One long argument: Azriel Rosenfeld and the genesis of modern image systems

Virginio Cantoni a, Stefano Levialdi b,*

a Dipartimento di Informatica e Sistemistica, Università degli Studi, Pavia, Italy
b Dipartimento di Informatica, Università Roma, La Sapienza, Roma, Italy

Received 9 July 2004; received in revised form 6 October 2004

Abstract

We present our vision on the history of multiprocessor architectures for image processing seen under the light of Darwin’s Theory of Evolution, remembering our fruitful and entertaining scientific meetings with Professor Azriel Rosenfeld along the nearly 30 years of our acquaintance. We experienced the blooming of novel systems (70s and 80s), the harvesting of such systems (90s) and later, the extinction of the massively parallel systems in favor of the adoption of large clusters of fast commercial processors to accomplish the same image processing tasks.

© 2004 Published by Elsevier B.V.

Keywords: Parallel image processing; Computer vision architectures; Multiprocessor systems evolution; Natural selection

In memoriam

One of us has known Professor Azriel Rosenfeld for many years since 1968 at the Tirrenia School in Italy of Pattern and Character Recognition, where he lectured on basic algorithms for image processing. Since then, we have met at the different International Conferences and Workshops that were held, most of them under his influence, in different parts of the world. He has been, for the whole Italian community, a scientific guide not only to the literature but also, and most of all, to the exciting subjects which were at the cutting edge of image processing research. His critical evaluation of our research lines together with a definite sense of humor, have helped us to design new architectures and algorithms, along the many years of our acquaintance. Among the many system projects that originated at Maryland University under his inspiration, Azriel also conceived and published a novel massive architecture (PRISM) alas, so massive, that remained a paper machine. At the Rome International Conference on Pattern Recognition, ICPR held on 1988, we

* Corresponding author.
E-mail addresses: virginio.cantoni@unipv.it (V. Cantoni), levialdi@di.uniroma1.it (S. Levialdi).
were happy to clap our hands when he received the first King Sun Fu Award. We believe that many generations to come will still be indebted to him for his open-minded view of how our field will be evolving.

1. Introduction

Professor Azriel Rosenfeld provided the state of the art for digital picture processing, year after year, starting in 1969: it rapidly became a standard reference for all researchers and Ph.D. students as well as for professionals in the field. His first survey paper (Rosenfeld, 1969) appeared on the ACM Computing Surveys, 1969. Every year he scanned and analyzed over 1500 papers from all the world conferences, workshops and scientific journals. Moreover, in 1971, Rosenfeld together with H. Freeman, founded the Computer Graphics and Image Processing journal that rapidly became the leading publication in this field, nowadays with the title Computer Vision and Image Understanding.

Among the many research lines that he fostered, we like to remember the first one related to his seminal work on digital geometry. In fact, during the 60s, images were perceived as structured sets of elements (as in a matrix) without considering the consequences of digitalization. In fact, adjacencies problems, tessellation and anisotropy properties on the digital plane, metric definitions and, generally, geometrical and topological properties of digital lines and curves were studied. AR pioneered research on these topics together with J. Pfaltz in a very well known paper (Pfaltz and Rosenfeld, 1966), providing formal definitions, including handling paradoxes, which pointed out the practical consequences of the achieved formalisms.

To allow efficient image processing during the 70s, it was soon realized that many identical local computations were required, so that it seemed natural to think about replicating the instruction executors, one for each picture element (pixel) or for each instruction (SIMD and MISD respectively). This led to the first class of massively parallel computer architectures, according to a well-known taxonomy (Flynn, 1972) introduced by Flynn, the so-called SIMD architectures (e.g. MPP (Butcher, 1981), CLIP4 (Duff, 1978)) and the MISD class (e.g. Cytocomputer (Lougheed et al., 1980), ZMOB (Rieger et al., 1981)). At those times (beginning of the 80s) programming was still a hard task for multiprocessor architectures even when using commercial microprocessors like the Z80. In this connection, ZMOB was programmed in Forth by Don Hopkins ¹ while he considered it as “the Computer of the Future, using the Processor of the Past”.

During those years there was a vast community of researchers, both academic and industrial, that committed themselves to the design and implementation of innovative parallel architectures that could be efficiently used for image processing. Following these lines it was hoped that real-time processing could be achieved. In one specific paper on computer architectures for image processing (Rosenfeld, 1988): “...it is well understood how to use most of these architectures to speed up processing at the array level. Much less is known, unfortunately, about how to use them to achieve speed-ups at the level of geometric representations or graph structures.

It is commonly believed that speed-up of array-level processing is all that is needed since that level involves the largest amount of data (images, transformed images, etc.), while the remaining levels involve only small numbers of entities (image parts, etc.). This belief may be dangerously misleading. Processing of image parts, for example, may lead to combinatorial explosions of the derived parts and of relations among collections of parts. It should not be assumed that the computational bottleneck in vision is entirely, or even primarily, at the array level. We need to learn how to use the proposed architectures at the higher levels, or failing this, to design new architectures appropriate for these levels”.

The real problem to be solved involved different issues: economical, technological, structural (pipeline/array/reconfigurable) and usable (Cantoni and Levialdi, 1983). Moreover, the technology was still

not ripe to allow real-time processing as required in many applications (e.g., surveillance and robot vision), so that artificial intelligence strategies were considered. One of the most successful was the planning strategy: processing images at low resolution, with a subset of data, and subsequently refining the solution at the required level of detail. AR described and evaluated the potentialities of multi-resolution in a book edited by him (Rosenfeld, 1984). This approach encouraged us to construct a family of pyramidal architectures (Cantoni and Levialdi, 1986; Cantoni and Ferretti, 1994), the PAPIA computers (Cantoni et al., 1985; Cantoni and Levialdi, 1988).

An interesting interconnection scheme chosen for PAPIA I is the Quad Pyramid shown in Fig. 1. The base and the first four levels are drawn illustrating the intralayer 4-connection as well as the interlayer quad connection so communicating with four PEs on the lower level. (Only the inner links corresponding to the subregion are shown.) Each layer has one quarter of Pes of the previous layer so that the total number of processing elements amounts to $4/3$ of the number of base PEs. Note that 9 intralayer connectivity has also been implemented with the bin-connectivity for the interlayer.

Following this approach, many groups around the world started designing and the construction of a variety of hierarchical architectures that were discussed and compared at a NATO Advanced Research Workshop (Cantoni and Levialdi, 1986) where AR participated praising the organization for its completeness, since all the important groups of the world involved in special projects attended this event. More specifically, AR presented, at this Workshop, a paper describing a collection of multiresolution, or “pyramid” techniques, for rapidly extracting global structures (features, regions, patterns) from an image. If implemented on these architectures, such techniques would require processing times on the order of the logarithm of the image diameter: $O(\log n)$.

AR was not only a driving force to model and describe both the power and limitations of those parallel computers, but also encouraged the construction of one specific architecture: the PRISM machine (Rosenfeld, 1985) as mentioned above.

The PRISM machine is shown in Fig. 2: a stack of $n$ cellular arrays each of size $2^n \times 2^n$ constitute the prism. Cell $(i,j)$ on level $k$ is connected to cells $(i,j)$, $(i+2^k,j)$ and $(i,j+2^k)$ on level $k+1$, with $1 \leq k < n$, where the sons are modulo $2^n$. Globally, it consists of $n4^n$ PEs, while the quad-pyramid only has $4^n + 1/3$ PEs. Fig. 2 presents the base (4-connected elements) and other three layers having 8-connected elements. The usual 4-adjacency at distance one within the same level and other four elements at distances power of 2 on the levels above.

The PRISM machine can perform various basic image processing operations, (e.g. Gaussian convolution or least square polynomial fits) on image neighborhoods of power of 2 sizes, and also the discrete Fourier Transform on the whole image in $O(n)$ time.

In practice, $n$ (number of levels) would be at most equal to 10, so that a PRISM machine has at most 10 times the number of computing elements with respect to the number of base Pes. Nevertheless a pyramid machine has 4/3 processors as the number of base PEs. We are then considering computer architectures that have a large number of processors with complex interconnection schemes where it is difficult to find a trade-off between computational power and cost for such a niche market (with respect to the general-purpose computer market), even using VLSI technology.

In spite of this, AR claimed that he was not interested in the practical outcome, but essentially looked at the computational properties of these innovative architectures (people really interested in building such machines could go along!). Note

![Fig. 1. The Quad Pyramid interconnection scheme with 4-connectivity on the plane.](image-url)
that, a number of multiprocessor computers on a single plane, were really built following the interconnection models he suggested (i.e. the hypercube topology; a good example of hypercube architecture is the Connection Machine. Such computer, the CM-2, \(^2\) was a massive, 5 feet tall cube formed in turn of smaller cubes, representing the 12-dimensional hypercube structure of the network that connected the processors together).

AR concludes his analysis of parallel architectures for image processing as follows (Rosenfeld, 1988): ‘‘...various forms of parallel processing can be used to speed-up some of the simpler computations that are commonly performed by vision systems; but much less is known about how to speed-up higher level computations many of which involve computational searches. To achieve real-time vision in complex environments we need to design computer architectures and algorithms that can make effective use of parallelism at all levels. As already pointed out humans can recognize objects—even complex objects whose presence was unexpected—in a fraction of a second, which is enough time for only a few hundred (!) “cycles” of the neural “hardware” in the human visual system. Computer vision systems have a long way to go before they will be able to match this performance’’.

In fact, massively parallel processors for image processing are not now designed and built any more for a number of reasons. In the next section we will account for these reasibs considering the evolution of such machines, market and media applying the evolutionist theory.

## 2. Evolution in computing architectures for image processing: a Darwinist metaphor

Starting from the ’60s, for over 40 years, computer scientists have suggested, designed, built and some times even marketed, new computer architectures for image processing. We may consider the history of the development of such solutions as an architecture evolution process driven by image processing applications and obviously the real market. Along the time many projects have failed, others have been modified, adapting to new application demands, while others have had a wide popularity for a limited lapse of time but no parallel architecture specimen has lasted over 5 years.

Among those systems which we may place in the hall of fame, we quote the Cytocomputer (Lougheed et al., 1980) following the pipeline approach, the PICAP machine (Kruse, 1973), following the multi-special processing unit approach. As for the massively parallel approach we recall the MPP (Butcher, 1980); for the pyramid (Burt et al., 1986).

The development of new computer architectures may be seen as a natural evolution script under the

---

\(^2\) [http://www.svisions.com/ss/cm-dv.html]
light of Darwin’s evolutionist framework. Evolution, as described by Darwin, consists in a gradual adaptation of living beings to the environment and to its specific features (food availability, different predators, weather catastrophes, etc.); those unable to cope with such events, are doomed to disappear. Quoting from Darwin’s Origin of the Species (Darwin, 1859): “I have called this principle, by which each slight variation, if useful, is preserved, by the term Natural Selection”.

2.1. Darwin’s metaphor

One may consider Darwin’s view of evolution as a set of theories that formed the basis of his thinking. Not all such theories have found immediate acceptance within the biological community, for a variety of reasons, since some hypotheses have not found experimental evidence. These critical theories do not fully cover Darwin’s thoughts but the formalization of such theories by Mayr (1991) may be considered a sound reference platform for all scientists working on the evolution theory.

In the sequel, we will discuss the philogenesis of computer architectures for image processing along similar lines to those elicited in Mayr’s framework. For this reason we have titled this paper following his reference article which is driving our metaphor. An interesting study of the various points of view on Darwinism, which may be taken as preliminary statements, can be found in (Hanes). Five sub-theories are listed next to point out the different components of the evolution theory.

2.1.1. Evolution as such

This is the theory that the world is not constant or recently created nor perpetually cycling, but rather it is steadily changing, and that specimens (living organisms/computer architectures) are transformed in time. As far as components are concerned, they evolve following the prediction of Moore’s Law (1965). This law states that the processing power of a microchip doubles every 18 months. Corollary: computers become faster and the price of a given value of computing power halves every 18 months. At the same time, there is a Law by Gilder (1993), stating that the total bandwidth of communication systems triples every 12 months. New developments seem to confirm that bandwidth availability will continue to expand at a rate that supports Gilder’s Law. This hardware evolution is, in fact, both a constant adaptation technology and demand-driven process along time, and may be considered as a steadily changing evolution.

2.1.2. Common descent

This is the theory that every group of organisms descended from a common ancestor, and that all groups of organisms, including animals, plants, and microorganisms, ultimately go back to a single origin of life on earth. In fact, the computational model due to von Neumann (1945) is the basis and reference for all future developments of computer architectures; even for multiprocessors with a different computing paradigm which are known as non-Von. The first non-Von Neumann computer was developed by NEC in 1984 (Creative Computing, 1984) on a microprocessor chip, an image pipelined processor uPD7281D, for the purpose of image processing; another system was developed during the 80s at Columbia University, had NON-VON as its name (Shaw, 1982).

2.1.3. Multiplication of species

This theory accounts for the origin of the enormous specimen diversity. It postulates that species multiply, either by splitting into daughter species or by “budding”, that is, by the establishment of geographically isolated founder populations that evolve into new species. Similarly, new different parallel computing systems were born, having an enormous processor variety (with different functions), a wide spread of interconnection schemes in 2, 3 and $n$ dimensions of the hypercube topology, reconfigurability and different memory sharing.

In the 80s, different authors tried to classify the different projects with the aim of measuring their computational power and specific features. Among these taxonomies, let us quote Maresca et al. (1988) for the more general case; whilst for the hierarchical architectures Cantoni (1986). Lastly, in order to evaluate the computational capabilities considering a number of architectural features
(number of processing elements, organization of such elements, number of bits per pixel, parallel access to neighbors) Danielsson and Levialdi (1981). Furthermore, the application domain (image processing) shapes the most suitable system solution in analogy to the geographical influence on the species.

2.1.4. Gradualism

According to this theory, evolutionary change takes place through the gradual modification of populations and not by sudden (saltational) production of new specimens that represent a new type. Starting from compact computing elements (integrated circuits, VLSI, optical, etc.) we are now progressing towards distributed approaches that are large-sized clusters, (e.g. Grid processing).  

The transition between one (compact, tightly coupled) extreme and the other (distributed, loosely coupled) has been a gradual one. The reason for this graduality is in the technological advances which have allowed better and faster communication between processors, which is twice as fast as the computing capability. Note that the improvement in performance of sequential computers permits the management and processing of new media (sound, voice, images, videos), gradually increasing their multimedia efficiency.

2.1.5. Natural selection

This theory states that evolutionary change comes about through the abundant production of genetic variation in every generation. The relatively few individuals, who survive, owing to a particularly well-adapted combination of inheritable characters, give rise to the next generation. On the artificial side, the amount of digital images that can be handled and edited every day by non-technical people is measured in gigabytes, such activity could not be foreseen only 10 years ago. Considering digital photography, the enormous popularity of image dedicated software has provided the users with sophisticated tools that can be naturally executed on standard computers. This large and pervasive market is the one deciding success or failure of a product, a program or a machine: this explains why only few products widely advertised, are able to survive.

2.2. Five theories: Discussion

We will compare biological evolution, following Darwin–Meyr’s thoughts with the evolution of computer vision systems, and we can see that the analogies hold for the various sub theories. Moreover it is worthwhile to consider also the remarks arisen in both fields (biological and computational) since these may turn useful to understand the past and predict the future.

The first two sub theories are commonly accepted and this holds both for the biological world and for our computer vision one. In fact, there has been a natural evolution in the computing systems since they perform better than in the past and they are all descendent from the original von Neumann architecture.

The third sub theory is also accepted by both communities (biological and computational) considering on one hand the climate adaptation—polar, equatorial, etc.—and on the other the different required solutions involving real-time or application constraints which are essentially offline (e.g. robot vision for automatic guided vehicles for the first case or video generation and post processing for a videogame for the second case). The following Fig. 3 illustrates the branching evolution of computers, starting from the ancestral von Neumann architecture through the extinct massively parallel ones (MPP) and the surviving digital signal processors (DSP), the accelerated graphic port (AGP), the GRID network of processors and the Berkeley MOTE project 4 based on miniature active sensors which exploits a mixed solution adopting distributed computing and allowing contextual interaction between elements.

The contents of Fig. 3 illustrate the different families of computer vision architectures following the taxonomy described in (Cantoni and Levialdi, 1983). Most of the projects extinguished and DSP remains evolving towards AGP. Among the new.

---


comers it is worthwhile to quote the MOTE and the GRID families positioned on the large MIMD axis distant from the origin due to their rich PE multiplicity.

The fourth sub theory is questionable in both areas, in fact evolution as a process in which similar species are born, sounds reasonable and many have endorsed the concept of linear evolution having a monotone progress towards improvement. Nevertheless, one may not exclude the birth of a completely different species (as for instance through human genetically manipulated organisms) not necessarily in a positive direction, as warned by bioethicists. Yet, such new organisms are difficult to be conceived if reasoning is constrained by a continuous equilibrium mode. Along similar lines, technology may also be able to provide radically different solutions by means of the introduction of new, materials, or by the introduction of different, better, computational models. Examples of radically different computational engines are mote computing (using fine grain distribution of sensors and processing elements), grid computing (massive distribution of processors on Internet), computerized clothing, and ubiquitous computing.

The last and best-known sub-theory was developed before Darwin’s age and, even if some authors disagree with it, may be considered as the foundation upon which the whole bio-evolution model stands. An interesting definition of such theory was given recently by Farnell during his Ph.D. thesis (2004). In fact he states: “This simple process (natural selection) explains how the appearance of design in the natural world was not due to the efforts of a conscious designer, but instead arose from a completely mindless process…” In this connection, a further explanation of this theory is given as follows: “During the process of natural selection, then, those genes that are most conducive to survival or reproduction will, by their own effects, tend to become increasingly common
in the population with each successive generation. And in the competition between individual organisms for finite resources and potential mates, this will be at the expense of those genes’ alternatives in the population that are less conducive to survival or reproduction, and which will eventually become ‘extinct’.

Other authors have considered natural selection in completely different areas; it is worth to highlight for example the concept of population thinking by Edelman (1992), in which the individual variance in a population is the source from which selection is active to improve the species, selection occurs by dropping negative features at the end. According to Edelman this process is called “ex post facto” and can be synthesized in “evolution works by selection and not by instruction”; A similar process, claims Edelman, may describe the brain’s working principles from which creativity may emerge as discussed in (Cantoni and Levialdi, 1994) where computer programs promote a set of unforeseeable solutions to help the user/artist in his personal choice.

In a similar way to the biological natural selection, the same phenomenon has appeared in computer vision architectures. The adaptation to commercial markets has driven, ex post facto, the selection process. No special funding was provided to research in this area by any National Government. In this connection, the massive architectures have been extinguished also due to a number of market factors:

- the small amount of groups having financial resources to cover the high expenses of such machines,
- the complexity of the ad hoc or customized software,
- the problems stemming from maintenance and updating a small number of systems,
- the input/output problems coming from non-standard systems.

In fact, at architectural meetings for image processing computers progressively fewer presentations appear on new hardware and more rely on standard machines (see the CAMP—Computer Architectures for Machine Perception) series of Workshops, or the hardware track at the ICPR (International Conference on Pattern Recognition, the last one held at Cambridge, UK, 2004).

The winning solution to the architectural problem for image processing tasks will be to choose standard processors (always with more computational power following Moore’s law) but in an effective cooperative distributed environment (taking advantage of Gilder’s law).

3. Conclusions

We started with the history of innovative architectures for image processing, as they were conceived in the 70s and 80s, which were rich of proposals that harvested a wide number of concepts, projects and implementations. Azriel Rosenfeld was instrumental as far as the first two items. A number of discussions and workshops that faced all the different aspects of matching algorithms to architectures for image processing took place yearly since 1981 (Duff and Levialdi, 1981) up to 1988 (Levialdi, 1988).

The market has choked most of the new systems which came out of production in the 90s while the computing power of off-the-shelf machines increased, their reliability and modularity also improved so as to satisfy, with present-day budgets, most of the image processing tasks: industrial, biomedical, and document analysis.

Remembering that Azriel Rosenfeld used a very minimal number of slides for his presentations, we ask ourselves if one slide would be sufficient to describe the present day situation of dedicated hardware for image processing.

References


