

Image segmentation based upon topological operators: real-time implementation case study

Ramzi Mahmoudi, Mohamed Akil

► **To cite this version:**

Ramzi Mahmoudi, Mohamed Akil. Image segmentation based upon topological operators: real-time implementation case study. Real-Time Image and Video Processing 2009, Jan 2009, San Jose, Californie, United States. 10.1117/12.812121 . hal-01294104

HAL Id: hal-01294104

<https://hal-upec-upem.archives-ouvertes.fr/hal-01294104>

Submitted on 30 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Image Segmentation Based Upon Topological Operators Real Time Implementation Case Study

R. Mahmoudi*, M. Akil*

IGM, Unité Mixte CNRS-UMLV-ESIEE UMR8049, Univ. Paris-Est ;
Cité Descartes, BP99, 93162 Noisy le Grand, France
corresponding author : akilm@esiee.fr

ABSTRACT

In miscellaneous applications of image treatment, thinning and crest restoring present a lot of interests. Recommended algorithms for these procedures are those able to act directly over greyscales images while preserving topology. But their strong consummation in term of time remains the major disadvantage in their choice. In this paper we present an efficient hardware implementation on RISC processor of two powerful algorithms of thinning and crest restoring developed by our team. Proposed implementation enhances execution time. A chain of segmentation applied to medical imaging will serve as a concrete example to illustrate the improvements brought thanks to the optimization techniques in both algorithm and architectural levels. The particular use of the SSE instruction set relative to the X86_32 processors (PIV 3.06 GHz) will allow a best performance for real time processing: a cadency of 33 images (512*512) per second is assured.

Keywords: Crest Restoring, Thinning, Segmentation, Hardware Implementation, Optimization Techniques, Parallelism, Pentium, Profiling, SSE instructions set.

1. INTRODUCTION

In order to avoid important significant information loss, due to the passage to binary image, numerous operators acting directly over greyscale image were proposed [2] [6] [10]. These kinds of operators that preserve the topology of the image are extremely useful in solving various problems such as medical image analysis, 3D model matching and character recognition. Different algorithms dealing with thinning and skeleton were studied [1] [4] [7] (in [7] fifteen 2D thinning algorithms were compared). To grade these algorithms, the majority of publications focus on the following criteria as quality criteria [8] of classification i) Homotopy: the skeleton must preserve the topology. ii) One-pixel-thickness: one-pixel-thick objects. iii) Noise immunity: Damage on skeleton caused by noise must be repaired. Other criteria relative to the implementation and the execution time can be introduced [3] [5] [9] to classify algorithms as sequential (each non-significant information is detected and removed immediately), or parallel (all non-significant information are firstly detected then removed). Among all the proposed algorithms, we are interested in both operators of thinning and crest restoring introduced by Bertrand and Couprie in [4]. Proposed operators [4] were inserted in a powerful chain, see figure 1, which offers high segmentation quality. But the major drawback of this chain is the processing time with is very considerable. It is also important to munch that no context for specific use has been defined. We propose to improve this chain not only for better legibility but also for an execution time which respects real-time image processing constraint.

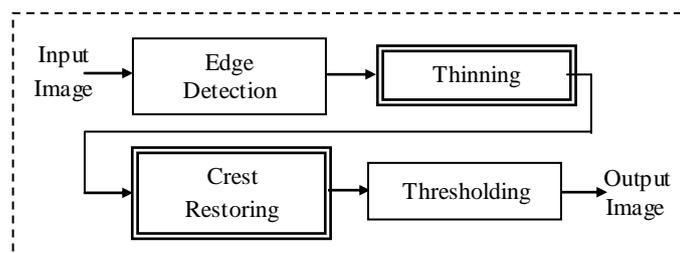


Fig. 1. Original segmentation chain

*{mahmoudr, akilm}@esiee.fr

Medical image processing will be the new defined application context. The original chain is based on edge detection approach. It starts by extracting edges from the input image thanks to Deriche filter, then it calls upon the thinning operator followed by the crest restoring one and finally a thresholding process is lunched. In order to improve legibility and also execution time, important ameliorations aiming not only algorithm's level but also architecture's level were applied: exploiting Pentium performance and specially the associated SEE instructions set allow us to assure real time segmentation.

The remainder of the paper is organized as follows. Section II gives a quick review of theory bases of chooses algorithms and explains the different steps of the original chain. Details about enhancement, implementation and applied optimization are given in section III. We show also obtained results . Finally, the conclusions recall our research work.

2. THEORY BASES & ORIGINAL CHAIN

In this section, we start by recalling some basic notion introduced by Bertrand and Couprie [4] to modify the topology of grayscale image in order to filter out the non-significant information. Then we develop different steps of the original chain.

Authors [4] introduce the notion of λ -destructible point which allowed them to selectively alter the topology, based on a local contrast parameter (λ). Inspired by the same principle, authors [4] introduce the notion of λ -end and λ -deletable points in order to apply the same procedure on grayscale skeleton. To eliminate non significant branches and region minima, λ -end and λ -deletable points will be lowered what allows to get filtered skeleton. But the major disadvantage of this procedure, as any others procedures based on the dynamic, is that if two neighboring minima are separated by a high crest which contains a narrow and deep pass, the dynamic between these two minima only depends on the altitude of the pass. To cure that a crest restoring method was introduced for detecting and eliminating narrow passes, obviously on grayscale image.

As we munched in the introduction the original chain is based on edge detection approach. The first step is edges extraction thanks to Deriche filter. Following this detection, λ -destructible points will be selected under the control of a local contrast parameter λ for a selective modification of topology. We lower values of the selected up pixels until the point that they are not λ -destructible any more. This operation will be repeated until stability and guarantees a strict preservation of the topology. Even obtained objects at this stage are thin and lengthened; they frequently suffer from cuts under the effect of remaining noise. The goal is to reconnect them again. If we regard an image as a topographic relief, a threadlike object can be described like a crest line and the cracks will correspond to the narrow passes in this line. To cure this problem, authors [4] propose an algorithm allowing detecting and raising these narrow passes. They introduce the notion of extensible separating point (points having under two connexities) and according to its local topological characteristics, separating point of the thinned image are classified in constructible point, saddle point, end point or an isolated point. Only saddle points will be raised on the same level as those of the crest. Obtained thinned skeleton with restored crest is the input of the thresholding procedure. The output image is a binary image and presents the segmentation that we are looking for.

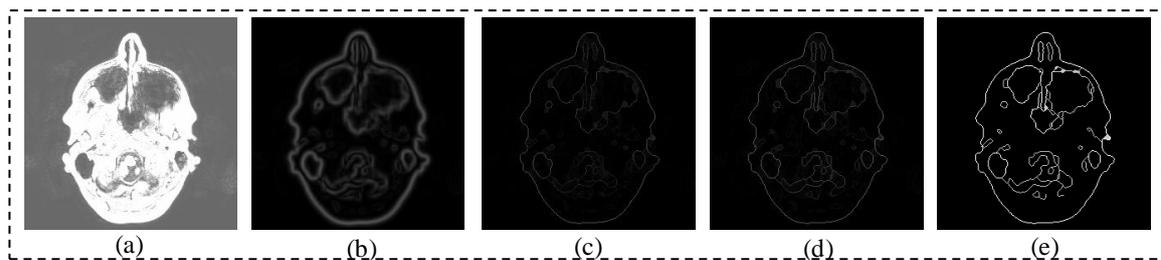


Fig. 2. Obtained results by the original chain

In figure 2, all steps of the original chain are illustrated. Gradient image (b) is obtained from (a) by the application of Garcia Lorca gradient operator. At first, crest has been reduced to thin lines (c) by a filtered thinning with $\lambda=5$, then restored (d) by executing the Crest Restoration algorithm until stability. To highlight object of interest a threshold has been applied (d) given (e). It is clear that no artifacts have been introduced by the executed chain. Compared to popular method known as hysteresis thresholding, the proposed method gives much better results. It is also important to recall that the origin chain is a general interest chain. Its use is not framed by a specific context even if we used 2D images of human brain cut for testing.

3. OPTIMIZATION & RESULTS

To obtain the best execution time, it is necessary to have an overall view of the intervening agents and to maintain in spirit reached objective: In this work, we deal with a segmentation process executed on Pentium microprocessor (P4 with 128 bits). We will start by giving a first evaluation of the original chain's execution time to be able to fix the first targeted functions for optimization. Then we explain the algorithmic improvements brought before finishing with the architectural ameliorations.

Within the framework of algorithmic improvements, a new segmentation's chain will be proposed. The new chain is inspired by the old one proposed by Bertrand and Couprie [4] and adapted for medical images processing. Also basing itself on the contour approach, it uses the two principal algorithms of thinning and crests restoration while including three new algorithms to ensure a dynamic redistribution phase, a double thresholding phase and a labeling phase. The new chain is a complete chain which offers a high segmentation quality and it is much faster than the original one. Concerning architectural side, applied optimization's techniques have allowed a spectacular profit in segmentation's times. Let take the case of an image of 768 Kbytes size (512*512), its segmentation's time passed from 3,6s to 0.03s that allows a possible cadency of 33 image/s.

3.1 First evaluation

In our analyses, we called upon Gprof as being the standard profiler of GNU world. After first evaluation, it proved that segmentation time of the original chain is extremely important. Indeed, only images having sizes lower than 48kbytes can be segmented in less than one second, the reason is the quasi reduced number of the treated pixels of these small sizes images. But since the moment that we process images having size higher then 750kbytes, the segmentation time exceeds one second. For example images (512*512) take 3.6s to be segmented. According to this profiling, we note that thinning phase consumes more than 50% of segmentation's total time. The iterative criterion (repetition until stability) of the two algorithms (thinning and crest restoring) can explain this percentage. We should not also forget that the pixels resultants from contours detection phase are numerous and cover the 255 gray's levels, this big number of level slows down the algorithm of thinning. The algorithm of crest restoring is slowed down by the extensive damage on the crest's lines. These damages are due to the presence of the noise in the acquired images. We will try to solve these problems and to remove obstacles which penalize the segmentation's time.

3.2 Algorithmic enhancement

The modifications associated to this level aim the thinning and crest restoring algorithms in first. The goals of these modifications can be classified as follows: the first priority to minimize the number of processed pixels by the thinning algorithm, then to reduce damage in crest lines by reducing noise area, to improve in term of quality the legibility of the final resulting image, and obviously to minimize throttling nodes in targeted algorithms.

3.2.1 New chain organization

The first factor which slows down the thinning phase is the number of treated pixels. The direct passage from the contour detection towards thinning penalizes enormously the segmentation time. For this reason, we propose to advance the thresholding phase before thinning and that will allow to reduce the number of treated pixels and consequently to accelerate thinning. Thanks to this reorganization, we obtain an average profit of about 72%. The segmentation's interval

of the whole tested images passed from [0:14]s to [0:2]s. The best profits were relating to big size images. Results of the new reorganized chain are promising and make it possible to look for real-times segmentation following other improvements. At this stage, only nine images (512*512) are segmented per second.

3.2.2 New dynamic redistribution phase

When we introduced crest restoring phase. We mentioned that cuts on thinned objects are due to the remaining noise in the image. As we explain it before if two neighboring minima are separated by a high crest which contains a narrow and deep pass, the dynamic between these two minima only depends on the altitude of the pass. Thus to reduce the number of processed pixels by the algorithm of crest restoring we propose to add a new dynamic redistribution phase at the beginning of the chain. To achieve this aim, we chose the dynamic's redistribution as pretreatment's method. This technique consists in transforming image's histogram so that this one occupies all dynamics. It is very important to mention that the topology will be preserved even the associated dynamic of each pixel will change. It is also important to mention that thanks to the thresholding process, enhanced noise by the dynamic's redistribution will be eliminated. The addition of this new dynamic's redistribution phase has improved the quality of objects resulting from the thinning phase. The objects are less damaged by noise. Crests lines are more continuous with fewer disruptions. Thus it will be fewer narrow passes to be detected and eliminated by crest restoring algorithm.

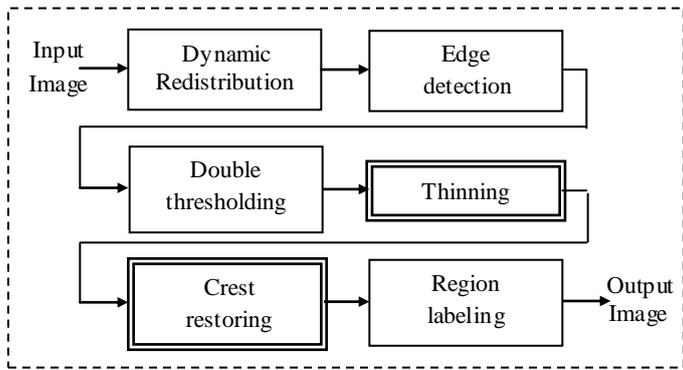


Fig. 3. Adding new labeling phase

3.2.3 New Labeling phase

The threshold, in the original chain, aimed to highlight the skeletal objects. When we advanced this phase before the thinning, we changed its algorithm and we have defined a new role (minimizing the number of processed pixels).

Thus its original purpose is no longer assured. For this we propose to add a new phase of labeling to address this problem. The new algorithm has to increase the readability of the final image to ensure a better quality of segmentation and it should not penalize the segmentation time that we seeks to improve since the beginning of this work. Once the restoration of crest is completed, we can associate to the scene the list of areas delimited by closed contours, see figure 3. These areas will be located by a label. This label is characterized a color. Operator of area labeling associates the same color for all pixels of the same area.

In figure 4, all steps of the new chain are illustrated. After the dynamic redistribution of the original image (a) we obtain (b). Gradient image (c) is obtained from (b) by the application of Garcia Lorca gradient operator then a double thresholding process is lunched to get (d). Crest line has firstly been reduced to thin lines (e) by a filtered thinning with $\lambda=5$, then restored (f) by executing the Crest Restoration algorithm until stability. To highlight object of interest a labelling algorithm has been applied to (f) given (g). The new chain provides great segmentation's quality. The various objects of interest are present in the final image. They are presented in more explicit way than ones offered by the original chain.

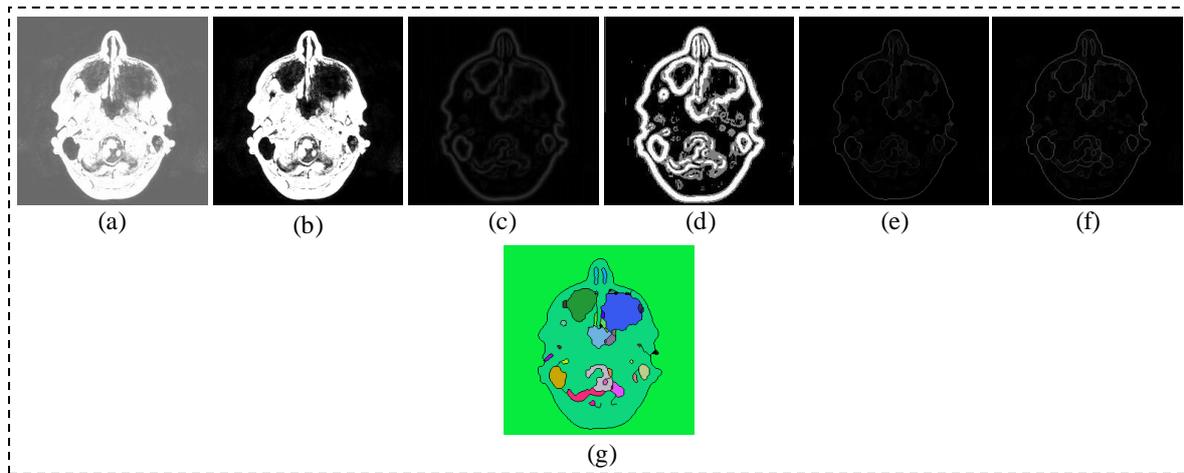


Fig. 4. Obtained results given by the new chain

3.3 Architectural level optimization

Algorithmic optimization can be necessary in very specific cases, but it remains always independent of target architecture. However on RISC machines which have a large registers and where the effectiveness requires the reunification of the identical instructions to profit from the effect pipeline, we can implement a code even more powerful than the one obtained after algorithmic optimization. This aim can be achieved by exploiting SIMD technology which rests on a revolutionary technique, namely to treat several data by loading only one instruction.

3.3.1 Compiler optimization options

We begin by activating the optimization options of the used compiler GCC. There is a very slight improvement only for large sizes images. Profit in segmentation's time is very limited, even with the activation of -O3 option.

3.3.2 Memory and register management

Some SSE2 instructions require that its arguments be aligned on 16-byte boundaries when moving to (/from) memory (otherwise the instruction throws exception). For that we have defines the following procedure allowing the redirection of `_aligned_malloc` (ms RCT) and the memorizing of allocated size. Others those variables of transitions, the memory alignment is very important at reservation's time of large tables: matrices of the pixels, table of the topological numbers.

```
void *benchmalloc (size_t size)
01: void *ptr = (void *) _aligned_malloc(4*size);
02: if (ptr) alloc_mem += size;
03: return ptr;
```

Even if we improve memory access time, it is certain that if operands are placed in the registers the execution time will be faster. But the number of available registers is limited, so it is almost impossible to eliminate all memory access for intermediates results. Let take the example of the associated values for each point in thinning algorithm. For each pixel, we need to compute its index (according to the eight neighbors) and we need also to compute associated alpha. To avoid double memory access, we coded both parameters in the same integer (by shift): index will be code on 24 bits and alpha on the 8 remaining bits. Thus for each point we will use only 32 bits what enables to charge values of four points at the same register (see figure 5).

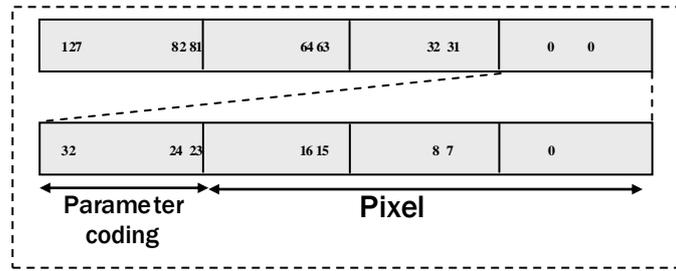


Fig. 5. Register management

3.3.3 Intra-Register vectorization

Pipeline promotion is one of the key points to improve the performance of programs. To exploit the potential parallelism offered by the Pentium 4, register vectorization can be applied by unrolling serial loops into a form that takes advantage of instructions provided by SSE2.

Anyone who looks carefully at the thinning algorithm realizes that the two phases of detection and removing of destructible points are independent. During the detection, the eight neighbors of each pixel are checked according to a precise chronological order. Thus, neighbors of any pixel (x) and its neighbor ($x+1$) can be examined at the same time without ambiguity. We also point out that each pixel indicated as a destructible point is coded with its alpha in the same integer (32 bits) then piled up. To benefit from the 128 available bits, we can launch up to four instructions of stacking in parallel. We draw attention to the fact that the vectorization is effective only in the case of large size loops. If the number of iterations is small, no contribution will be visible. It is also very important to check variables sharing within the loop. Even in the case of a conditional dependency, the probability to increase latency, instead of decreasing it, is great.

3.3.4 External loop unrolling

The whole of applied operations to characterize a point's nature uses piles. Let us take the case of the crest restoration algorithm, which uses in a first stage a pile to store the pre-divergent points with their alpha. The order in which points are stored has no importance thus we can launch this loop several times in parallel. For efficient loops unrolling, extra branches must be eliminated from the loops, unrolling also enhances other optimizations, such as prefetch, scalar replacement, and instruction scheduling. Finding the unroll factor is important to determine the amount of loop expansion. In the present case, the management of the pile is the only obstacle. If we launch more than two loops at the same time, it will be difficult to manage the pile's sharing resulting in an automatic increase in latency. Thus, we are limited to a double external loop unrolling for the better management of the pile.

4. CONCLUSION

In this work, we proposed a new image processing chain, which offers high quality segmentation respecting real time constraints. The new chain is perfectly suited to medical applications. Acting directly over grayscale images while preserving their topologies, all topological operators and in particular thinning and crest restoration operators have given birth to a powerful image segmentation of 2D brain's image cuts.

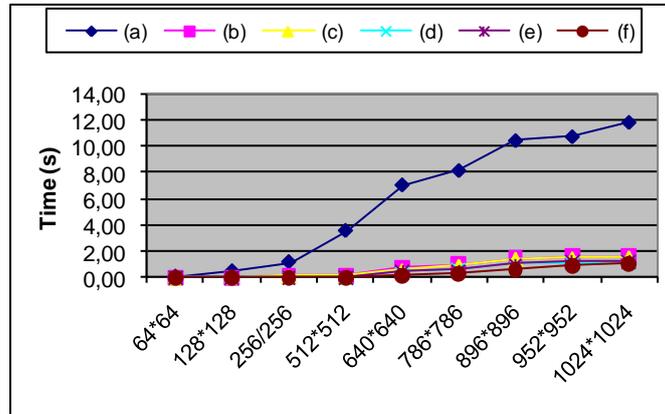


Fig. 6. Summary of performance – (a) the original chain’s performance, (b) after reorganizing, (c) after dynamic redistribution, (d) after algorithmic optimization, (e) after compiler optimization and (f) after architectural optimization.

The association of algorithmic optimization techniques to architectural optimization techniques relative to Pentium processor (P4 3.06Ghz), allows to expect a cadency of 33 images (512 * 512) per second, see figure 7. These real-time performance confirms the interest of an “Algorithm Architecture Adequation” approach, see figure 6. For better adaptation between the chain and the target architecture, we begin by carving algorithms to ensure better stability and readability of code. Then, we make calls to techniques that exploit the best performance of the processor.

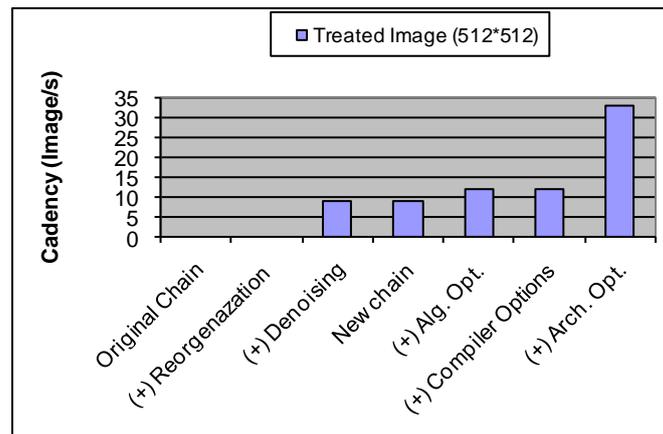


Fig. 7. Summary of performance in terms of cadency

REFERENCES

- [1] G. Bertrand and M. Couprie, "Two-Dimensional Parallel Thinning Algorithms Based on Critical Kernels", Journal of Mathematical Imaging and Vision, Volume 31 , Issue 1 (May 2008), Pages: 35 - 56
- [2] L. Huang, G. Wan and C. Liu, "An Improved Parallel Thinning Algorithm", Seventh International Conference on Document Analysis and Recognition (ICDAR'03) , 2003, icdar, Volume 2, p. 780
- [3] M. Couprie, F. N. Bezerra et G. Bertrand, Topological operators for greyscale image processing. Journal of Electronic Imaging, October 2001, Volume 10, Issue 4, pp. 1003-1015
- [4] M. Couprie, F. N. Bezerra and G. Bertrand. Grayscale image processing using topological operators. SPIE Vision Geometry VIII., 1999, volume 3811, pages 261-272.

- [5] C. Lohou and G. Bertrand , "New parallel thinning algorithms for 2D grayscale images", Vision geometry Conference No9, San Diego CA , ETATS-UNIS (30/07/2000) 2000, vol. 4117, pp. 58-69.
- [6] T. Ojala, M. Pietikäinen and T. Mäenpää, "A Generalized Local Binary Pattern Operator for Multi-resolution Grayscale and Rotation Invariant Texture Classification", advances in pattern recognition No2, Rio de Janeiro , BRESIL (11/03/2001) 2001, vol. 2013, pp. 397-406.
- [7] M. Couprie, "Note on fifteen 2D parallel thinning algorithms", University of Marne-la-Vallée, IGM2006-01, 2006.
- [8] J. Canny, "A computational approach to edge detection", IEEE PAMI, 1986, Volume 8 , Issue 6, pp. 679 - 698 .
- [9] F. Garcia Lorca, L. Kessal, D. Demigny "Efficient ASIC and FPGA implementation of IIR filters for real time edge detections", Proc. International Conference on Image Processing, IEEE ICIP 1997.
- [10] G. Bertrand, J.C. Everat and M. Couprie, "Image segmentation through operators based upon topology", J. Electron. Imaging 6 (4), 1997, pp. 395-405.