Introduction

Engineers constantly encounter technological problems which are becoming increasingly complex. These problems may be encountered in different domains such as transport, telecommunications, genomics, technology for the healthcare sector and electronics. The given problem can often be expressed as one which could be solved by optimization. Within this process of optimization, one or several "objective functions" are defined. The aim of this process is to minimize the "objective function" in relation to all parameters concerned. Apart from problems of optimization, i.e. the problem's objective function which is part of this topic (e.g. improving the shape of a ship, reducing polluting emissions, obtaining a maximum profit), a large number of other situations of indirect optimization can be encountered (e.g. identification of a model or the learning process of a new cognitive system). When looking at this issue from the angle of available methods used to resolve a given problem, a large variety of methods can be considered. On the one hand, there are "classic methods" that rely purely on mathematics, but impose strict application conditions. On the other hand, digital methods that could be referred to as "heuristic" do not try to find an ideal solution but try to obtain a solution in a given time available for the calculation. Part of the latter group of methods is "metaheuristics", which emerged in the 1980s. Metaheuristics has many similarities with physics, biology or even ethology. "Metaheuristics" can be applied to a large variety of problems. Success can, however, not be guaranteed. The domain of optimization is also very interesting when it comes to its functions within the field of application. In the domain of optimization, the processing of signals and images is especially varied, which is due to its large number of different applications as well as the fact that it gave rise to specific theoretical approaches such as the Markov fields, to name just one example.

These ideas have influenced the title of this book *Optimization in Signal and Image Processing*. This book has been written for researchers, university lecturers

and engineers working at research laboratories, universities or in the private sector. This book is also destined to be used in the education and training of PhD students as well as postgraduate and undergraduate students studying signal processing, applied mathematics and computer science. It studies some theoretical tools that are used in this field: artificial evolution and the Parisian approach, wavelets and fractals, information criteria, learning and quadratic programming, Bayesian formalism, probabilistic modeling, the Markovian approach, hidden Markov models and metaheuristics (genetic algorithms, ant colony algorithms, cross-entropy, particle swarm optimization, estimation of distribution algorithms (EDA) and artificial immune systems). Theoretical approaches are illustrated by varied applications that are relevant to signals or images. Some examples include: analysis of 3D scenarios in robotics, detection of different aggregates in mammographic images, processing of hand-written numbers, tuning of sensors used in surveillance or exploration, underwater acoustic imagery, face recognition systems, detection of traffic signs, image registration of retinal angiography, estimation of physiological signals and tuning cochlear implants.

Because of the wide variety of different subjects, as well as their interdependence, it is impossible to structure this book – which contains 13 chapters – into distinct divisions, which might, for example, separate traditional methods and metaheuristics, or create a distinction between methods dealing with signals or with imagery. However, it is possible to split these chapters into three main groups:

– the first group (Chapters 1 to 5) illustrates several general optimization tools related to signals and images;

– the second group (Chapters 6 to 10) consists of probabilistic, Markovian or Bayesian approaches;

– the third group (Chapters 11 to 13) describes applications that are relevant for engineering in the healthcare sector, which are dealt with here through the use of metaheuristics.

Chapter 1 deals with the benefits of modelization and optimization in the analysis of images. After the introduction of modelization techniques for complex scenes, the analysis of images has become much more accurate. In particular, traditional means of image analysis, such as the segmentation of an image, need to be revised. Jean Louchet creates a link between two domains that have been developing independently. These are the synthesis and the analysis of images. The synthesis of images relies on a wide range of different modelization techniques which are based on geometrics, depiction and movement. The author shows that some of these techniques can also be used for the analysis of images, which would broaden the possible applications of these techniques. Jean Louchet also shows how artificial evolution can lead to a better exploitation of models, create new methods of

analysis and push back the limits of Hough transform using a stochastical exploration of the model's parameter space.

In Chapter 2 Pierre Collet and Jean Louchet present the so-called "Parisian" approach of evolutionary algorithms and how these algorithms are used in applications when processing signals and images. Evolutionary algorithms are reputed to take a long time to perform calculations. The authors, however, show that it is possible to improve the performance of these algorithms by – if possible – splitting the problem into smaller sub-problems. When using the "Parisian" approach to analyze a scene, the objects which have been modified by genetic operators are not the vectors of the parameters that determine a complete model of an image. These objects are elementary entities which only make sense when merged together as a representative model of the scene that will be studied. In other words, a problem cannot be represented by a single individual but by several individuals, or even the entire population. The "Parisian" approach is successfully used in the field of robotics when analyzing 3D scenes via stereovision. The so-called "Fly algorithm" allows for the detection of obstacles in real time and much more quickly than when using traditional approaches. Other visual applications based on models can be processed by evolutionary methods. Here, the authors discuss the identification of models of mechanical systems based on sequences of images.

Chapter 3 deals with the use of wavelets and fractals when analyzing signals or images. The application of these techniques is becoming increasingly frequent in natural science as well as in the study and research carried out in the scientific fields of engineering and economics. Abdeljalil Ouahabi and Djedjiga Ait Aouit show that multifractal analysis and the exploitation of techniques of multiresolution based on the concept of wavelets lead to a local as well as global description of the signal's singularities. On a local level, the criterion of punctual regularity (rugosity) based on Hölder's inequality can be characterized by the decrease of the wavelets' coefficients of the analyzed signal. On a global level, the distribution of a signal's singularities can be estimated by global measures when using the auto-similarity of multifractals. In other words, the spectrum of singularities is obtained when localizating the maxima of the module of the wavelet transform of a signal. The authors give two examples of the aims and applications of this formalism. One example in the healthcare sector is a multifractal analysis which allows for the detection of different aggregates in mammographic images. The second example is fracture mechanics. In this field the formalism described above is used to study the resistance of materials.

Chapter 4 deals with the information criteria and their applications when processing signals and images. Here, the model of a random signal should be optimized. An information criterion is a description or formulation of an objective

function that should be minimized. The information criteria are an improvement on the traditional technique of the maximum likelihood. This improvement is due to the focus being shifted towards simultaneous research on the optimal number of free parameters in the model as well as the ideal values for these parameters. Christian Olivier and Olivier Alata first give a general overview of the main information criteria as well as the relevant literature. The majority of the criteria were introduced for research using 1D auto-regressive (1D AR) models. In Chapter 4, this case is illustrated by an application that involves the segmentation of natural images. The information criteria were then transferred to the 2D AR model. Two applications resulted from this. These are the modelization of the image's texture and the unsupervised segmentation of textured images. The authors then look at the extension of the information criteria to other models based on parameters. These are a mix of Gauss's laws n-D, which are here applied to unsupervised classification as well as Markov's modes. Last but not least, this chapter deals with the application of information criteria in the case of non-parametrical problems, such as the estimation of distribution via histograms or the search for antiderivatives that carry a maximum amount of information depending on the form of the information. The information criteria finally offer a means to justify the choice of parameters which are linked to a large number of problems when processing signals or images. The information criterion deals with a high number of observations. This is why the time required to carry out the calculation might be high (particularly in an unsupervised context). Dynamic algorithms, however, are able to reduce the number of operations that need to be carried out.

Chapter 5, written by Gaëlle Loosli and Stéphane Canu, deals with an aspect of optimization that can currently be encountered within signals and images, for example in shape recognition, i.e. learning processes. More precisely, the chapter focuses on the formulation of learning as a problem in convex quadratic programmation on a large scale (several million variables and constraints). This formulation was obtained by the "nucleus methods", which emerged about a decade after neural networks. Its main aim is linked to the fact that the solution in question is often "parsimonious", i.e more than 99% of all unknown variables are zero. Using this specific feature enables learning algorithms to solve large scale quadratic programming problems within a reasonable amount of time. The best-performing methods, known as "active constraints", work "off-line". In other words, they are able to determine the exact solution of a given problem if they are provided with all the data that is used in the learning process. To carry out an "online" learning process, a method of iterative stochastic optimization is used, which allows us to obtain an approximate solution. This chapter describes one of the methods which is part of the "support vector machine" (SVM) type. The efficiency of this technique is illustrated by results of experiments which focused on the problem of recognizing handwritten numbers.

Chapter 6 deals with the problem of planning within time and space the use of sensors with the aim of optimizing the exploration and surveillance of a specific zone; given the rather low number of available sensors as well as their capacity, this zone is large. Due to the problem being rather extensive, exact methods cannot be used. An approximate solution can, however, be obtained with the help of metaheuristics. In this case, Frédéric Dambreville, Francis Celeste and Cécile Simonin, the authors of this chapter, recommend the use of "cross-entropy". This method was initially created to evaluate the probability of rare events and has been adapted to "difficult optimization" problems (many local minima need to be considered). The solution is obtained with the help of a probability law that continually approaches the global optimum. This method is applied to the problem of planning sensors via *a priori* modeling mainly under the form of different groups of probability laws, of possible planning policies. In this chapter, three examples are explained in detail. The first example looks at how to ideally array search units in the context of military operations. The aim is to maximize the probability of locating the target which does not move but is hidden. In the second example, cross-entropy is used for an exploratory mission. The movement of the vehicle needs to be planned based on maps that show the environment. The third example is the problem of optimal control in an environment where only certain parts of the environment can be observed. Cross-entropy is particularly useful when dealing with data that are very difficult to formalize. Optimization via cross-entropy therefore means to "learn" an optimal strategy.

The topic of Chapter 7 is linked to that of the previous chapter. Chapter 7 deals with a surveillance system such as a maritime patrol aircraft that needs to locate a moving target. In order to do this, all resources, i.e. passive as well as active sensors (e.g. a radar), need to be used. Passive measures do not involve any cost. However, they only determine the direction of the target. Active measures provide much more information since they can evaluate the distance to the target. These measures, however, need to be used sparsely because of their cost (emitting a wave) and with discretion. The author of this chapter, Jean-Pierre Le Cadre, gives a general outline of the problem of optimal and temporal repartition when using active measures. He futhermore describes the general mathematical tools (e.g. multilinear algebra) that allow for the analysis of this problem. The study focuses on the explicit calculation of objective functions while expressing the quality of the estimation (or tracking) of the trajectory's location by using non-linear observations of state. First of all, this chapter examines the case of targets that contain a determined trajectory. Their movement is rectilinear and uniform, or in other words the target is "maneuvering". When dealing with certain types of approximations, the problem of convex optimization comes into play. This problem can easily be resolved. The author also looks at the stochastic evaluation of this case. He shows that it is possible to directly calculate the objective function of a target of Markovian trajectory without having to use simulations.

Chapter 8 deals with segmentation methods of images which exploit both the Markovian modeling of images and the Bayesian formalism. For every image under observation there is an infinite number of combinations of objects that can be associated with it. These combinations of objects represent, or in other words create ,the image. To reduce the number of possible solutions that should be integrated in the stage of segmentation, prior local or global knowledge is required. The aim of Markovian modeling lies precisely in its capacity to locally describe global properties. Due to the equivalence between Markov's field and Gibbs's distribution, the optimal segmentation can be obtained by the minimization of a function linked to energy. Christophe Collet, the author of this chapter, applies this formalism to the context of underwater acoustic imagery. To detect small objects on the seabed, the author exploits images that have been taken by a lateral multibeam sonar. The images that were obtained were distorted by noise. A segmentation of good quality therefore requires the nature of noise to be taken into consideration during the process of image modeling. This chapter shows different examples of application. These are the segmentation of sonar images into two different groups (shadow, reflections of the seabed) or segmentation into three different groups (shadow, seabed and echo). Due to the third group, echo, physics, which forms the basis of the creation of sonar images, is also taken into consideration. Two other examples are the differentiation between manufactured and natural objects, as well as the subdivision of the seabed into different regions (sand, mid-ocean ridges, dunes, stones and rocks). All tasks linked to detection and classification are first of all united in the fact that the function of energy, which integrates the prior knowledge required to obtain a solution, needs to be minimized. The technique used for this optimization is a deterministic method or a genetic algorithm, depending on whether an initial good quality solution is available or not.

Chapter 9 was written by Sébastien Aupetit, Nicolas Monmarché and Mohamed Slimane and describes the use of hidden Markov models (HMM) for the recognition of images. Hidden Markov models are statistical tools which allow for the modelization of stochastic phenomena. This type of phenomenon may, for example, consist of several sequences of images. Images of the same sequence are taken from different angles but show the same scene, e.g. a person's face. After a learning phase, HMM is prepared for the process of recognition. During this learning phase several sequences of images, let us say four sequences of four photographs each showing the faces of four different people are processed. When confronted with a new photograph of a face, HMM is able to distinguish which person is shown in the picture from the four previous pictures. At the same time, the risk of HMM making a mistake is minimized. More precisely, a discrete HMM corresponds to the modeling of two stochastic processes. The first process is hidden and perfectly modeled by a discrete Markov chain while the second observed process is dependent on the state of the first process, i.e. the hidden process. This chapter focuses on learning processes, a crucial aspect of HMM. It provides an overview of the main

criteria of existing learning processes and the possible solutions for HMM learning processes. Furthermore, the principles of three metaheuristics inspired by biology and population-based are also addressed by the authors and analyzed in light of HMM learning processes. These three metaheuristics are a genetic algorithm, ant colony algorithm and particle swarm optimization (PSO). Several versions of these types of metaheuristics (which are different to one another because of the mechanisms which are implemented, or simply due to the settings of the respective methods) are examined and tested in great detail. These tests are carried out on a set of test images as well as samples of literature. The chapter emphasizes the fact that results can be improved if metaheuristics used for learning processes are combined with a method dedicated to local optimization.

In Chapter 10 Guillaume Dutilleux and Pierre Charbonnier use different metaheuristics inspired by biology for the automatic detection of traffic signs. The aim is to make an inventory of road signs currently used in the French secondary road network. The data used are images that have been collected by vehicles inspecting the roads that are part of the respective network. The application does not face any real time constraint. However, the application needs to be robust when faced with changes in the conditions under which the images are collected. Problems might occur due to differences in light, backlighting, worn out or partially hidden traffic signs. The method that has been proven to be successful includes the technique of "deformable models". This technique consists of a mathematical model, a prototype of which the object research is carried out upon. This model's shape can be manipulated and changed to such an extent that it is adapted to the respective image that should be analyzed. The quality of this adjustment and to what extent manipulation can be accepted are, in the case of Bayesian formalism, respectively measured by a likelihood and an *a priori*. The problem of localizing an object therefore comes down to the problem of optimization in the sense of a maximum *a posteriori*. The residual value of a minimized objective function gives an indication of the effective presence of the object in the scene which is to be analyzed. In practice, the presence of numerous local minima justifies the use of metaheuristics. The authors have carried out experiments with three different techniques in the field of metaheuristics. These are an evolutionary strategy, PSO and a method of clone selection (the latter is relevant to a more general field of "artificial immune systems"). The performance of automatic detection is compared to a number of different algorithms when dealing with a sequence of traffic signs. (For these test images the real data had already been obtained manually.)

The majority of metaheuristics were initially created for the processing of problems that arise when dealing with discrete optimization. Chapter 11, written by Johann Dréo, Jean-Claude Nunes and Patrick Siarry, looks at their adaptation to applications with continuous variables, which are encountered frequently, especially in the field of signals and images. The techniques suggested in the literature for this

adaptation are linked to each specific form of metaheuristics. These techniques cannot be generalized, i.e. it is not possible to apply these techniques to another application. Furthermore, no metaheuristic, whether it is continuous or discrete, is the ideal technique, i.e. most efficient, for all possible sorts of problems. This is why hybrid methods, which combine different forms of metaheuristics or metaheuristics with downhill simplex techniques, often need to be used. This chapter describes two "continuous metaheuristics". These are an ant colony algorithm and EDA. Furthermore, a local technique, which is frequently used in continuous cases to refine the search within a "promising valley" of solutions, is Nelder and Mead's downhill simplex method. These methods are used for image registration in the field of retinal angiography. Before a doctor can actually interpret a sequence of images, the problem of inevitable eye movement during the procedure needs to be dealt with. In the example given in this chapter, image registration is carried out by using only translatory motions between different images. Metaheuristics were found to be particularly appropriate for image registration in angiography with a high resolution. The time required for calculations only increases a little when increasing the resolution of images.

Chapter 12, written by Amine Naït-Ali and Patrick Siarry, describes the introduction of a genetic algorithm used for the estimation of physiological signals, the Brainstem Auditory Evoked Potentials (BAEP). BAEP is an electric signal which is generated by the auditory system as a response to acoustic stimulation. Studying this signal allows for the detection of pathologies such as acoustic neuroma. Measuring BAEP is, however, a problem as this signal is of a very low energy and covered by electric noise that stems from spontaneous electric activity of the cerebral cortex (these signals can be measured using electroencephalograms (EEG)). To identify a patient's effective BAEP, several hundred signals need to be exploited. These signals are obtained as a result of acoustic stimulation. They also have to be synchronized before being simply added to one another in order to eliminate the noise. The synchronization process is expressed in the form of an optimization problem in which unknown variables are the random delays of different signals. Here, the problem is solved with the help of a genetic algorithm. The authors show that a significant acceleration of this technique can be obtained when creating a model for the variation law of these delays. This can, for example, be performed using a set of sinusoids.

Chapter 13, written by Pierre Collet, Pierrick Legrand, Claire Bourgeois-République, Vincent Péan and Bruno Frachet, presents an evolutionary algorithm that allows for the adjustment of parameters for a cochlear implant. This adjustment is carried out in interaction with the patient using the device. This type of implant enables deaf people, whose cochlear plate is still intact, to hear. The device works as follows: a group of electrodes is implanted into the patient's cochlear plate. These electrodes stimulate the auditory nerve. The electrodes are connected to a digital

signal processor (DSP) that receives the sound as signals through a microphone situated next to the patient's ear. The parameters of DSP need to be adjusted in a way that reconstructs the patient's auditory ability to a point that he/she might even be able to understand spoken language. Adjusting these parameters is usually undertaken by a human and becomes increasingly complicated as technology progresses. A current implant consists of 20 electrodes and several hundred parameters. The effort for adjusting these parameters is dependent on the patient's ability to understand spoken language. This is why this study looks at the performance of an interactive evolutionary algorithm which should take over the task of adjusting the parameters of a cochlear implant. There are a large number of difficulties that lie within this application. These are the subjective evaluation of every single patient, the quality of every single solution produced by the algorithm, the necessity of a rapid convergence of the algorithm in order to strictly limit the amount of solutions to be evaluated by the patient (as every evaluation takes a few minutes) as well as the fact that the search space is very broad. This chapter presents experiments undertaken by the authors with the help of a small number of patients following a methodical protocol. The first results are promising. They show the disadvantages of manual adjustment in cochlear implants which is increasing because the number of available electrodes is currently increasing.

Chapter 1

Modeling and Optimization in Image Analysis

1.1. Modeling at the source of image analysis and synthesis

From its first days, image analysis has been facing the problem of modeling. Pioneering works on contour detection led their authors to refer to explicit models of edges and noise [PET 91], which they used as a conceptual basis in order to build their algorithms. With an opposite approach to these phenomenological models, a physical model of light diffusion on surfaces has been used as the basis for Horn's works [HOR 75] on shape from shading. More generally, a phenomenological model aims at describing a directly computable property of the geometric configurations of gray levels on an image; the physical model then tries to use the knowledge corpus of physics, or even sometimes to create an *ad-hoc* conceptual system, as we will see later. Between these two extremes, there is a large number of approaches to modeling. Here, we shall try to illustrate them using some examples.

It is important to first show the links between image analysis and synthesis. For a number of years, these two domains have been undergoing largely independent development processes. In spite of their conceptual similarity, they have been dealt with by two separate scientific communities, with different origins and centers of interest, which did not address the same applications. Robotics is one of the few fields of application that has played an important role in moving them closer to one another. Image synthesis addresses another large panel of approaches to modeling, in particular in the fields of geometry, rendering and motion modeling – to such an extent that there are important international communities, journals and conferences that specialize in each of these approaches to modeling. One of the benefits of

Chapter written by Jean LOUCHET.

connecting image analysis and image synthesis comes from the fact these two domains often use common modeling techniques which they use as bridges to their constructive interaction.

1.2. From image synthesis to analysis

One of the difficult points in image analysis and machine vision is algorithm validation methods. In most cases it is not possible to access the "ground truth" that corresponds to the image or the image sequence on which we want to evaluate the quality of the analysis process. A frequent compromise consists of evaluating an algorithm by referring to another algorithm, which is seldom acceptable. The ideal, straightforward solution would consists of creating a synthesis tool – and therefore a model – able to build dependable test data from any ground truth.

Once the effort of building such a model has been carried out, we naturally arrive at the idea of incorporating into the image analysis algorithm, the knowledge of the physical world and its rules, following an "artificial intelligence" approach. This can be done implicitly (using the same knowledge corpus and coding it in a way suitable to the analysis algorithm) or alternatively by explicitly incorporating the model into the analysis process. Of course, this raises the delicate ethical question of mutually validating two algorithms running the same model, and therefore susceptible to containing the same errors or clumsy simplifications. In any case, this is, when pushed to its extreme, the basis of the so-called "analysis by synthesis" approaches where rather than the model (whether photometric, geometric, kinematic or physical), it is the whole image synthesis process that is embedded into the analysis algorithm. Between merely using general physical knowledge in the analysis process, and at the other end, embedding a complete synthesis process into the analysis algorithm, it appears that the image analysis techniques explicitly exploiting a model are undergoing an important development, particularly thanks to modern optimization techniques, as we will see in several examples.

First, we will examine the classical approaches to image segmentation and show how they have built their organization, often implicitly, after the way scene models used in image synthesis are naturally organized.

In the next section, we will revisit the Hough transform [HOU 62], which is probably the best known example of image analysis and model inversion, through deterministic, exhaustive search in a parameter space; the model used here is phenomenological (visual alignment). We show the Hough transform and its generalizations may be rewritten into an evolutionary optimization version; as a stochastic exploration of a parameter space, here each point represents a particular instance of a model. This considerably widens the field of potential applications of the original Hough transform.

The following part will quickly examine the contribution of physical models to image analysis. This is a promising yet little known topic we will discover through two examples using photometric and dynamic models.

In some applications, the model underlying the analysis technique may be taken apart into elementary objects whose collective behavior actually represents the object to be modeled. A specific evolutionary optimization method, called "Parisian Evolution", can then be implemented. This is a change in the semantics of the evolved population but a classical evolutionary process is still applied to the elementary objects. This will be the subject of Chapter 2.

1.3. Scene geometric modeling and image synthesis

As discussed earlier, image contour segmentation took its foundations from hypotheses about image signals, resulting into a wide use of differential operators as main analysis tools. Region segmentation, which was developed later, probably because of its greater need for computational power, brought more evidence of the strong link between the structures of a 3D scene and the image entities directly accessible to calculation. It is therefore tempting to revisit the notion of image segmentation [COC 95, GON 92] through its possible interpretations in terms of scene models.

Seen from this point of view, segmentation into regions could be defined as any image partitioning technique such that each region entity it extracts is a good candidate image projection of a 3D or space-time varying physical object in the scene.

Similarly, it is possible to give a new definition of contour segmentation: it describes any image line extraction technique such that each line extracted is likely to be the image projection of an edge of a physical object present in the scene.

With each level of primitives in the polyhedral model (vertex, edge, facet, etc.) it is possible to associate a *probable* local property of the image, such as the contrast along a line, the homogenity of a region, etc., and a corresponding calculation technique. Classical segmentation techniques are often a decisive step in the process of instantiating the model as efficient model exploration heuristics; contours usually are the projection of the subset of the scene where the probability of finding an edge is highest, thus the knowledge of contours contributes to the efficiency of the exploration of the space of parameters which describe the possible positions of edges. Similarly, interest points give useful hints on where to look for polyhedron vertices, and so on.

One of the consequences is that the pertinence of a segmentation technique on a class of images essentially depends on whether it actually corresponds to an observable characteristic of the model underlying the class of scenes and how the

images have been captured, An illustration of this is given by fluid flow imaging, where polyhedral models are irrelevant and classical contour or region segmentation techniques are just as irrelevant. Segmentation techniques are the translation of the scene-specific description language.

The primary role of image analysis is to instantiate or identify the parameters of a general scene model. If we consider scenes made from opaque objects, which is not too bad in most familiar scenes, the most widely used modeling language in image analysis, as in synthesis, is based on polyhedral objects. The ultimate goal of image segmentation should ideally be to provide a description of the scene using the same primitives and language as in image synthesis: a geometrical description (polyhedra, facets, edges, vertices), completed (if useful to the application) with a photometric description (light sources, radiance factors, diffusion coefficients, etc.). In the case of time-dependent sequences, it will be necessary to include object motion and deformations, and the analysis may even include the building of a description of the scene in terms of agents, individual behaviors and physical interaction [LUC 91].

In all these cases, "informing the model" means optimizing the likeness between real data and data synthesized from the scene model, and therefore will generally involve the optimization of a cost or resemblance function.

1.4. Direct model inversion and the Hough transform

1.4.1. *The deterministic Hough transform*

One of the main motivations of the development of image segmentation techniques is the difficulty of directly resolving the problem of optimizing a scene model using classical methods. However, a well known exception to the rule is given by the Hough transform [HOU 62, BAL 82], which may be described as a direct parameter space exploration technique. It consists of filling in the space of model parameters (also known as the *Hough accumulator*) using a vote technique, where each locally detected primitive in the image results in incrementing a sub-variety of the parameter space. The parameter vector that will be eventually elected is the one which has received the greatest number of votes, and therefore is likely to represent the best possible model, i.e. the most satisfying *a priori* global explanation of all the primitives that have been previously extracted from the image.

In spite of a reasonable success story, the Hough transform and its priority-to-image philosophy imply that for each image primitive considered to be relevant, the accumulation process will go into the n-dimensional parameter space to modify the values of all the points belonging to a variety with dimension $n - 1$ (within the limits of the search space). The heuristics that have been found in order to improve the algorithm's speed have a limited effect and the generalized Hough transform becomes unusable when the dimension of the search space becomes greater

than 3 or 4, mainly due to available memory space, memory access time and the complexity of the dual space incrementation task.

1.4.2. *Stochastic exploration of parameters: evolutionary Hough*

If we consider the Hough technique, which involves calculating voting scores throughout the parameter space then exhaustively exploring this space in order to find the best optima, it becomes an attempt to directly explore the search space [ROT 92, LUT 94]. Evolutionary programming provides us with very welcome exploration techniques which, in our case, allow us only to calculate the values in the dual space where individuals of the evolving population actually are, rather than on all of the dual space.

According to the general principles of artificial evolution (also known as evolutionary programming), a function to be optimized is given, though not explicitly. We then consider an arbitrary set of individuals ("population") which belong to the space where this function is defined. This population is then evolved, in a way reminiscent of biological evolution, using genetic operators such as mutation crossover and selection. The selection criterion is the individual's performance ("cost" or "fitness") as given by the function to be optimized. The expression "evolutionary strategy" [REC 94, BÄC 95] refers to the artificial evolution algorithm, where gene coding is performed using real numbers as variables, unlike genetic algorithms which, strictly speaking, use Boolean variables.

Thus it is possible to define an evolutionary version of the Hough transform:

– the population is a finite subset of the parameter space;

– for each individual in the population, the fitness function gives a measurement of the pertinence of the corresponding image pattern;

– classical selection operator (tournament);

– barycentric crossover;

– mutation: Gaussian noise.

The fitness function used in the evolutionary version of the Hough transform is calculated according to the same criterion as in the classical versions of Hough; for example, if the criterion for a point to participate into the incrementation of the accumulator is that its contrast is greater than a given threshold, then in the evolutionary version the fitness of an individual (representing e.g. a straight line) will be the number of points on this line with a contrast higher than the same threshold.

In the case of the classical Hough transform, there is no very clear advantage to either approach. Indeed, the classical Hough method in which each image point (x, y) votes for the set of points (θ, ρ) in the dual space, such that $\rho = x \cos\theta + y \sin\theta$, is relatively fast and needs a reasonable memory allocation (Figures 1.1 and 1.2); on

Figure 1.1. *Result of the classical Hough transform (image 288×352)*

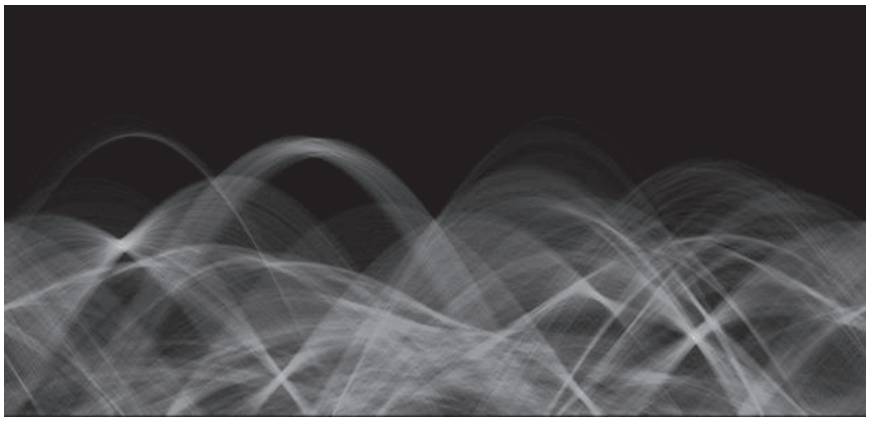

Figure 1.2. *The* (θ, ρ) *Hough accumulator corresponding to the same image (image 628×300)*

the other hand, the evolutionary process which, thanks to the sharing operator, is able to find several different solutions (Figure 1.3), suffers from not having a canonically defined ending: it is therefore difficult to compare processing times. In practice, the number of generations needed to ensure convergence results in similar calculation times.

Figure 1.3. *Result of the evolutionary version of Hough*

However, when the dimension of the parameter space becomes higher, dual space storage and exploration soon become prohibitive, while the evolutionary version is much less greedy in terms of memory and computing time. The evolutionary version of the Hough transform really becomes interesting when considering more complex parametric optimization problems as shown in the following examples, where a classical Hough approach would fail.

1.4.3. *Examples of generalization*

The following example[1] consists of detecting circles with unknown diameters (a moving ball) in an image sequence. The individuals are triples (a, b, r) which define circles with equation $(x-a)^2 + (y-b)^2 = r^2$. The fitness of a particular individual is defined as the average gradient norm taken on 40 points randomly distributed on the circle. The algorithm parameters are:

population size	100
selection	2-tournament
mutation rate (%)	15
r mutation amplitude	10
a, b mutation amplitude	40
barycentric crossover rate (%)	5
number of generations per frame	init. 800 then 240

1. These results were produced by A. Ekman, a PhD student at KTH Stockholm in October 2004.

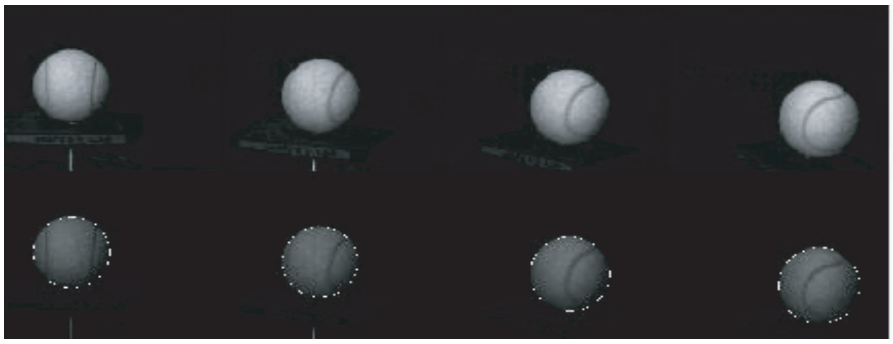

Figure 1.4. *Four original images from the "tennis ball" sequence (top) and results of the "evolutionary Hough" detection of circles with unknown radius (bottom)*

It is worth noting an interesting property of this approach: if motion is small enough between two consecutive frames, the evolutionary algorithm is able to track the object's motion, unlike its deterministic counterpart which has to resume its calculations from the very beginning, even with a tiny image change, whatever the degree of redundancy in their information contents. In other words, in spite of the fact artificial evolution is often regarded as slow, the evolutionary version of the Hough transform possesses true real time properties that its deterministic version does not have.

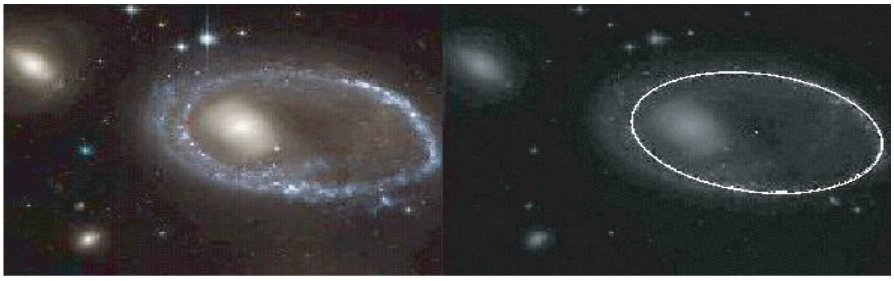

Figure 1.5. *Image of the galaxy AM 0644-741 taken by the Hubble telescope (left) and the result of the evolutionary Hough ellipse detection algorithm (right)*

Another example (Figure 1.5) consists of detecting a conical section (ellipse) in an image, using the same method and the same adjustment of genetic parameters as with the circle detection, but with a 5-parameter genome corresponding to the ellipse parametric equation:

$$x = a + r_x \cos\alpha \cos\theta + r_y \sin\alpha \sin\theta$$
$$y = b - r_x \cos\alpha \sin\theta + r_y \sin\alpha \cos\theta$$

It is not easy to compare the theoretical performance of classical (filling up parameter spaces) Hough methods to their evolutionary versions, as strictly speaking the number of generations required for convergence depends on analytical properties of the fitness function, which is image-dependent. However, let us re-examine the practical examples given above. With a square $N \times N$ image and an accumulator (parameter space) with dimension n, where each parameter can take P different values, the classical Hough transform needs a memory with size P^n, and roughly $N^2 \times P^{n-1}$ calculations to fill the accumulator. The evolutionary version does not require large memory resources, and it needs $G \times E \times N$ calculations, where G is the number of generations and E the population size. In the case of ellipse detection ($n = 5$), the images (213×161 pixels) correspond to $N \approx 200$ and the precision of quantization to $P = 200$; the classical Hough method would use a 300 GB accumulator and about 6×10^{12} calculations, against about 5×10^6 calculations with the evolutionary method, which gives a gain factor about 10^6. In the simpler case of circles (3 parameters) the computing time ratio falls to about 100, with a Hough accumulator size about 10 MB. From these examples it is safe to say that with three parameters, the evolutionary method is more efficient than the classical Hough method, but with 4 or more parameters, the only realistic method is the evolutionary version. This gives quite an important extension to the field of potential applications of the Hough transform.

1.5. Optimization and physical modeling

In the discussion above, we considered image analysis as the instantiation of a scene model. This leads us to wonder whether all the aspects of modeling used in image synthesis and computer graphics are still relevant in image analysis. While it still looks premature to give an exhaustive answer, we will examine two aspects of this question: the photometric models and the motion models.

1.5.1. *Photometric modeling*

Photometric modeling is essential to image synthesis. Rather surprisingly, photometry is not yet in wide use in image analysis. The main explicit application of photometry to image analysis is shape from shading, which consists of building the shape of a surface from the variations of the luminance. In his pioneering works, Horn [HOR 75] showed that it is possible, through the resolution of differential systems initialized on local luminance extrema, to recover the 3D shape of a surface, under relatively strong hypotheses: the surfaces are assumed to have Lambertian scattering properties, the light sources to be at infinity, etc. Practical applications have been limited by these constraints and the ill-conditioning of the problem. We may however note that recent research by Prados [PRA 04] showed that taking into account the distance-dependent reduction of lighting (light source at a finite distance) allows us to obtain a well-conditioned problem and become free of some of the usage restrictions.

Another example has been given by G. Maurin [MAU 94] with the 3D location of a light source. In this preliminary study, the author exploits one of the images from a stereo pair of an interior scene which has previously been analyzed in three dimensions; thanks to homogenity hypotheses concerning the regions detected, he calculates the position of the (single) light source in a room.

Another original example of how to exploit a simple photometric model in image analysis was given by J.B. Hayet [HAY 98] with the Robocop project (obstacle location through observation and calculation of shadows). In this application, rather than using a second camera, one or several computer-controlled light sources are set on the robot: detecting the shadow edges in each lighting configuration enables a cheap and fast 3D analysis of the scene and an elementary obstacle detection.

1.5.2. *Motion modeling*

1.5.2.1. *Kinematic modeling*

Motion modeling is the heart of image animation. Motion can be modeled at several levels, the kinematic level being the simplest and the most widely used. It consists in analyzing, in a purely geometric way, the motion of an object in the scene. It may be analyzed as a planar motion (e.g. for a pure translation, through the exploitation of the apparent motion constraints equation [HOR 81]); concerning more general planar movements, it is possible to exploit an equation which delivers the instantaneous planar rotation centers [LOU 96a]; 3D movements can also be analyzed directly [HUA 83].

However, this is not all about motion analysis and modeling. To each possible level of motion modeling, there is a corresponding image sequence analysis technique. Classical motion analysis (resulting in optical flow data) corresponds to a purely kinematic modeling of motion; similarly, physical or behavioral modeling approaches may be matched to a corresponding semantic level of image sequence analysis.

1.5.2.2. *Physical modeling*

Here, we will examine how physical modeling of motion may be used in image sequence analysis, using the mass-link modeling paradigm as it was developed by the Acroe team in Grenoble, France [LUC 91]. The CORDIS-ANIMA model and system were created in the 1970s with the main application objective of multi-modal man-machine interfacing, incorporating acoustical visual and gestural modalities. In addition to specialized peripherals such as their retroactive gestural transducer, the conceptual heart of the system is a physical modeling language based on two primitives: the mass (fitted at each instant, with a position and a velocity) and the link (between two visco-elastic and generally non-linear masses). An important conceptual, theoretical and experimental construction has allowed us to demonstrate the power of this approach through its applications in particular to the interactive

synthesis of sounds and images. One of the problems to have been addressed around this project is the inverse problem which consists of building a physical model able to reproduce as accurately as possible the given motion, deformation and interaction of an object. This problem has been partially resolved through a decomposition of the mass-link structures and a multi-objective evolutionary strategy to optimize the model [LOU 94, STA 02] using individual cost functions associated with each of the masses.

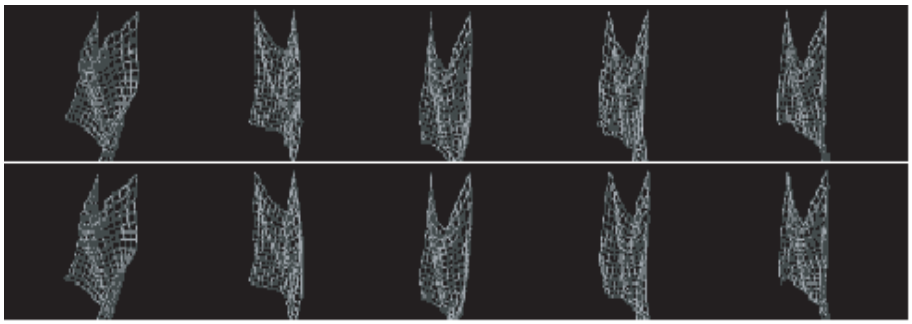

Figure 1.6. *Reconstruction of an image sequence of a cloth: original synthetic sequence (first line) and reconstructed sequence (second line) using cloth physical parameters identified from the original image data (modeling and images: X. Provost)*

One of the applications consisted of identifying the internal mechanical parameters of a cloth from an image sequence of a hanging cloth sample, then from these parameters re-synthesizing images of cloth from the same fabric (Figure 1.6) in more complex configurations [LOU 95]. In a similar way, from an image sequence representing a compressible viscous fluid flow, Jiang Li showed [LOU 96b] it is possible to identify the Cordis-Anima parameters that characterize the fluid's viscosity and compressibility and re-synthesize other flows from the same fluid (Figure 1.7).

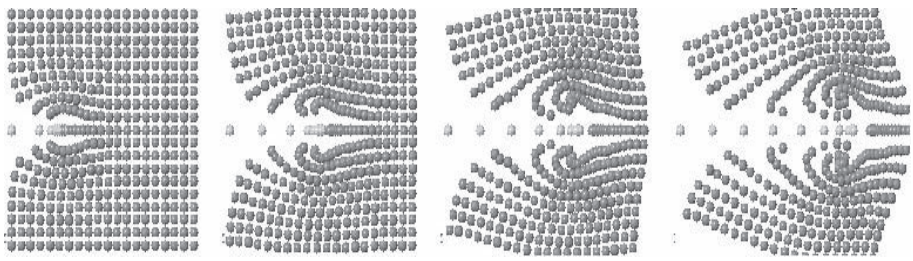

Figure 1.7. *Four successive synthesized images of the turbulent meeting of two compressible fluids (images: L. Jiang)*

Another application in the same spirit, consisted of detecting heart stroke scar zones from X-ray scanner image sequences. The heart is modeled as a mass-link system where the internal parameters are then identified using image sequence data. Anomalous values of the internal parameters of viscous-elastic links indicate a high probability of having a scar zone in the corresponding region (Figure 1.8).

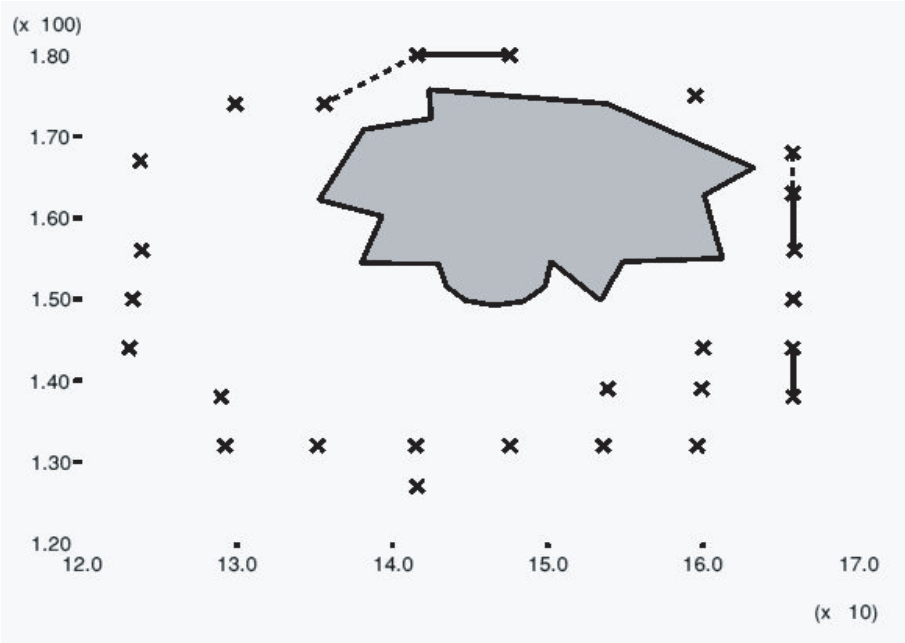

Figure 1.8. *Heart X-ray image sequence analysis: the gray zone represents the necrosed area, the crosses are calibration points. Dotted lines show where the algorithm found high stiffness using a planar model based on a single slice of the 3D data. Continuous lines represent links with high stiffness in a 3D model based on all the slices*

1.6. Conclusion

In this chapter, we revisited some classical approaches to image processing in the light of implicit or explicit scene modeling. This allowed us to outline a rarely described logical organization of the existing methods in image processing, setting some light into certain research directions still little exploited and sometimes promising. It looks like many modeling techniques developed by researchers in other communities, in particular by the image synthesis community, have potential applications to the domains of image or image sequence analysis. In a way, the challenge of modeling is not so much adding a specific technique to the existing panoply of image processing, but rather reconsidering the extension of the semantic field of what is commonly called “image processing” and “machine vision”.

1.7. Acknowledgements

Thank you to Pierre Collet, Anders Ekman, Jean-Loup Florens, André Gagalowicz, Jiang Li, Annie Luciani, Evelyne Lutton, Xavier Provot, Georges Stamon, Bogdan Stanciulescu and all the colleagues who contributed the ideas, algorithms and results described or quoted in this chapter.

1.8. Bibliography

[BÄC 95] BÄCK T. and SCHWEFEL H.-P., "Evolution Strategies I: variants and their computational implementation", *Genetic Algorithms in Engineering and Computer Science*, John Wiley & Sons, 1995.

[BAL 82] BALLARD D.H. and BROWN C.M., *Computer Vision*, Prentice Hall, 1982.

[CAD 94] CADOZ C., *Les réalités virtuelles*, Dominos/Flammarion, 1994.

[COC 95] COCQUEREZ J.-P., PHILIPP S. *et al.*, *Analyse d'images: filtrage et segmentation*, Masson, 1995.

[COL 00] COLLET P., LUTTON E., RAYNAL F. and SCHOENAUER M., "Polar IFS + Parisian Genetic Programming = Efficient IFS inverse problem solving", *Genetic Programming and Evolvable Machines*, vol. 1, pp. 339–361, 2000.

[DEL 93] DELNONDEDIEU Y., LUCIANI A. and CADOZ C., "Physical elementary component for modelling the sensory-motricity: the primary muscle", *4th Eurographics Workshop on Animation and Simulation*, pp. 193–207, Barcelona, September 1993.

[GON 92] GONZALEZ R.C. and WOODS R.E., *Digital Image Processing*, John Wiley & Sons, 1992.

[HAR 80] HARALICK R.M., "Using perspective transformations in scene analysis", *Computer Graphics and Image Processing*, vol. 13, pp. 191–221, 1980.

[HAY 98] HAYET J.-B. and TADJADIT M., Projet ROBOCOP: repérage d'obstacles par observation et calcul des ombres portées, rapport de stage ENSTA, PPL 99/11, 1998.

[HOR 75] HORN B.K.P., "Obtaining shape from shading information", *The Psychology of Computer Vision*, McGraw-Hill, 1975.

[HOR 81] HORN B. and SCHUNCK B., "Determining optical flow", *Artif. Intell.*, vol. 17, pp. 185–203, 1981.

[HOU 62] HOUGH P.V.C., Method and Means of Recognising Complex Patterns, U.S. Patent no. 3 069 654, 18 December 1962.

[HOU 92] HOUSE D.H., BREEN D.E., GETTO P.H., "On the dynamic simulation of physically-based particle-system models", *Proceedings of EuroGraphics'92 Workshop on Animation and Simulation*, Cambridge England, 5-6 September 1992.

[HUA 83] HUANG T.S., *Image Sequence Processing and Dynamic Scene Analysis*, Springer Verlag, Berlin, 1983.

[LOU 94] LOUCHET J., "An evolutionary algorithm for physical motion analysis", *British Machine Vision Conference*, York, BMVA Press, pp. 701–710, September 1994.

[LOU 95] LOUCHET J., PROVOT X. and CROCHEMORE D., "Evolutionary identification of cloth animation models", *Computer Animation and Simulation '95, Proc of the Eurographics Workshop*, Maastricht, Springer, pp. 44–54, September 1995.

[LOU 96a] LOUCHET J. and BOCCARA M., CROCHEMORE D. and PROVOT X., "Building new tools for synthetic image animation using evolutionary techniques", *Evolution Artificielle/Artificial Evolution 95*, Brest, September 1995, Springer Verlag, 1996.

[LOU 96b] LOUCHET J. and JIANG L., "An identification tool to build physical models for virtual reality", *IWSIP Manchester*, UK, November 1996.

[LUC 91] LUCIANI A., JIMENEZ S., CADOZ C., FLORENS J.-L. and RAOULT O., "Computational physics: a modeler-simulator for animated physical objects", *EuroGraphics '91 Conference*, Vienna, Elsevier Science Ed., 1991.

[LUT 94] LUTTON E. and MARTINEZ P., "A genetic algorithm for the detection of 3D geometric primitives in images", *12th ICPR*, Jerusalem, Israel, October 9-13, 1994/INRIA technical report # 2210.

[MAU 94] MAURIN G. and GAGALOWICZ A., Localisation 3-D de sources de lumière par utilisation des variations lentes d'illumination dans les images, rapport de stage INRIA/ENSTA, EPR, 1994.

[PRA 04] PRADOS E. and FAUGERAS O., "Unifying approaches and removing unrealistic assumptions in shape from shading: mathematics can help", *Proc. ECCV04*, 2004.

[PET 91] PETROU M. and KITTLER J., "Optimal edge detector for ramp edges", *IEEE Pattern Analysis and Machine Intelligence*, vol. 13, no. 5, pp. 1483–1491, 1991.

[REC 94] RECHENBERG I., "Evolution strategy", in ZURADA J.M., MARKS R.J. and ROBINSON C.J. (Eds.), *Computational Intelligence Imitating Life*, IEEE Press, pp. 147–159, 1994.

[ROU 81] O'ROURKE J., "Motion detection using Hough technique", *IEEE Conference on Pattern Recognition and Image Processing*, Dallas, pp. 82–87, 1981.

[REY 87] REYNOLDS C., "Flocks, herds and schools: a distributed behavioural model", *Computer Graphics (Siggraph)*, vol. 21, no. 4, pp. 25–34, 1987.

[REE 83] REEVES W.T., "Particle systems – a technique for modelling a class of fuzzy objects", *Computer Graphics (Siggraph)*, vol. 17 no. 3, pp. 359–376, 1983.

[ROT 92] ROTH G. and LEVINE M.D., "Geometric primitive extraction using a genetic algorithm", *IEEE CVPR Conference*, pp. 640–644, 1992.

[SER 99] SER P.K., CLIFFORD S., CHOY T. and SIU W.C., "Genetic algorithm for the extraction of nonanalytic objects from multiple dimensional parameter space", *Computer Vision and Image Understanding*, vol. 73, no. 1, pp. 1–13, 1999.

[STA 02] STANCIULESCU B., Modèles particulaires actifs pour la synthèse d'images animées, PhD Thesis, Joseph Fourier University, Grenoble, 2002.