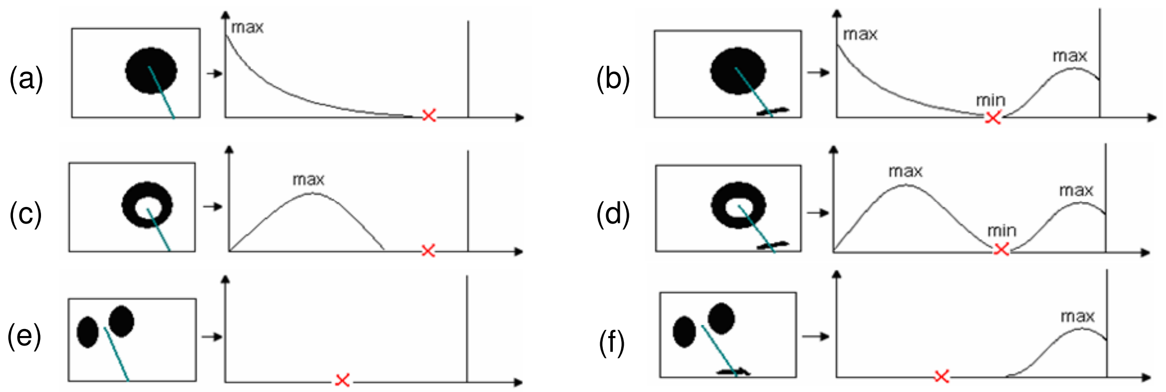


Fig.4. Illustration of the six possible "ideal" object-artifact configurations expected after a large enough smoothing along the shown ray:

- on the left (a, c and e) are the ideal configurations containing a smoothed object predominantly (only), where no more than 1 maximum is encountered, and
- on the right (b, d and f) – similar configurations but with an artifact-noise left after smoothing, and where 1 extra maximum could be evaluated respectively.

Let's first examine situations (b) and (d) from Fig.4. If we assume that the input 1D signal is continuous, then its minimum could be determined using its first derivative. In the discrete case, a 1D differential filter of mask (-1 0 1) could be used for simplicity, while for higher precision a larger mask could be implemented. In any case, two consecutive 1D convolutions over the signal will be carried out: the first one for smoothing the signal by *M*(*x,m*), and the second one - for differentiating the result of the smoothing.

Using the associativity property of convolution, we can optimize the above operations by differentiating the *M(x,m)* filter first, and convolving it with the input signal afterwards:

$$D(f * s) = Df * s\,, \quad s = s(\rho),\;\; f = M(\rho, m) \tag{4}$$

If we go with differentiation to locate the minimum (between the two maxima) of the smoothed signal, then the choice of *M(x,m)* is perhaps not very appropriate. The (first) derivative obtained by this choice will give too big a variation in the minimum location – in the area between the two consecutive zero crossings. For this reason, we will change the smoothing filter *M(x,m)* with the 1D Gauss filter (Gaussian):

$$G(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}}\exp(-(x-\mu)^2/2\sigma^2)\,, \tag{5}$$

where μ is the filter center, corresponding to the current filtering position $\rho \in (0, \rho_{max})$, and σ defines the *effective length m* of the filter*,* $m \approx 6\sigma$.

The Gaussian allows the derivative to be estimated with much higher precision (in the zone of the zero crossing). Considered discretely as a series of weight coefficients for averaging (smoothing), the Gaussian filter stresses on the neighbours located closer to the center μ. This contributes to the evenness (and continuity) of the next estimation of the derivatives.

The already described approach for determining m_{opt} will be used for the remaining four configurations of mutual disposition of objects and artifacts (see (a), (c), (e) and (f) from Fig.4). Yet, it will be enriched with additional heuristics that are not emphasized here.

In general, the procedure for determining the optimal length m_{opt} and, through it, the segmenting point for each ray of the image (a horizontal in the PM), starts with $m = \rho_{max}$. If for a current length *m,* the processed signal may be juxtaposed to any of the configurations in Fig.4, then we keep looking for the minimal *m*, for which the juxtaposition is still possible. Should a length with no juxtaposition to any of the configurations be reached, we break the search and return the last *m,* for which a match was possible.

This is from where the approach obtained its name – *adaptive segmenting filter.* We also call it *harmonic*, since it can only be expressed through convolutions, i.e. it can be represented in the Fourier space, and quickly calculated via FFT (Fast Fourier Transform).

3.3. Comparison with the 2D image filtering

Using differential filters in image recognition is most often related to problems concerning image contour detection [6]. Contours are usually defined as curves of the maximum intensity change for given standard vicinity. Thus, the most natural approach to calculate the contours is via the gradient of the image. The gradient is a vector, i.e. it has 2 components (magnitude and direction), and usually only its magnitude is used for contour representation. Yet, since the magnitude depends nonlinearly from the gradient position, it cannot be represented by a series of convolutions, i.e. by a linear filter. In cases when the latter is needed (e.g. to ensure the associative property in successive filtrations), the Laplace filter is used. The Laplace 2D filter corresponds to the 2nd derivative filters in 1D. Of course, an additional operation of "zero crossing" will be necessary, to finally determine the contours [6].

If we search for an alternative approach to solve the problem of defining the "segmenting curve, which separates object components from artifact-noise," and if we also ignore the already shared preferences for radial treatment of the image, then we can seek the segmenting curve as a contour in 2D. Hence, we will preliminary smooth the "masses" of the objects and the artifacts, by filtering with a low-pass 2D filter. The next step will be to differentiate twice, by using a Laplace filter, in order to ensure the above mentioned associativity, and to also be ultimately able to directly apply the LoG (Laplacian of Gaussian) filter [5]:

$$LoG * f(x,y) = (\nabla^2 * G) * f(x,y) = \nabla^2 * (G * f)(x,y)\,, \tag{6}$$

where $\nabla^2 = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right)$ is the Laplace operator, *G* is the 2D analog of the Gaussian from (5), and $f(x,y)$ is the original image (in a Cartesian coordinate system).

Of course, to the extent that the segmenting curve has to represent a unique and closed contour, we will also need to define optimal dimensions of the filter (6), but now in 2D. These optimal sizes will most probably turn out to be dependent on the filter position in the image and in the already described *centrally radial aspect***.**

In short, when comparing the above two approaches, we can claim the following:

• In respect to effectiveness in processing speed, we can consider both approaches similar. Both need to use the FFT, since the search for the "optimal sizes" of the filters presumes values of the definition domain that are big enough. Therefore, the direct convolution calculations (in the space domain), may turn out to be unacceptably long.

• In respect to effectiveness in functionality, ignoring the central-radial aspect, which is in the heart of the method we propose, will lead to an increase in the number of errors, when dealing with the artifact-noise suppression.

4. AUXILIARY EXTERNAL PROCESSING

4.1. Necessity of image preprocessing

As we already mentioned in the method analysis part above, we are considering grayscale images. Therefore, we convert input color (RGB) images to gray scale ones, by using a standard algorithm [6]. This approach reduces the processing complexity of the method, and manages to maintain the essential and sufficient characteristics, like contrast and differentiation among the respective components (the object and the noise ones).

To avoid the "positive-negative" duality we use a method for evaluating the grayscale image in analogy with the Otsu method for optimal global binarization [6]. Namely, if the average image intensity is above the Otsu threshold, then we define the image as *positive* (i.e. dark objects and/or artifacts on a light background), and alternatively we define it as *negative* (i.e. light objects and/or artifacts on a dark background). If we are dealing with a negative, the image is transformed into a positive using the formula in (255 – *i*) for every pixel, where *i* is its intensity.

4.2. Final image postprocessing

The last stage of the proposed method for eliminating noise-artifacts consists in the final restoration of the image, i.e. removing the segmented noise. It is accomplished by applying the color and intensity of the background to the area between the segmenting curve and the image frame. To do that, we first transform the segmenting curve already PM established, back to the original image (see Fig.2b). The curve should be preliminary smoothed out in PM by a simple 1D low-pass filter to avoid possible erroneous outliers.

The shade of the fill-in color can be defined by averaging the color/intensity over: the area to be filled in; the segmenting curve only; by the Otsu method [6] for the entire image.

5. EXPERIMENTAL RESULTS

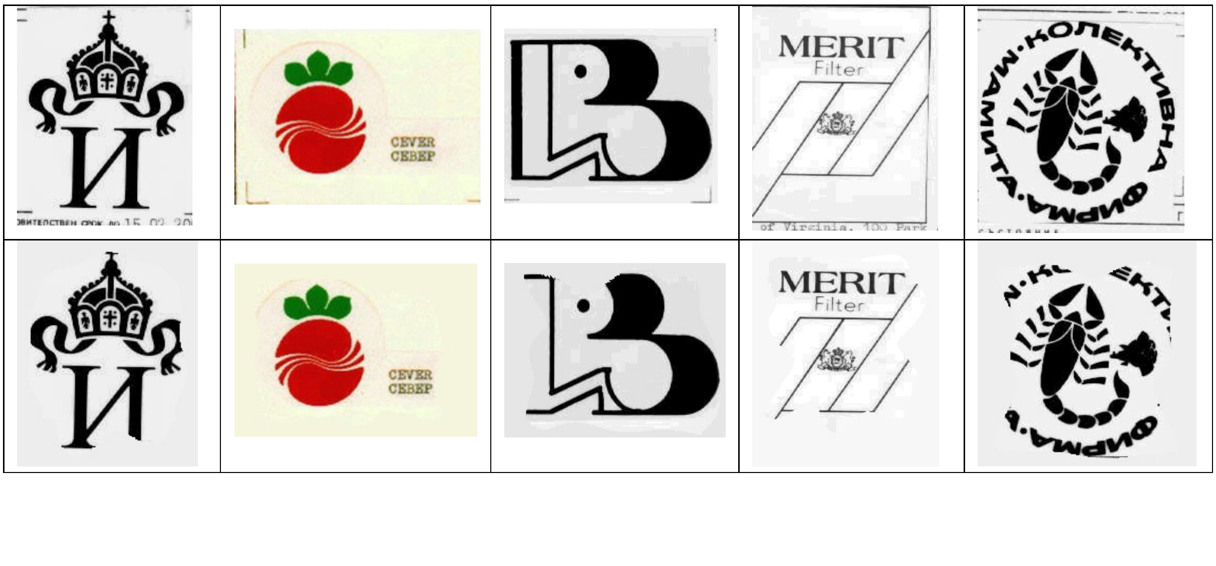
Test experiments were undertaken to evaluate the proposed *harmonic filter* contribution to noise-tolerance of a CBIR system in operation over real images. The system was EFIRS. Three of its CBIR strategies [1, 2] were tested with the filter. The test IDB was organized on 146 representative trademark images from the practice of PORB.

Experiments consist in analyzing the two primary cases: (i) Filter + EFIRS and (ii) EFIRS without preliminary filtering. This approach allows for primary evaluation of the system improvement by the proposed filter application. Besides, the test results are independent on the current status of EFIRS development.

We will skip the experiment details, since the idea and the organization are similar to those of a competing approach (the *geometrically-morphologic* filter) to resolve the task in question, which is also submitted to this conference (paper 3B.15). Instead, we will concentrate on the herein achieved results.

Analyzing comparatively the number and type of errors resulting from both tests, (i) and (ii), we find out an improvement of about 2 times, i.e. the percentage of error reduction when applying the harmonic fitter is about 50%. Moreover, these results are similar or a bit better than those of the geometrically-morphologic filter under the same test premises.

Table 1. Some test images (first row) and their noise-filtered version (second row)



In Table 1 we compare a few input images with their final, noise-removed versions. Images with different noise levels were selected as well as images whose character and noise-artifacts do not abide by the pre-conditions already listed in 2.1.

Having in mind the results we can conclude that the *harmonic adaptive-segmenting filter* is significantly drastic in terms of noise-artifacts and risks removing object components located closer to the periphery of the image. In such case, we rely on the fact that the most essential information of the image will be preserved, and that it will be sufficient to find the particular image via the chosen CBIR strategy (of EFIRS).

6. CONCLUSION

We propose a method for noise-artifact filtering in the image periphery, which we refer to as *harmonic adaptive-segmenting filter*. It is based on the first derivative of a 1D Gauss smoothing filter.

The method transforms the aimed task into an equivalent one of locating a segmenting point in a 1D signal, which divides the signal optimally in two areas – one for the object, and the other for the noise-artifacts. Any horizontal line of the image polar transformation (whose center, in the particular case, is the center of gravity of the image) is considered 1D signal. For smoothness in the image decomposition, the discrete intervals $\Delta\theta$ and $\Delta\rho$ are picked to be sufficiently small. Hence, the series of segmenting points, which have been determined by the set of decomposing rays, forms the so-called segmenting curve, which separates the key object in the image from the possible noise-artifacts in the periphery.

When the method was applied to EFIRS (an experimental system for fast and reliable CBIR), the resulting error level of images' retrieval was reduced to 50%, and even more. This corresponds to our original goal.

The following directions are considered for future work on the subject matter:

- Elaboration of alternative, fuzzy theory based approach for artifact-noise segmentation that better spares the possible object components near the image periphery.
- Elaboration of the alternative approach through the 2D LoG filter for optimal estimation of the segmenting curve (contour).
- Combining the method with other approaches, like geometrically-morphologic ones, and/or statistical ones.

REFERENCES

[1] Dimov D.T., Experimental Image Retrieval System, *Cybernetics and Information Technologies*, BAS, Sofia, **2**, 2 (2002) 120-122.

[2] Dimov D.T., Harmonic and Wavelet Transform Combination for Fast Content Based Image Retrieval, Int. Conf. on PDE Methods in App. Math. & Image Proc., Sunny Beach, Bulgaria, Sep. 7-10, 2004 (http://www.math.bas.bg/~kounchev/2004_conference/lectures/ddimov.pdf)

[3] Marr D., Vision – A Computational Investigation into the Human Representation and Processing of Visual Information, Henry Holt & Company, 1982.

[4] Otsu N., A Threshold Selection Method from Gray-Level Histograms, *IEEE Trans. on Systems, Man, and Cybernetics,* **9**, 1 (1979) 62-66.

[5] Reiss, T. H., Recognizing Planar Objects Using Invariant Image Features. Lecture Notes in Computer Science, Vol. 676, Springer, Berlin, Heidelberg, New York, 1993.

[6] Sonka M., Hlavac V., Boyle R., Image Processing, Analysis, and Machine Vision, 2^{nd} ed., Brooks/Cole & ITP, Pacific Grove, 1999.

ABOUT THE AUTHORS

Dimo Dimov, Assoc. Prof., Ph.D., Institute of Information Technologies, Bulgarian Academy of Sciences, Tel: (+359 2) 8706493, E-mail: dtdim@iinf.bas.bg

Nadezhda Zlateva, Ph.D. Student, Tel: (+359 2) 771830, E-mail: nadia.zlateva@aviq.com