FILTERING OF IMAGE SEQUENCES:

ON LINE EDGE DETECTION AND MOTION RECONSTRUCTION

MATTEO LUCCHETTI

DOTTORATO DI RICERCA IN **INGEGNERIA DEI SISTEMI** – XVII CICLO

DIPARTIMENTO DI INFORMATICA E SISTEMISTICA 'ANTONIO RUBERTI'

UNIVERSITÀ DI ROMA 'LA SAPIENZA'

Coordinatore del Dottorato:	Chiar.mo Prof. Carlo Bruni[1]
Tutore:	Chiar.mo Prof. Carlo Bruni
Docenti esaminatori:	Chiar.mo Prof. Alfredo Germani[2] Chiar.mo Prof. Patrizio Tomei[3] Chiar.mo Prof. Stefano Panzieri[4]

Riassunto.

L'argomento della Tesi riguarda l'elaborazione di sequenze di immagini, relative ad una scena in cui uno o più oggetti (possibilmente deformabili) si muovono e acquisite da un opportuno strumento di misura. A causa del processo di misura, le immagini sono corrotte da un livello di degradazione. Si riporta la formalizzazione matematica dell'insieme delle immagini considerate, dell'insieme dei moti ammissibili e della degradazione introdotta dallo strumento di misura. Ogni immagine della sequenza acquisita ha una relazione con tutte le altre, stabilita dalla legge del moto della scena. L'idea proposta in questa Tesi è quella di sfruttare questa relazione tra le diverse immagini della sequenza per ricostruire grandezze di interesse che caratterizzano la scena.

Nel caso in cui si conosce il moto, l'interesse è quello di ricostruire i contorni dell'immagine iniziale (che poi possono essere propagati attraverso la stessa legge del moto, in modo da ricostruire i contorni della generica immagine appartenente alla sequenza in esame), stimando l'ampiezza e del salto del livello di grigio e la relativa localizzazione.
Nel caso duale si suppone invece di conoscere la disposizione dei contorni nell'immagine iniziale e di avere un modello stocastico che descriva il moto; l'obiettivo è quindi stimare i parametri che caratterizzano tale modello.
Infine, si presentano i risultati dell'applicazione delle due metodologie succitate a dati reali ottenuti in ambito biomedicale da uno strumento denominato pupillometro. Tali risultati sono di elevato interesse nell'ottica di utilizzare il suddetto strumento a fini diagnostici.

[1] Università di Roma 'La Sapienza'
[2] Università de L'Aquila
[3] Università di Roma Tor Vergata
[4] Università di Roma Tre

Università degli Studi di Roma
"La Sapienza"

Dottorato di Ricerca in Ingegneria dei Sistemi
XVII Ciclo

Tesi di Dottorato

Filtering of Image Sequences: on line edge detection and motion reconstruction

Dottorando

Matteo Lucchetti

Tutore

Chiar.mo Prof. Carlo Bruni

Coordinatore

Chiar.mo Prof. Carlo Bruni

Dicembre 2004

A Simona,

mia passione,

mia gioia,

mia speranza.

Ringraziamenti

Al termine del mio percorso di studi sento l'esigenza di esprimere riconoscenza verso coloro che hanno contribuito in questi anni alla mia formazione.
Desidero innanzi tutto ringraziare il Professor Carlo Bruni, vero e proprio punto di riferimento durante tutto il periodo del Dottorato, per aver messo a disposizione la sua esperienza in un clima di fiducia e amicizia e per avermi dato insegnamenti che vanno ben al di là di quello che queste pagine possono racchiudere.
Un vivissimo ringraziamento va anche al Professor Giorgio Koch, per i continui stimoli offerti alla luce delle sue preziose competenze.
Desidero poi ringraziare Daniela Iacoviello, con la quale ho condiviso spunti e riflessioni che hanno dato luogo ad una proficua collaborazione.
Per i dati sperimentali riportati nell'ultimo capitolo di questa Tesi ringrazio l'ing. Giovanni Calcagnini e l'ing. Federica Censi, dell'Istituto Superiore di Sanità.
Un particolare ringraziamento va a tutti i colleghi di stanza che si sono succeduti in questi anni, in particolare Stefania, Claudia, Amit e Luigi, per aver reso il clima di lavoro costruttivo e spensierato al contempo.
Un sentito grazie ai miei due amici storici, Alessandro e Federico, sempre al mio fianco, nella buona e nella cattiva sorte.
Ogni più piccolo passo di avvicinamento a questo traguardo sarebbe stato impraticabile senza il continuo ed incondizionato appoggio della mia famiglia. A mia madre, a mio padre e a mio fratello la mia più profonda gratitudine.
Grazie a Simona, per la passione, la gioia e la speranza che ha saputo trasmettermi e che continua a infondermi.

CONTENTS

Introduction

The problem of analysing a sequence of digital images has grown in popularity since advances in technology made it easier to handle, store and process the huge amount of data coming out from the acquisition devices. Imaged objects can move and change their shape because of possible deformation. Furthermore, the measurement process brings into the available data some level of degradation which has to be accounted for. Both theoretical and applicative relevance of this kind of studies were appealing features for the scientific community and led to an intense investigation in the field during the last two decades.
Mainly three classes of problems have been faced in literature as far as sequence analysis has been concerned and represent to date open research fields. The first one was born as an extension of a well known problem arisen in single image processing, which is the sequence content interpretation, in terms of both contours reconstruction and objects segmentation. The second one, developed in parallel, was focussed on identification and estimation of the motion characteristics of the imaged objects. Third, the joint problem of reconstructing both image content and motion pattern out of a somehow degraded acquisition has been tackled. It's worth to point out that this classification includes also the case in which a single image has to be processed as well as the case of a sequence of images with no motion at all.

From a theoretical point of view all these three classes fall into the general framework of estimation theory. In particular, the general set up of the problem turns out to be non linear. The available measures depend on three variables, the two spatial coordinates and the time. Furthermore, in order to extract significant information from the image, a discontinuous variable is usually considered for estimation. In fact, real information about content of the scene is carried by contours of the present objects, corresponding to discontinuity in image intensity levels. Consequently, such characteristics make the problem of image sequence processing not trivial to solve and much effort has to be devoted to suitably cast the general results in dynamic estimation theory (such as filtering techniques and hypothesis testing).
From applications point of view, a huge amount of examples can be cited, ranging in lots of real world contexts which largely benefited from introduction of results coming from image processing. Just to list the most relevant fields, we can mention space applications, from remote sensing to rover localization, biomedical applications, from diagnostics to vision guided surgery and industrial applications, from visual servoing of robotic arms to automatic inspection.
A review of the currently implemented methodologies both from a theoretical point of view and from the applications one is presented in Chapter 1.

In Chapter 2, we introduce a unifying mathematical formulation of the general problem of sequence processing, when the images are corrupted by some suitably modelled kind of degradations, which are typically introduced by the acquisition devices (blur and additive noise). Admissible motion and deformation for objects in the scene are defined, so as to be parameterized in a finite dimensional dynamical model. From a methodological point of view, this general framework represents a unifying step forward with respect to the existent theoretical achievements, because it allows looking at the three classes above as particular instances of the same set up.

The first problem we deal with (Chapter 3) is edge detection of objects in the sequence, given exact knowledge about the underlying motion pattern (obviously including the cases of a single image and of a sequence characterised by null motion). Such a problem is of key interest, in that it includes all those applications in which the imaged scene doesn't move, while the measurement device does, with a known motion.

We propose a procedure which includes a preliminary pre-processing of the measured image, aimed to localize the detection problem and to improve the signal-to-noise ratio. The proposed pre-processing is based on multi-scale analysis, performed with suitable wavelets. The edge identification turns out to be a linear estimation problem and is consequently accomplished by an algorithm which implements recursive linear quadratic estimation and hypothesis testing. Finally the procedure is tested against simulated and real data. On the basis of two suitable metrics introduced in the previous chapter for the set of edge points, simulation results are evaluated. Effectiveness of the proposed algorithm depends on the capability of exploiting the relationship introduced by the known motion between the different frames. To this purpose, a first experiment with real data is also presented with very promising results. This test put also in evidence an interesting property of robustness of the proposed procedure with respect to possible wrong estimates of image degradation parameters. In fact, the real edges of the image emerge, after few frames being processed, out of an initial very large number of fake edge points, which are in turn filtered out.

The second problem we face (Chapter 4) is motion parameter identification and estimation when the initial uncorrupted image is known. Differently from the previous case, the estimation problem turns out to be non linear, in that the measurement equation is such. Moreover, we have no hypothesis on the probability distribution we have to deal with, so that we can't assure the resulting filtering equations to be finite dimensional. As a matter of fact, to make the problem feasible for numerical solution, we present two different ways to suitably approximate the general formulation.
The first one is based on a Gaussianity hypothesis for the probability distribution of the motion vector conditioned upon the available information, which guarantees finiteness of the resulting filter dimension. The updating step of the resulting filter is achieved by using the Bayes theorem, while the prediction is simply given by the propagation of the dynamical model of the motion. Unfortunately, due to high computational complexity, this method turns out to be unfeasible for the implementation on a personal computer. Consequently, a further investigation is requested in order to introduce possible procedures to make this approach effective.
Alternatively, we introduce a second approach, based on the Extended Kalman Filter. In this case, hypothesis on Gaussianity is preserved, but in addition we suppose to approximate the measurement equation with a linear equation. The resulting filter is the well known general EKF, whose equations are considered for our problem.

In the last chapter of the Thesis (Chapter 5) we present an application of the above procedure of edge detection to a biomedical problem of real applicative interest. Images coming from a pupillometer are considered. A pupillometer is a biomedical acquisition device essentially made up of a camera shooting the human pupil. It has been shown in medical contexts that both the contour of the pupil and its reactions to light variations can be promising diagnostic indexes for several neuro-vegetative diseases. Moreover, the advantage for the pupillometer to be a non invasive device has increased interest around its potentialities.
We start with processing a single image captured by the pupillometer. Preliminary to edge detection is the degradation parameters estimation and to this aim a procedure is presented to reconstruct both blurring and noise characteristic parameters. The above procedure for edge detection is then implemented in the case of a single image, with good results, also due to a suitable fitting of the detected edge points with an elliptic model.
A remarkable improvement is obtained by processing a sequence of pupil images, acquired from the pupillometer over sufficiently short time instants, so as to justify the hypothesis of null motion. In particular, this experiment confirms robustness of results with respect to inaccurate estimates of the parameters which describe degradation introduced by the pupillometer.

1. Main available results in image processing

§ 1.1 – The problem of edge detection and motion reconstruction

Image processing has been stimulating the whole scientific community since the improvement of technology made it easier to acquire data and cheaper to handle them. The increasing interest in this field was motivated by both the challenging theoretical framework in which it can be situated and the relevant impact it presents in a lot of applicative fields.
The first and principal problem dealt with in image processing was that of *edge detection*, in order to extract suitable information for giving an interpretation of the captured scene. The principal physical edges correspond to sharp variations in the reflectance, illumination, orientation and depth of scene surfaces. Since image intensity is often proportional to scene radiance, edges are represented in the image by discontinuous changes in the intensity function, which is usually referred to with terms *grey level*. With the term *edge* we will address in this work the set of points of a 2D image in which the grey level has a step discontinuity, even if other kinds of edge have received attention in literature, such as line edges (local extremes of the grey level) or junctions (features formed in the points where at least two edges meet).
The purpose of edge detection is to localize these variations with an efficient and reliable procedure. The importance of detecting edges arises in all applications involving images acquisition and understanding. For an overview of edge detection methodologies one can see [1]. The problem is significant in that the acquired images are captured by measurement devices such as cameras and they consequently are (more or less sensitively) degraded by some kind of measurement errors. Such errors are typically modelled by additive and/or multiplicative noises and blurring effects due to the device itself, to the objects motion in the scene, and/or to distortions due to the lens optical acquiring system. In the next Chapter we will give a more rigorous definition of the image set we will deal with in this work, with detailed modelling of the degradation effects.

In first studies on image processing, much effort has been devoted to the analysis and edge detection of a single frame, possibly coming out from a single shot acquisition or extracted from a suitable sequence. Nowadays, attention is given to sequences of images as the result of the acquisition step, so that the problem arises of how to deal with the motion of the objects and their consequent possible deformation in time.
With respect to the case of a single frame to be processed, through analysing a sequence of images, more problems can obviously be investigated, which include as a particular case the basic one of detecting edges in a single image.
In fact, information about the motion of the scene can be successfully exploited to account for relationships between the current frame and the previous one (or the following), trying to implement an online filtering procedure to detect discontinuities. On the contrary, in case we don't know any relation between two consecutive frames, processing a sequence of images could contribute to identify the motion parameters. We can therefore consider the following three categories of problems, when processing a sequence of images:

- edge detection when the motion is known, with the aim of reconstructing the grey level function of each pixel or at least the edges of the initial image; a particular instance of this class of problems concerns the case in which a single image (instead of a sequence) is dealt with;
- motion identification, when the initial image (and the relevant edges) is known;
- both motion parameters identification and location and strength of edges estimation for the initial image, when only some statistical properties of disturbances affecting both the motion and the measured images are known.

In the following our focus of attention will be on the first two classes of problems. Both will be investigated in the framework of estimation theory and will be solved by means of online filtering techniques.

§ 1.2 – Edge detection for a single image

As we said, great effort has been devoted in the past to developing methodologies for edge detection of a single image. Sometimes we even find an image sequence considered as a simple collection of independent frames; each image is processed separately, thus ignoring the relationship given by the motion. In this section we want to present an overview of the main edge detectors introduced in literature to analyse a single image.
Most of attention has been devoted to step type edges, which can be defined as the set of points at which the grey level discontinuity occurs in an image. This type of edge generally occurs between two regions having continuously varying grey levels and it can result from various phenomena: for example when one object hides another, or when there is a shadow on a surface, or, as obvious, when objects in the scene have different colours with respect to the background.

The edge detectors can be defined as tools which accept discrete, digitised images as input and return an edge map as output. The edge map of some detectors includes explicit information about the position and strength of edges, their orientation, and the scale. During the history of image processing, a variety of edge detectors have been devised which differ in their applicative ranges (the photometrical and geometrical properties of the edge) and in their mathematical and algorithmic properties.

Since edges are characterized by high values (ideally infinite) of the first order derivative of the grey level, typically the detectors base their algorithms on evaluating such a quantity. Conceptually, all the most commonly proposed schemes for edge detection include three operations: *smoothing, differentiation* and *labelling*. Smoothing consists in reducing noise in the image and regularizing the successive numerical differentiation, in order to ensure robust edge detection. Differentiation consists in evaluating the desired derivatives of the grey level image. Labelling involves localizing edges and increasing the reliability of the result by suppressing false edges. Smoothing and differentiation of an image are usually realized by convoluting the image itself with a suitable smoothing and/or differentiation kernel [2]. We will present the most acknowledged filters introduced in literature in the following of this Section, when speaking of popular edge detectors.
Generally speaking, the implementation of an edge detector which satisfactory performs in many contexts turns out to be not an easy task, since differentiation is an ill-conditioned problem, in the sense that the existence, uniqueness and robustness of a solution cannot be guaranteed in the absence of additional constraints. Moreover, smoothing generally results in a loss of information, so that the ultimate goal we want to pursue is to find optimal detectors which ensure a favourable compromise between loss of information and noise reduction.
Poggio and Torre ([3] and [4]) showed that regularizing differentiation can be achieved by convolution of the image with a sutable cubic spline (or its derivative), with area controlled by a parameter which is called ***scale***. The scale is a parameter characteristic of the convolution kernel, characterizing its smoothing power versus its capability to enhance discontinuities. As a matter of fact, particular attention has to be devoted to the choice of the scale factor. In addition to the cubic spline, two other regularization filters have been proposed in [4], which are the Green function and the Gaussian, or their derivatives.
In particular, the Gaussian n-th derivative is usually referred to with the term ***wavelet*** and its properties have been largely investigated in the context of generic signal processing for

detection of discontinuities and smoothing of noise [5]. In such kernels the scale factor is represented by the variance of the Gaussian (or its covariance matrix). The approach we present in this work to detect edges is based on wavelets too. As we will show, these filters ensure a compromise between noise elimination and the preservation of image structure, even if the problem remains to determine the scale factor.
Several kinds of filtering techniques have been proposed other than the convolution with such kernels. In particular non-linear filters [6] have been proven to be more successful than linear filtering with respect to particular kind of noises (e.g. impulse noise).
The order in which differentiation and smoothing are performed is immaterial in the case of a linear filter, since they are commutative operators, while in the case of non-linear filters smoothing must precede the differentiation operation.

The last operation which has to be accomplished is labelling, that is localizing the edges and increasing the reliability of the output by suppressing false edges. Edges are usually localized by thresholding the gradient modulus of the regularized image, but the resulting output must be skeletized, since the detected edges are not filiform. Last, the elimination of false edges follows. This operation is very important, even though little work has been currently done to accomplish this task. One of the achievements of this work concerns this aspect. As we will see in the third chapter, processing of a sequence of images can be exploited to refine the first edge detection and to strongly reduce false edges.
The rule commonly used to classify edges as true or false depends on a plausibility threshold, which is the minimum acceptable measure for the estimated discontinuity size. Due to fluctuation of the discontinuities estimation, edges resulting from this binary decision are often broken lines. This rule has been improved by Canny [7], who introduced the so-called hysteresis algorithm to take into account edge continuity. Two thresholds are used: a given edge is true if the estimated grey level of every edge point is above a low threshold and at least one is above a high threshold; otherwise the edge is false. This was the first continuous computational model of an edge.
Another aspect of the elimination of false edges concerns threshold computation. Usually, a threshold value is found using a trial-and-error process and is used for all edges of the image. The threshold is a function of the edge characteristics, properties of the smoothing filter, and properties of the differentiation operator. Consequently it is not easy to find a single threshold value for an image. In this work we propose a rule based on a statistical hypothesis test to clean the image, with a confidence level based on the variance of the estimation error of the discontinuity size for each pixel of the image.

As we said above, a theoretical issue generally relevant in edge detection procedures is the choice of the scale for smoothing filter. Having no information about the context in which the detector has to work, the approach typically used to determine the scale consists in trial-and-error experiments, till the right scale is found. Once this parameter is fixed, this scale is applied to the whole image to detect edges. Obviously, this approach based on a single scale factor can lead to unsatisfactory results, and it has been improved by the introduction of *multi-scale analysis*.
As claimed by Rosenfeld and Thurston [9] as well as by Marr and Hildreth [10] we can obtain the description of an image at different scales by applying an edge detector at different scales and combining the recovered edge information. This is usually called multi-scale approach.
The two main requirements for multi-scale edge detection are regularization and nice scale behaviour. About regularization we already spoke before. With the terms nice scale behaviour we intend that edges are not created as the scale increases [11]. Many schemes for multi-scale edge detection have been proposed (see for example [12], [13], [14] and [15]). The use of this approach involves two problems:

(1) the number of scales of an edge detector which can be used for a given image and the way in which those scales are selected;
(2) the way in which the edge information recovered at different scales can be efficiently combined.

Various methods for the automatic estimation of scales have been proposed (see for example [16] and [17]). About the combination of edges detected at different scales, the operation is made easier if the detector has nice properties in the scale space: no edges are created as the scale increases. Theoretically it has been shown [18] that this property is fulfilled if the smoothing is performed through linear convolution between the image and suitable kernels, represented by a Gaussian or its derivatives (wavelets).

Another issue of relevance about edge detectors is the way to evaluate their results. Several authors have proposed performance measures to evaluate the output of edge detectors. Abdou [19] has proposed a measure called the figure of merit, which is a combination of three factors: non-detection of true edges, detection of false edges, and edge delocalization error. Venkatesh and Kitchen [20] extend these kinds of error with the introduction of the one due to detection of several edges instead of an edge one pixel wide. Finally, Kitchen and Rosenfeld's measure [21] combines errors that arise due to an edge's thickness and lack of continuity. For recent results on this topic one can see [22].
In this Thesis, chapter 3, we propose two indexes to evaluate detection performance, the first based on Haussdorff distance between two curves, and the second on the evaluation of number of errors in considering true a false edge and vice versa.

Since the appearance of image processing, the number of edge detectors has increased continuously. It is very difficult to make an inventory of the available algorithms. Most of them include the above mentioned three main steps, but they differ for the choice of the smoothing filter, the differentiation operator used, the labelling process, and even for the goals they aim to.

About thirty years ago the first detectors were proposed, based on approximate evaluation of the first derivative (gradient) and of the second derivative (Laplacian) of the grey level. These detectors are limited to the differentiation operation. Estimation of the gradient vector has been based on the use of masks. These are low-dimensional convolution matrices, and the convolution with the grey level is performed by sliding the kernel over the image, generally starting at the top left corner, so as to move the kernel through all the positions where the kernel fits entirely within the boundaries of the image. If the image has M rows and N columns, being $I_{i,k}$ the grey level of the image in the pixel (i,k), and the kernel has m rows and n columns, being $K_{i,k}$ its generic entry, the output image will have $M-m+1$ rows and $N-n+1$ columns, whose generic element $O_{i,j}$ is given by the convolution

$$O_{i,j} = \sum_{k=1}^{m}\sum_{l=1}^{n} I_{i+k-1,j+l-1} K_{k,l}, \quad i = 1,2,\ldots,M-m+1 \quad j = 1,2,\ldots,N-n+1.$$

The first detector, the Roberts Cross edge detector, used the following 2×2 masks, respectively for the two coordinate directions:

$$K_x = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad \text{and} \qquad K_y = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}.$$

It was suddenly replaced by more clever ones, which used 3×3 masks. In particular, according to [23], the most acknowledged ones are the following:

$$K_x = \begin{bmatrix} -1 & 0 & 1 \\ -a & 0 & a \\ -1 & 0 & 1 \end{bmatrix}, \qquad \text{and} \qquad K_y = \begin{bmatrix} -1 & -a & -1 \\ 0 & 0 & 0 \\ 1 & a & 1 \end{bmatrix},$$

where a is a positive real number, which is equal to 1 in the case of Prewitt's masks [24], to 2 in the case of Sobel's masks, and to $\sqrt{2}$ in the case of Frei-Chen masks. These masks are quite famous in edge detection operations and have been intensively used for a large number of applications.
The Roberts gradient masks are more sensitive to diagonal edges. Prewitt's, Sobel's and Frei-Chen's produce better results than the Roberts' operator, because the mask is larger, and provide averaging of small luminance fluctuations. The Prewitt operator is more sensitive to horizontal and vertical edges than diagonal edges, and the reverse is true for Sobel operator. The Frei-Chen edge detector has the same sensitivity for diagonal, vertical and horizontal edges. The performance of these operators deteriorates due to increasing noise in image, since, for example, both Prewitt's and Sobel's masks are derived by assuming that white noise is additive and image surfaces are linear.

Second derivative operators were also considered. The most frequently used operator is the Laplacian one, and for its estimation the following 3×3 masks have been introduced:

$$K = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad K = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \qquad K = \begin{bmatrix} 1 & 2 & 1 \\ 2 & -4 & 2 \\ 1 & 2 & 1 \end{bmatrix}.$$

Marr and Hildreth [10] proposed the use of zero-crossings of the convolution of the image with the second derivative of a Gaussian. This method is also known with the term LoG.
About the above mentioned problems with noise, Rosenfeld and Thurston [9] were the first to introduce the smoothing operation, in order to reduce the noise of the image. Their first attempt consisted in replacing the value of the pixel by the average computed on a squared window. This method, even if gives better results with respect to the previous ones, remain noise-sensitive, since size and coefficients are fixed and cannot be adapted to the considered specific image. It should be recalled that smoothing has a positive effect, noise reduction, thus ensuring robustness in edge detection, and a negative effect, loss of information (due to discontinuity flattening). Clearly, we have a fundamental trade-off between information loss and noise reduction. The ultimate goal is to find optimal detectors that ensure an acceptable compromise between noise reduction and edge conservation.

Optimal edge identification techniques involve designing edge detectors with the best performance in some sense. There are two problems related to optimal edge detection: the definition of performance criteria and the design of a filter which optimizes these criteria. Usually, the definition of performance depends on detection accuracy and delocalization error of edges. The design of optimal detectors requires the specification of these criteria to give mathematical models, and the use of the optimization theory to derive the optimal detector. Different mathematical models and thus different optimal detectors have been proposed: just as an example, the above mentioned LoG detector belongs to this class, because it has been derived from the solution of an optimization problem.
In spite of reviewing all the detectors in this class we will focus our attention on the most cited and acknowledged one, which is the one proposed by Canny [7].

Canny derived his filter by optimizing a certain performance index that favors true positive, true negative and accurate localization of detected edges. Its optimality is related to three criteria:

- the **detection** criterion accounts for the fact that important edges should not be missed and that there should be no spurious responses;
- the **localization** criterion says that the distance between the actual and located position of the edge should be minimal
- the **one response** criterion minimizes multiple responses to a single edge (also partly covered by the first criterion, since when there are two responses to a single edge one of them should be considered as false).

His first analysis was restricted to linear shift invariant filter that detect unblurred 1D continuous step and the first two optimality criteria. A closed form solution was found using calculus of variations. If the third criterion (multiple responses) is added, the best solution can be approximated by numerical optimization. Canny showed that the resulting filter can be effectively approximated by the first derivative of a suitable Gaussian smoothing filter.
The detector is then generalized to two dimensions. The step edge is given by its position, orientation, and possibly magnitude. In two dimensions, this edge detector uses two kernels representing the derivatives along horizontal and vertical directions. Canny's operator is not unique, as it varies with the edge profile, but for step edges, it can be easily shown that it is similar to the LoG operator because the maxima in the output of a first derivative operator correspond to the zero-crossings in the Laplacian operator used by Marr and Hildreth.
According to [25], the problem with Canny's edge detection is that his algorithm marks a point as an edge if its amplitude is larger than that of its neighbours, without checking that the differences between this point and its neighbours are higher than what is expected for random noise. As a consequence, his technique causes the algorithm to be slightly more sensitive to weak edges, but it also makes it more sensitive to spurious and unstable boundaries whenever there is an insignificant change in intensity (e.g., on smoothly shaded objects and on blurred boundaries).

Basing on Canny's approach, in 1992 Mallat and Zhong [26] published a paper presenting a generalization of this framework. In particular they introduced a numerical technique for the characterization of one- and two-dimensional discrete signals in terms of their multiscale edges. In this work the authors show that performing a multiscale Canny edge detection is equivalent to finding the local maxima of a wavelet transform of the image. Consequently, the authors study the properties of multiscale edges through the wavelet theory. Given that for pattern recognition one often needs to discriminate different types of edges, they show that the evolution of wavelet local maxima across scales characterizes the local shape of irregular structures. Numerical descriptors of edge types are derived and the completeness of a multiscale edge representation is also studied.
Their work inspired many investigators so that multiscale wavelet processing is up to now one of the most studied methodologies. As said, the edge detection method we present in the third chapter follows this line of research too.

Last, it's worth to mention another approach introduced in literature for edge detection, which is based on edges parameterization and optimal fitting.
As regards parametric fitting, these techniques involve adjusting the image by an edge model and selecting the edge pixels for which the fitting error is minimal (see for example [27]). The accuracy of edge fitting is measured in terms of the mean square error criterion. An advantage is that the differentiation problem is thus avoided, but the computational load is very high. Consequently, instead of the parametric model of an edge, a dimensionless edge parametric manifold has been introduced in [28]. The edge parametric model is compared to image pixels. If the distance between the manifold and an image window is sufficiently small, there

is an edge. This scheme has been used for the extraction of five edge models: step, line, roof, corner and blob.
Another technique consists in adjusting the image by a set of given basis functions. In [29] the image is fitted by a linear combination of Tchebychev's polynomial, of order less than or equal to three. First and second-order derivatives are computed and used to locate edges.

§ 1.3 – Different types of processing for a single frame

In this paragraph we want to point out some other types of detection possible for a single frame. Differences with the previous methods lie in the feature to be detected. As we said in section 1.1, the aim of detection could be not to recover just step edges, but junctions or lines. Furthermore, much interest has been devoted in the past to design algorithms able to detect different regions in the image so to understand the basic shapes composing the scene. Such a problem is usually addressed with the term *segmentation*. In the following we will first review some method introduced in literature to detect lines and then we will just sketch the main segmentation techniques.

Lines correspond to local extrema of the grey level image and are of great use in the identification, for example, of roads and rivers in remote sensing images. Most schemes for the detection of lines are limited to thinning algorithms and are designed for binary images. The main problem with such implementations is that they usually yield edges which are not accurately located and they do not perform well in complex images such as remote sensing images.
To detect lines in an image, many researchers made use of the Hough Transform. The Hough Transform [30] was originally intended to be a method of finding straight lines in images, but the methods have evolved to take into account other more complex shapes in an image, such as for examples circles. It is often used both for pattern recognition, and for post-processing other edge detection techniques to reduce false information in the resulting edge image.
The general principle of the Hough Transform is that regular curves in image space, when parameterised, can be represented as a point in parameter space. So, for example, a straight line in an image can be parameterised as $y_i = mx_i + c$ for some m and c and for all points (x_i, y_i) on the line. If we take every pair of edge points from our edge image and calculate the m and c which correspond to a straight line between those points, many more points will map to the same m and c if a straight line is present. So, if we discretize parameter space as a grid, and increment a counter for each time a given value of m and c result from a pair of edge points, the counter in parameter space corresponding to that line will be much larger that other counters. Thus, the problem of finding a straight line in the image can be reduced to that of finding local maxima in parameter space (*searching through voting*). The only drawback in this methodology is that we need to know beforehand what type of curve we are looking for (in order to calculate parameters from our points).
Several approaches other than the Hough transform were proposed. Haralick in [29] extended the approach presented before for edge points detection, using an optimal fitting strategy with Tchebychev's polynomials of order less than or equal to three. Lines occur at pixels having zero-crossings of the first directional derivative taken in the direction that maximizes the second directional derivative.
Using a 1D ideal roof model and Canny's criteria of optimality, Ziou [31] derived an optimal line detector. In 2D the image is convoluted with two directional filters operating in the x direction and in the y direction separately. The resulting images are combined and lines are located in this image at the maxima in the direction that maximize the grey level variance.
Last, we would like to mention Koundinya and Chanda [32], who proposed an algorithm-based combinatorial search. The basic idea behind their methodology is to locate lines that

maximize an *ad hoc* confidence measure. The confidence measure of a candidate pixel is proportional to the number of pixels in its vicinity having a different grey level than the candidate pixel.

Instead of lines and junctions, one could be interested in identifying regions in the image, in order to detect the shape of objects in the imaged scene. As far as segmentation problems are dealt with, we want here to mention four of the most popular approaches: *threshold techniques*, *edge relaxation methods*, *region growing and merging techniques*, and *connectivity-preserving relaxation methods*.

Threshold techniques make decisions based on local pixel information and are effective when the intensity levels of the objects fall squarely outside the range of levels in the background. Because spatial information is ignored, however, blurred region boundaries can create havoc.

Edge relaxation techniques center around contour detection. The method involves producing a confidence measure on how likely every edge detected pixel is to be a true edge as opposed to a false one. The probability associated with each edge point considers a number of factors, including the proximity and direction of adjacent edge pixels, as well as the strength and length of the edge. For example, a weak edge point between two strong edge points is very likely to be a true edge point, while an isolated edge point, even a strong one, is likely to be false [33]. The problem is usually addressed iteratively, with an edge pixel being given a label of yes, no or maybe at each iteration, until there are no maybes left and the segmentation is clear. Gaps in edges can also be filled in by this method, which also evaluates the possibility of unmarked pixels belonging to an edge.
Edge relaxation techniques appear to work well, but are computationally expensive. Also, for several methods [34], non-convergence of confidences to zero or one is possible, resulting in a degradation of results after several iterations. Moreover, weakness in connecting together broken contour lines make them, too, prone to failure in the presence of blurring.

Region growing and merging [35] is the most popular segmentation method for the majority of image classes. The method relies on the homogeneity of an image segment, in color, grey level, texture, or any other criteria.
Methods in this class involve defining some initial method of splitting the image into a large number of small regions, and recursively merging smaller regions into larger regions by some criterion. These regions could then be used as a starting point for the iterative merging of other regions. Among the methods used in this family, there are merging by the nearest neighbor or the k-nearest neighbor rules, other cluster analysis algorithms, boundary melting and splitting and merging. Since region merges are local operations in a region growing process, it has been suggested that many merges could be performed independently at the same time, so that the overall computing process results to be faster [36].

A *connectivity-preserving relaxation-based* segmentation method, usually referred to as the *active contour model*, was proposed recently.
An Active Contour [37] is an energy-minimizing spline that detects specified features in an image. It consists of a set of *control points* connected by straight lines. The Active Contour is defined by the number of control points as well as the coordinates of each control point. It is held together by internal forces and is guided toward image features, such as an object's boundary, by external forces. The energy function that describes Active Contours is composed of two components, the internal energy and the external energy.
The internal energy deals with intrinsic properties of the contour and is a smoothness constraint that keeps the points contained in the contour; in particular, it is usually made up of

two terms, which account for the elastic energy of the contour (shrinking or expansion) and the bending energy (in order to maintain the contour as smooth as possible).
The external energy forces the contour to behave as we desire, that is to respond to image features. Since an edge is defined where a large grayscale gradient exists, this energy should decrease as the gradient increases. Nevertheless, the definition of the external energy depends on the type of application we consider, so that in literature many other definitions can be found [38]. Although the energy-minimizing model is not new, coupling it with the maintenance of an "elastic" contour model gives it an interesting new twist. As usual with such methods, getting trapped into a local minimum is a risk against which one must guard and this could represent a hard task.

§ 1.4 – A review of image sequence analysis

An issue of widely acknowledged importance as well as of large applicative interest is the one of processing a sequence of blurred noisy images from one or more objects undergoing some (possibly non rigid) motion. Within this context, the problems naturally arise of identification of the objects in the scene and/or of their motion as well.
The issue of processing sequences of images corresponding to moving objects is dealt with in a very large and rich literature. A comprehensive report for the work up to 1981 is given in [39]. In recent times, technological development and spreading of applications led to branching of scientific work in various research lines.

An area which received considerable attention is the issue of estimating motion from image sequences; for a survey one can see [40]. Relative motion of a viewing system and the environment it is observing causes *image motion*, that is the motion of the projected environment on the imaging surface of the viewing system. The velocity vector which can be associated to the projection of each visible environmental point is expression of the so-called *optical flow*. Great effort has been devoted in recent times to the estimation of the optical flow from a sequence of images. Estimation is done through the resolution of a set of constraints derived from hypotheses made on the nature of the observed brightness pattern and on the optical flow. For instance, the basic assumption made, in most of the works in this field, is represented by the gradient equation of Horn and Schunck [41], which is a constraint written under the hypothesis that image brightness is invariant along image motion trajectories; this constraint is usually supplemented with a regularization constraint derived from the hypothesis that optical flow is a smooth vector field. Of course, since we are dealing with an estimation problem, we obtain an approximation to the optical flow that will depend on the constraints used and the mode of resolution of these constraints.

According to the classification reported in [42], analysis of image motion can be divided into four fields of interest:
- *image motion detection*;
- *image motion estimation*;
- *structure identification from motion*;
- *deformable objects motion analysis*.

The **motion detection** problem is particularly pertinent when the viewing system is static and the goal is to locate mobile objects in the observed scene. A binary map is then sought, indicating motion or absence of motion at every point in the image positional array. In earlier formulations, temporal change detection was primarily stated as a simple pixel-based frame difference evaluation, followed by thresholding [43]. Further improvement in modeling has led to statistical frameworks, in which statistical tests have been implemented [44].

Nonetheless, motion detection should not be reduced to temporal change detection. Of greater practical interest is the recovery at each instant of the image regions corresponding to moving objects (the so-called *moving objects masks*). A temporal change caused by a moving object can give rise to three types of regions in the image, corresponding respectively to the uncovered background, the covered background and the overlap of the two successive object images. Earlier methods have proceeded by extracting successive temporal change maps, by means of heuristics referring to the intensity images. In Wicklund and Granlund [45], three successive images, at instants t_1, t_2 and t_3, are considered. Two binary temporal change maps, between t_1 and t_2, and between t_2 and t_3, are then determined, on which a logical AND is performed. This means information from 'past, present, and future' is necessary to retrieve object masks at a given time.
In a quite different class of approaches a reference image of the static background is taken into account [46]. In this way, maps of temporal change coincide with the maps of moving object masks. However, the reference image must be continuously updated over time, due to possible changes in the background and in the illumination conditions.
More recent investigations have been concerned with the development of motion detection methods which are robust to perturbations like background temporal clutter, and to moderate amounts of viewing system motion interference [47].

As far as **motion estimation** is concerned, two distinct approaches have been developed.
The first of these is based on extracting a set of relatively sparse, but highly discriminatory, two-dimensional features in each image of the sequence, corresponding to three dimensional object features in the scene, such as corners, boundaries and so on. Inter-frame correspondence is then established between these features. Constraints are formulated based on assumptions such as for instance rigid body motion. Their solution is obtained using the observed displacement of the 2-D image features, leading to the computation of motion parameters of objects in the scene. Both parallel projection and perspective projection image model (see for example [48] and [49] respectively) were used to estimate the motion of an object undergoing rigid translation and rotation. Implicit in all the above discussion is the assumption that the scene contains a single rigid object, even if examples can be cited with more than one rigid object.
The other approach is based on computing the optical flow of the two-dimensional field of instantaneous velocities of brightness values (gray levels) in the image plane. The optical flow is then used in conjunction with added regularity constraints about the optic flow or information regarding the scene to compute the actual three-dimensional relative velocities between scene objects and camera. To this purpose, popular assumptions [41] made on optical flow include one of the following:

- optical flow is smooth and neighboring points have similar velocities;
- optical flow is constant over an entire segment of the image
- optical flow is the result of restricted motion (for example, planar motion).

If such assumptions are violated, e.g. when an object translates in front of a stationary background, there exist boundaries where local smoothness of optic flow will not hold. If the boundaries can be detected then the technique may be limited to smooth regions. In [50] some techniques for determining such boundaries are reported. Further developments in image motion estimation can be found in [51].

The goal of **motion-based segmentation** is to partition the image into regions that have different motion characteristics or properties (usually referred to as *motion regions*). A motion segmented image is useful as it allows subsequent kinematical analysis to be performed selectively within each of its segments. Approaches in this field have been classified into two classes: *border placement schemes*, which look for edges occurring between regions with a

significant difference in some motion characteristic or property (see for instance [52], [53] and [54]), and *region extraction schemes*, which look for maximal regions satisfying some motion-related homogeneity criterion ([55], [56], [57]).
A task that is closely related to the estimation of motion is the task of structure estimation of the imaged scene (the so called **structure-from-motion** task). In the case of the optical flow method, this consists of grouping pixels corresponding to distinct objects into separate regions, i.e. segmenting the optical flow map, and then computing the three dimensional coordinates of surface points in the scene corresponding to each pixel in the image at which the flow is computed. In the case of the feature-based analysis, computing structure corresponds to forming groups of image features for each object in the scene and then computing the 3-D coordinates of each object feature associated with each image feature. Although structure can be computed independent of motion, knowledge of motion parameters for features/regions of interest can aid segmentation of image features/regions corresponding to distinct objects and knowledge of object motion can facilitate establishment of feature correspondence within a pair of stereo images, thus aiding the determination of structure. For reviews of the main achievements in this research line, one can see [51] and [58].

Most of the previously classified works deal with rigid bodies' motion. The study of **motion of deformable objects** has been considered too in a wide bibliography, being motivated by applications in domains such as health care, environment monitoring and so on. There are three broad classes of motion of deformable objects, pertaining to articulated bodies, elastic bodies and fluids. Processing images in the presence of deformation increases difficulties because motion can be complex and object shape can be fuzzy. Representation of shape and of its deformable motion is obviously the crucial matter. Both the two can be considered separately or jointly.
A great deal of attention has been devoted to techniques based on active contours (*snakes*) or deformable surfaces. These are conceptually attractive methods, and able to integrate several kinds of requirements, such as adequacy to data and conformity to physical properties, even though they suffer from shortcomings. In fact, initialization must be chosen close to the desired solution; as a result, the initial approximation is usually supplied by hand, or exploits strong a priori knowledge related to the considered applications. The methods are also affected by the proximity of other contours and the presence of textured image regions.
An alternative method consists of using 2D parametric templates, involving a compact global parameterization to represent shapes of interest [59]. As obvious, such methods are strongly problem-specific, as a priori information about the shape is used.

Another research line, which extends the aim pursued by the structure-from-motion task concerns the **image reconstruction** (and in particular edges reconstruction) problem given full information about the motion.
The problem of a single image reconstruction can be reduced to the one of signal restoration, that is linear deblurring and noise suppression, and it constitutes a classic chapter in image processing (see previous paragraph). This can be generalized to the restoration of an image sequence, where the basic assumption is that the sequence contains one filmed scene (i.e. it's a continuous sequence). Here the applicative relevance of the problem stems from the opportunity of processing several images of the same object to increase the available information. Despite its importance, this restoration problem has not been as extensively treated as the single image counterpart problem. This can be explained by the amount of computations and memory required when applying even the simplest restoration algorithm to such signals. The approaches proposed in literature consequently focused to make the problem easier, by assuming stationarity, causality in the image plane and locality of the filter. Among these algorithms are the 3-D local Kalman filter [60], 3-D median filter [61] and motion-compensated simple averaging in the time axis [62].

In [63] in addition, the issue of superresolution is dealt with, namely the issue of reconstructing images from observed decimated versions of them. Here the richness of the concerned signal is exploited to combine different frames with each other, in order to create an enhanced resolution output image.

§ 1.5 – Overview of the main applicative fields

Since its birth, Image Processing has always been considered an ideal bridge between theoretical results in computer and systems science and practical problems in several applicative fields. This is due to the most important aim pursued, the one of controlling an artificial vision system by a computer, extracting the entire visual information one can think necessary for a particular goal and feeding with them the successive processing step.
Consequently, a huge amount of applications have been developed in past years, and it is not an easy task to try to catalogue them. In order to give an overview, we will refer to a classification of the main applicative fields, which is the one that can be found on the website of CVonline (http://homepages.inf.ed.ac.uk/rbf/CVonline/applic.htm). This last is by far the most widely acknowledged catalogue of bibliographic references for investigators in Computer Vision.
It is worth pointing out that, among all the possible applications implemented in literature, the most largely investigated fields so far have been four: human, industrial, medical and remote sensing. However, many other disciplines received profit from the introduction of image analysis techniques. For each topic we are going to list, we will briefly sketch the most significant implications of edge detection and we will survey the main achieved results for a selected group of applications.

Agricultural.
In agricultural applications we find several examples of the use of edge detection techniques. Maybe the most interesting ones are linked to robotic vision and guidance, aimed to online *crop treatment*. An example can be found in [64] where image analysis is used to automatically select and remove weed by precision robotic application, in order to reduce or even eliminate chemicals used in farming. Another class of applications concerns the one of *fruit grading*, to classify the production. In [65] we find application to lentils grading, based on their dimension and average colour. Other remarkable applications in this field, which are worth to be mentioned, dealt with *plant health monitoring and parasite identification*, *pollen identification*, *yield estimation* and *tree counting*.

Animal Husbandry.
Mainly two kind of problems have been faced by means of image analysis and edge detection techniques: *fish movement modelling and tracking*, and *flock tracking*. About the first, in [66] a vision system to monitor fish eggs abundance in particular regions of the ocean while staying aboard of a ship is presented. This system, called REFLICS, uses a segmentation algorithm to locate fish-egg-like objects in the image and then a classifier to determine fish egg, species and development stage (age).

Character Recognition.
Both printed and handwritten characters have been analysed for the possibility of automatic recognition. Many different alphabets have been considered as well, such as Arabic, Chinese, Japanese, Korean and obviously Roman. In each application, once the character is detected, there has to be an interpretation phase, during which the detected edge is recognized as a character of the undertaken alphabet. Such a classification is usually done by means of neural networks. For an interesting survey of the achievements in this field one can see [67].

Commercial Applications.
This class of applications has received attention in recent times, because of the great concern about *signature identification* and/or *verification* [68]. Within this area we also want to mention the attempts to build systems for automatic processing of handwritten *bank checks*, or for *bar-code reading*.

Human Body Analysis.
This is definitely one of the most important fields of application for edge detection methodologies. Lots of examples can be cited; we will classify the implemented applications on the basis of the feature they address.
Body. Human body is analysed to recover useful information about both its behaviour and the model of its movement; person counting, as well as walker identification have been taken into account for surveillance aims, as we will see in the following; furthermore, a large number of investigators focused on the possibility of performing motion tracking for gait identification and posture analysis (see for example [69]).
Hands. Hands too have been intensively dealt with in image processing, either in order to analyse gesture, or to recognise hand sign [70], or to verify the palm print [71].
Heads and Faces. Nevertheless, the most studied parts of the human body in computer vision are represented by heads and faces. Several applications can be thought of, such as authentication and identification (from frontal identification of faces, to iris identification), or expression understanding. An interesting problem faced is that of gaze location and tracking for optical control of a PC mouse pointer [72]. Moreover, applications for lips shape detection, tracking and reading have been developed.

Image Enhancement.
In facing this problem, edge detection and image segmentation have provided useful tools to restore either noisy or corrupted frames. Such reconstruction has been successfully exploited in old movies digitisation and in super-resolution applications.

Industrial Applications.
As we already said, this is one of the fields which image processing has entered more massively. Problems dealt with in industry range from *factory automation* and *process control*, to *part pose estimation* and *part recognition*, from *inspection* to *manufacturing*. We just want to cite some example of applications in industry: particle counting, check of printed circuit boards, as well as of solder joints, visual servoing for robotic arms [73] and paint spraying [74].

Medical (Veterinary and Human) Applications.
Maybe the area most benefited by the introduction and the development of image processing techniques is the one concerning biomedical applications. Within this subject there has been a real step forward in all those tests which involved extracting and understanding some feature from one or more biomedical images. These images can be classified on the basis of the devices which produced them.
CCD-based. These are all those acquisitions made by an external device, such as a CCD camera. In this area we can mention skin analysis for cancer detection [75], or retinal images segmentation. We will present in last section an application of our algorithm to the analysis of an image sequence of a human pupil shot by a CCD-based device called pupilometer, aimed to the non invasive diagnosis of neuro-vegetative diseases.
Microscope-based. Interest has been focussed on cell nucleus segmentation, as well as on chromosome analysis. Lymphoma analysis has been presented too, with promising results.

Scanner-based. Many different kinds of image are available from various scanner devices and as a consequence a large number of application has been presented, ranging from aneurysm detection [76] to arterial tree segmentation, from cardiac shape extraction and tracking to cartilage analysis. Much effort has been also dedicated to tumours detection as well as to the design of real-time surgical support systems.
Ultrasound-based. Both cardiac monitoring and fetal monitoring [77] have been effectively accomplished analysing images coming from ultrasound acquisitions.
Xray-based. Here we find interesting achievements in the study of angiography and of bone densitometry [78], in addition to breast scan (mammogram) for general tumour detection.

Military Applications.
As far as military applications are concerned, methods of image analysis have been used mainly for moving target identification and tracking. Consequently, we find either aircraft, or missile, or vehicle, effectively detected, identified and tracked. For an interesting example one can see [79].

Navigation, Robot and Vehicle Control.
The principal aim for researchers in this field has been automatic guidance of both vehicles and robots. Consequently, either vehicle control or trajectory correction have been undertaken. A very hard task that arose in all these applications has been that of (possibly moving) obstacles detection and avoidance. In particular, in robotics, a new line of research has started, which is called *visual servoing*, and is now considered one of the most efficient techniques for robot localization and moving in an unstructured environment. For an accurate overview of this field one can see [80].

Remote Sensing and Space Applications.
Remote sensing is one of the applications which image processing seems to be best suited for. To such purposes as airport analysis and landing strip identification, clouds identification and tracking [81], road detection, extraction and matching, identification of diamonds mines, sea/ocean wave analysis, terrain analysis and understanding and so forth, all methodologies for contour extraction gave a significant impulse to developing affordable and efficient solutions. It is worth to mention here the field of space exploration, where image processing techniques were massively involved. To recall some applications we want to cite here planning of the path for all the planetary rovers implemented up to now (PathFinder, Opportunity, etc.), as well as the detection and mapping of the shape of the Moon's craters.

Security and Surveillance.
In the previous subsection about human applications of image processing, we have already spoken about intruder detection and tracking. Other security and surveillance topics addressed in literature concern anomalous behaviour detection [82], pedestrian and car park surveillance, metal detection and background labelling.

Traffic and Transport.
Last, we want to refer to subjects regarding traffic and transport. As we already saw in previous topics, image processing has been mainly used for monitoring aims (from airport to seaport, from vehicles to pedestrians). Interest has been also addressed to driver alertness monitoring for online alarming [83], as well as to traffic analysis and vehicle detection for intelligent traffic lights support. Furthermore, in order to accomplish the automatic vehicle guidance, road sign identification and road structure analysis have been considered, with promising results.

2. The set of images with motion and deformation

§ 2.1 – The image set

We want here to provide the mathematical definition of the image set we will consider in this work. This issue has been developed in [84] and we will present in this section the most relevant results therewith established.
In the sequel, we shall use the following notations: I_A is the indicator function of the set A, i.e.,

$$I_A(x) = \begin{cases} 0, & x \notin A, \\ 1, & x \in A; \end{cases}$$

∂A denotes the boundary of the set A, $\mathring{A}$ denotes its interior, while $\overline{A}$ denotes its closure.
We first introduce the family $\mathcal{D}_0$ of 2-D functions f suitable to model the grey level of the set of images we are interested in.

Definition 2.1 – The set $\mathcal{D}_0$ contains all the 2-D, possibly discontinuous functions $f : D \to [0, E]$ which admit a representation

$$f = \sum_{k=1}^{N_f} \gamma_k^f I_{A_k^f}, \quad N_f < \infty \tag{2.1}$$

where E is a compact connected subset of $\mathbb{R}$; D is a compact connected subset of $\mathbb{R}^2$; $\left\{A_k^f, k = 1,2,\ldots,N_f\right\}$ is a finite partition of D such that the atoms A_k^f are connected, with $\overline{A_k^f} = \overline{\mathring{A}_k^f}$ and the boundary ∂A_k^f of A_k^f has zero 2-D Lebesgue measure, $k = 1,2,...,N_f$; each function γ_k^f is continuous on D and its modulus of continuity is uniformly bounded over $\mathcal{D}_0$ by $M < \infty$. □

When considering images, $f(x)$ obviously represents the grey level of the image at point $x \in D$.
For any $f \in \mathcal{D}_0$, we denote by C_f its discontinuity set. Clearly, we have

$$C_f \subset S_f = \bigcup_{k=1}^{N_f} \partial A_k^f .$$

On $\mathcal{D}_0$ we consider the topology induced by an L_2 – norm so that two images f, g will be assumed to coincide whenever $f(x) = g(x)$ a.e. in D.
In order to guarantee the existence of a solution for an optimal reconstruction problem we need to stay within a suitable compact subset $\mathcal{D}_0'$ of $\mathcal{D}_0$. The problem of compactness for subsets of possibly discontinuous 2-D functions has been dealt with in [85]. We outline here one of the results there reported.
Here and in the following, μ denotes the Lebesgue measure of suitable dimension.

Definition 2.2 – Let I be a compact interval in $\mathbb{R}$. We define by Φ a set of right continuous functions $\phi : I \to D$ which admit a representation

$$\phi = \sum_{k=1}^{N_\phi} \phi_k I_{A_k^\phi}, \quad N_\phi < \infty,$$

where $\{A_k^\phi, k = 1,2,\ldots,N_\phi\}$ is a finite partition of I; the A_k^ϕ's are intervals with uniformly positive Lebesgue measure $\mu(A_k^\phi)$; and ϕ_k's are uniformly differentiable over I. □

We now consider a curve C, belonging to the family $\mathscr{C} = \Phi(I)$, defined over D and inducing a partition $\{A_k, k = 1,2,\ldots,N\}$ in it. Let us introduce a particular subset $\mathscr{C}_\psi$ of this family, which will be useful in the following.

Definition 2.3 – For a given function $\psi : \mathbb{R}^+ \to \mathbb{R}^+$, continuous and nondecreasing, we define

$$\mathscr{C}_\psi = \{C \in \mathscr{C} : \forall x \in C, \mu(B_i) \geq \psi(r), r > 0\}$$

where B_i is any atom of the partition of $\{y \in D : \|x - y\| \leq r\}$ induced by C. □

Note that curves in $\mathscr{C} = \Phi(I)$ are generally regular curves, with a finite number of multiple points; adding the extra constraint of Definition 2.3, we prevent the set of curves $\mathscr{C}_\psi$ from having too sharp cusp (i.e., cusp narrower than some uniform limit).
Given these definitions, the following theorem holds [85].

Theorem 2.4 – For a given ψ, let $\mathscr{D}_0'$ be defined as

$$\mathscr{D}_0' = \{f \in \mathscr{D}_0 : C_f \in \mathscr{C}_\psi\}.$$

Then, $\mathscr{D}_0'$ is compact. □

Unfortunately, $\mathscr{D}_0'$ as defined in theorem 2.4 is not convex: this follows from the fact that union of two admissible sets (in $\mathscr{C}_\psi$) need not be in $\mathscr{C}_\psi$. Nonconvexity of $\mathscr{D}_0'$ is a severe drawback in that it prevents us from guaranteeing uniqueness of the solution for the above mentioned optimal reconstruction problem. This in turn might imply nonrobustness of the solution against small perturbations of the measured data.
To overcome this problem, one might search for a solution within a convex closed subset of $\mathscr{D}_0$. One instance might be a family of $f \in \mathscr{D}_0$ with all the C_f's belonging to a given admissible discontinuity set sufficiently comprehensive (such as a sufficiently fine lattice). To that purpose, the following result holds [84].

Theorem 2.5 – Let T be a curve in D inducing a partition of D, as the one of Definition 2.1, and define

$$\mathscr{D}_0'' = \{f \in \mathscr{D}_0 : C_f \subset T\}.$$

Then, $\mathscr{D}_0''$ is a convex compact subset of $\mathscr{D}_0$. □

It's worth to point out that $\mathbb{D}_0''$ coincides with the subset of functions $f \in \mathbb{D}_0$ which admit a representation 2.1 such that $S_f = T$. The assumption of this theorem appears to be not so reductive if one recasts the reconstruction problem in a setup ready for feasible algorithms. As we will notice in section 2.3, compactness and convexity of $\mathbb{D}_0''$ imply well posedness of the image reconstruction problem, i.e. existence and uniqueness of its solution, when it is formulated as a minimum distance problem from the available data or more generally as a minimization problem for a suitably defined continuous and strictly convex functional.
Following these results, in Chapter 3 we will present a possible solution to the problem of edge reconstruction for a sequence of images, where, in order to consider a compact and convex set of images, we will assume that discontinuities could only be located on a suitable lattice.

§ 2.2 – Possible metrics for the discontinuity set

In order to evaluate the performance of the edge reconstruction procedure we will present in the next chapter, it is useful to introduce two suitable metrics. To do that, we first need to choose a suitable subset $\mathbb{D}_0'''$ of $\mathbb{D}_0$, satisfying conditions of Theorem 2.5.

Definition 2.6 $\mathbb{D}_0'''$ is the set of all $f \in \mathbb{D}_0$ such that C_f belongs to a grid G defined by a given Δ-size lattice L, where Δ is a positive real number:

$$L = \{(x_1, x_2) \in D,\ x_1 = l\Delta,\ x_2 = m\Delta,\ l, m = 0,1,2,.....\}$$

C_f is the union of intervals whose lengths admit a uniform positive lower bound and ∂D belongs to the grid itself. □

Given this definition, it is worth to point out that, due to compactness of D, the number of segments which constitute C_f is finite, so that we can relate C_f with Γ_f, defined as the set of the middle points of such intervals. Consequently, Γ_f suffices to identify the edges of the image. As a matter of fact, this will be the real output of the edge detection algorithm. When f varies in $\mathbb{D}_0'''$, $card\{\Gamma_f\}$ can vary between 0 and the maximum number of pixels of the image. We now call $\mathcal{G}$ the set of all possible Γ_f's associated with images in $\mathbb{D}_0'''$. The assessment of the quality of edge detection can be thus obtained by introducing suitable metrics on $\mathcal{G}$. In particular, we considered two different metrics.
The first one accounts for the maximum distance between two different sets of edge points Γ_f and is provided by the Hausdorff distance normalized with respect to its maximum value. Denoting by Γ, Γ' respectively the sets of original and reconstructed edge points, and by N^2 the number of points in the image, the normalized Hausdorff distance $d(\Gamma, \Gamma')$ is defined as:

$$d(\Gamma, \Gamma') = \frac{1}{\sqrt{2}N} \max\left\{\max_{x \in \Gamma} \min_{x' \in \Gamma'} \|x - x'\|, \max_{x' \in \Gamma'} \min_{x \in \Gamma} \|x - x'\|\right\}. \tag{2.2}$$

Even though it is a very reliable method to measure distance between two different sets of points, the main drawback of such a metric lies in that it is not sensitive with respect to the number of misplaced edge points. To account for that, we also consider the normalized number of misplaced edge points $m(\Gamma,\Gamma')$, defined as:

$$m(\Gamma,\Gamma') = \frac{1}{N^2}\left(card\{x \in \Gamma : x \notin \Gamma'\} + card\{x' \in \Gamma' : x' \notin \Gamma\}\right). \tag{2.3}$$

This quantity tackles just the number of errors in considering false edges as true and vice versa.
Both the two metrics are significant and will be used to evaluate the performance of the edge detection algorithm.

§ 2.4 – Modelling the motion

As we were mentioning in the introduction, we are interested in processing time sequences of images reproducing a moving scenario. Thus, let us introduce a motion operator.
Let $\varphi(t,x) \in \mathbb{R}^2$ denote the position at time $t \in [t_0,t_f]$ of a point which was in $x \in \mathbb{R}^2$ at time $t = t_0$. The position is evaluated with respect to an arbitrary fixed point $\bar{x}$ which acts as (and as a matter of fact is identified with) the origin of the plane.

Definition 2.7 – A "regular" motion operator is a function $\varphi : [t_0,t_f] \times \mathbb{R}^2 \to \mathbb{R}^2$ such that:

a) $\varphi(\cdot,x)$ is continuous on $[t_0,t_f]$, $\forall x \in \mathbb{R}^2$

b) $\varphi(t,\cdot)$ is a bijective C^1 function on $\mathbb{R}^2$, $\forall t \in [t_0,t_f]$ □

For any C^1 curve passing through x, the rotation of tangent vector in x due to φ depends on x itself.
For each fixed $t \in [t_0,t_f]$, $\varphi(t,\cdot)$ admits an inverse in $\mathbb{R}^2$ which will be denoted by $\varphi_t^{-1}(\cdot)$.
Thus at time t the initial image f is transformed into:

$$f_t(x) = f\left(\varphi_t^{-1}(x)\right) = \left(f \cdot \varphi_t^{-1}\right)(x) \tag{2.4}$$

As a matter of fact, Definition 2.7 corresponds to a general continuous topological motion. In special cases, such as a rigid motion or elastic deformation, φ admits a finite dimension parameterisation:

$$\varphi(t,x) = \psi(\lambda(t),x)$$

where $\lambda : [t_0,t_f] \to \mathbb{R}^p$ is a suitable continuous function and $\psi(\lambda(t),\cdot)$ is bijective.
For instance, in the rigid motion $\lambda(t)$ has three components: the rotational angle $\lambda_1(t)$ of the moving object and the shift $(\lambda_2(t) \quad \lambda_3(t))^T$:

$$\varphi(t,x) = R(t)x + T(t) \tag{2.5}$$

where $R(t)$ accounts for rotation effect:

$$R(t) = \exp\left\{\begin{pmatrix} 0 & -\lambda_1(t) \\ \lambda_1(t) & 0 \end{pmatrix}\right\}, \tag{2.6}$$

and $T(t) = (\lambda_2(t) \quad \lambda_3(t))^T$ accounts for the shift.

More in general, for $\varphi(t,\cdot)$ defined by any affine transformation, (so that φ_t^{-1} is affine as well) the following representation holds [86]:

$$\varphi(t,x) = R(t)P(t)x + T(t), \tag{2.7}$$

where $P(t)$ accounts for a deformation described by the positive definite symmetric matrix:

$$P(t) = \begin{pmatrix} \lambda_4(t) & \lambda_5(t) \\ \lambda_5(t) & \lambda_6(t) \end{pmatrix}, \tag{2.8}$$

and all entries λ_i, $i = 1,2,...,6$ are continuous real functions on $[t_0, t_f]$.

Note that, according to the Helmholtz's fundamental theorem of kinematics [87], an affine transformation such as (2.7) may always be assumed for a motion operator φ as given in Definition 2.7, provided we limit its validity to a sufficiently small element of a deformable body [88].
The representation (2.7) is equivalent to the representation

$$\varphi(t,x) = P'(t)R'(t)x + T(t),$$

where R' and P' are matrices of the same type of R and P, but not coincident with them.
The "multiplicative" representation (2.7) for an affine transformation φ is equivalent to an "additive" representation:

$$\varphi(t,x) = R''(t)x + P''(t)x + T(t), \tag{2.9}$$

where $R''(t)$ and $P''(t)$ are respectively the (uniquely defined) antisymmetric and symmetric components of the matrix $R(t)P(t)$. In turn, the representation (2.9) may be put (not uniquely) in the form:

$$\varphi(t,x) = R'''(t)x + P'''(t)x + T(t),$$

where $P'''(t)$ is symmetric and $R'''(t)$ is proportional to a rotation matrix.

Note that the affine character of $\varphi(t,\cdot)$ allows to add to rigid motion a deformation which is such that any two parallel vectors keep being parallel and their moduli are incremented by the same factor. Therefore, under $\varphi(t,\cdot)$, the two original coordinate directions identified by the two unity length vectors $(1 \quad 0)^T$ and $(0 \quad 1)^T$ are transformed into two new vectors $k_1(t)$ and

$k_2(t)$, with lengths $\Delta_1(t)$ and $\Delta_2(t)$ respectively, and directions (unity length vectors) $\kappa_1(t)$ and $\kappa_2(t)$. It is possible to check that:

$$\Delta_1(t)=\sqrt{\lambda_4^2(t)+\lambda_5^2(t)},\qquad \kappa_1(t)=\frac{1}{\Delta_1(t)}\begin{bmatrix}\lambda_4(t)\cos\lambda_1(t)+\lambda_5(t)\sin\lambda_1(t)\\ \lambda_5(t)\cos\lambda_1(t)-\lambda_4(t)\sin\lambda_1(t)\end{bmatrix} \tag{2.10}$$

$$\Delta_2(t)=\sqrt{\lambda_5^2(t)+\lambda_6^2(t)},\qquad \kappa_2(t)=\frac{1}{\Delta_2(t)}\begin{bmatrix}\lambda_5(t)\cos\lambda_1(t)+\lambda_6(t)\sin\lambda_1(t)\\ \lambda_6(t)\cos\lambda_1(t)-\lambda_5(t)\sin\lambda_1(t)\end{bmatrix} \tag{2.11}$$

Correspondingly, the angle between the two directions, originally equal to $\frac{\pi}{2}$, becomes:

$$\vartheta(t)=\cos^{-1}\left[\frac{\lambda_4(t)\lambda_5(t)+\lambda_5(t)\lambda_6(t)}{\sqrt{\left(\lambda_4^2(t)+\lambda_5^2(t)\right)\left(\lambda_5^2(t)+\lambda_6^2(t)\right)}}\right],\qquad \vartheta(t)\in(0,\pi) \tag{2.12}$$

Moreover, because of the definite positiveness property of $P(t)$, parameters $\lambda_4(t),\lambda_5(t),\lambda_6(t)$ have to comply the constraints:

$$\lambda_4(t)>0,\qquad \lambda_4(t)\lambda_6(t)-\lambda_5^2(t)>0\,. \tag{2.13}$$

Boundedness of $\lambda_i(t)$, $i=1,2,\ldots,6$ and definite positiveness of $P(t)$ may be strengthened in the sense of uniformity w.r.t. t if one wishes the image set $\mathcal{D}_t=\{f_t, t\in(t_0,t_f]\}$, with f_t given by (2.4) to be a compact set, provided that $\mathcal{D}_0$ is such. In fact, from (2.13) one easily gets:

$$\sin^2\vartheta(t)=\frac{\left(\lambda_4(t)\lambda_6(t)-\lambda_5^2(t)\right)^2}{\left(\lambda_4^2(t)+\lambda_5^2(t)\right)\left(\lambda_5^2(t)+\lambda_6^2(t)\right)}$$

which, under the assumed uniformity, is uniformly positive. This in turn guarantees, as stated in the previous section (cfr. Theorem 2.4), compactness of $\mathcal{D}_t$ [85].

In view of the formulation of an estimation problem for the motion parameters, it is convenient to introduce the unconstrained parameters vector $\gamma(t)=(\gamma_1(t)\quad\gamma_2(t)\quad\gamma_3(t))^T$:

$$\begin{aligned}\gamma_1(t)&=\ln\lambda_4(t)\\ \gamma_2(t)&=\lambda_5(t)\\ \gamma_3(t)&=\ln\left(\lambda_4(t)\lambda_6(t)-\lambda_5^2(t)\right)\end{aligned} \tag{2.14}$$

with the inverse transformation:

$$\lambda_4(t) = \exp(\gamma_1(t))$$

$$\lambda_5(t) = \gamma_2(t) \tag{2.15}$$

$$\lambda_6(t) = \left[\exp(\gamma_3(t)) + \gamma_2^2(t)\right]\exp(-\gamma_1(t))$$

Obviously, a finite parameterisation of φ allows its modelling in terms of a (finite dimensional, possibly stochastic) dynamical system, describing the time evolution of the parameter λ.
As it is usually done in a Newtonian framework, we introduce the rotation and translation velocities vector $\mu(t) = (\mu_1(t) \quad \mu_2(t) \quad \mu_3(t))^T$ as time derivatives of (respectively) the position parameters vector $\lambda(t) = (\lambda_1(t) \quad \lambda_2(t) \quad \lambda_3(t))^T$:

$$d\lambda(t) = \mu(t)dt\,. \tag{2.16}$$

We now consider a simple stochastic dynamical model for the motion parameters, which assumes μ and γ as (vector) diffusion processes. More specifically, we assume the motion to be modeled by the stochastic differential equations:

$$\begin{aligned} d\mu(t) &= A'(t)\lambda(t)dt + A''(t)\mu(t)dt + b'(t)dt + \Sigma'(t)dw'(t) \\ d\gamma(t) &= b''(t)dt + \Sigma''(t)dw''(t) \end{aligned} \tag{2.17}$$

In (2.17), matrices $A'(t)$ and $A''(t)$ respectively account for possible elastic and viscosity effects; vectors $b'(t)$ and $b''(t)$ account for possible known trends; w' and w'' are 3-dimensional independent Wiener processes which affect the dynamics by the diffusion matrices $\Sigma'(t)$ and $\Sigma''(t)$.

The overall model for the motion, as it results from (2.16), (2.17), is described by a stochastic affine differential equation:

$$dX(t) = A(t)X(t)dt + b(t)dt + \Sigma(t)dw(t) \tag{2.18}$$

where $X(t)$ is the state vector:

$$X(t) = (X_1(t) \quad \cdots \quad X_9(t))^T = \left(\lambda^T(t) \quad \mu^T(t) \quad \gamma^T(t)\right)^T, \tag{2.19}$$

$A(t)$ is a block matrix:

$$A(t) = \begin{pmatrix} 0 & I & 0 \\ A'(t) & A''(t) & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{2.20}$$

$b(t)$ is a vector:

$$b(t) = \begin{pmatrix} 0^T & b'^T(t) & b''^T(t) \end{pmatrix}^T, \tag{2.21}$$

$\Sigma(t)$ is a block matrix:

$$\Sigma(t) = \begin{pmatrix} 0 & 0 \\ \Sigma'(t) & 0 \\ 0 & \Sigma''(t) \end{pmatrix}, \tag{2.22}$$

$w(t)$ is a 6-dimensional vector with independent Wiener processes as entries:

$$w(t) = \begin{pmatrix} w'^T(t) & w''^T(t) \end{pmatrix}^T. \tag{2.23}$$

The initial condition for (2.18), $X(t_0)$, is a 9-dimensional random vector with distribution $\pi(t_0)$.

The motion transformation φ depends on t only through the value of the state vector $X(t)$ and by an abuse of notation from (2.7) we may write:

$$\begin{aligned} &x(t) = \varphi(x(t_0); X(t)) = \\ &= \begin{pmatrix} \cos X_1(t) & -\sin X_1(t) \\ \sin X_1(t) & \cos X_1(t) \end{pmatrix} \begin{pmatrix} e^{X_7(t)} & X_8(t) \\ X_8(t) & \left(e^{X_9(t)} + X_8^2(t)\right)e^{-X_7(t)} \end{pmatrix} x(t_0) + \begin{pmatrix} X_2(t) \\ X_3(t) \end{pmatrix}, \quad t \in [t_0, t_f] \end{aligned} \tag{2.24}$$

From (2.24) we easily compute the inverse transformation:

$$\begin{aligned} &x(t_0) = \varphi^{-1}(x(t); X(t)) = P^{-1}(t)R^{-1}(t)(x(t) - T(t)) = \\ &= e^{-X_9(t)} \begin{pmatrix} X_8(t)\sin X_1(t) + & -X_8(t)\cos X_1(t) + \\ + e^{-X_7(t)}\left(e^{X_9(t)} + X_8^2(t)\right)\cos X_1(t) & + e^{-X_7(t)}\left(e^{X_9(t)} + X_8^2(t)\right)\sin X_1(t) \\ -X_8(t)\cos X_1(t) - e^{X_7(t)}\sin X_1(t) & -X_8(t)\sin X_1(t) + e^{X_7(t)}\cos X_1(t) \end{pmatrix} \cdot \\ &\cdot \left(x(t) - \begin{pmatrix} X_2(t) \\ X_3(t) \end{pmatrix} \right) \end{aligned} \tag{2.25}$$

§ 2.3 – The measurement equation

We now proceed to model those degradation factors which occur because of the data acquisition equipment.

We assume we perform measurements of the image at discrete measurement times t_i, $i = 1,2,\ldots,M$, $t_0 \le t_1 < t_2 < \ldots < t_M \le t_f$. By an abuse of notation suffix i will indicate evaluation of functions at time t_i.

Consequently, at each $t_i \in [t_0, t_f]$ we have a noisy, blurred, dimmed measured version z_i of f_i, which we define as follows:

$$z_i(x) = (H_i \cdot f_i)(x) + n_i(x), \quad x \in \mathbb{R}^2, \ \ i = 1,2,...,M \tag{2.26}$$

where H_i is a linear operator with suitable point spread function h_i which accounts for possible blurring effects in the measurement device and suitable dimming function l_i which accounts for possible local brightness variation; $n_i(\cdot)$ is a Gaussian process with zero mean and variance $\sigma_{n_i}^2$ which takes measurement uncertainty into account. We further assume that $n_i(\cdot)$ and $n_j(\cdot)$ are independent of each other, for $t_i \neq t_j$, and also that, for any t_i, $n_i(x)$ and $n_i(y)$ are independent of each other, for $x \neq y$.

In order to cast the problem into a better detailed framework, the operator H_i which transforms the image f_i into its (noiseless) measured version and accounts for possible blurring of the measurement instrument, is here assumed to be of the form [89]:

$$(H_i \cdot f_i)(x) = \int_{R^2} N_{\Sigma_b(i)}(x - \xi) f_i(\xi) d\xi \tag{2.27}$$

where the point spreading function $N_{\Sigma_b(i)}$ is a 2-dimensional Gaussian kernel with zero mean and (positive semidefinite)[5] covariance matrix $\Sigma_b(i)$. Furthermore the dimming function l_i is assumed known positive continuous, so that the same is not explicitly shown in (2.27), but it is included in f_i and will be disregarded in the following. Thus the measurement equation for each time instant t_i turns out to be:

$$z_i(x) = \int_{R^2} N_{\Sigma_b(i)}(x - \xi) f_i(\xi) d\xi + n_i(x). \tag{2.28}$$

[5] Here and in the following, integrals with singular covariance Gaussian kernels will be intended in the sense of distributions.

3. Edge detection from an image sequence with known motion

§ 3.1 – Problem set up

A significant class of problems, which is relevant to the image estimation, given full information about the motion, is the one which focuses on edge identification. Considering that the motion we suppose to know refers to the relative movement between the scene and the acquisition device, this problem turns out to be of particular relevance in all those applications in which the scene doesn't move at all, while the measurement device does.
The image reconstruction problem is the problem of estimating f at time k from the observed signals $z_i, \quad i = 1,2,\ldots,k$. As usual in estimation problems, and more generally in the so called inverse problems, we first need to investigate the well posedness of the problem itself, that is uniqueness and continuity of the solution in the absence of noise, or equivalently existence and continuity of the inverse operator H^{-1}. Such properties are equivalent to existence and robustness of the estimate obtained when considering noise. To that purpose, we state the following results, for which we refer to [84].

Theorem 3.1 – Let $f', f'' \in \mathcal{D}_0$ and $y', y'' \in H(\mathcal{D}_0)$ such that $y' = H(f')$, $y'' = H(f'')$. Given an open subset $A \subset \mathbb{R}^2$ if it results that

$$y'(x) = y''(x), \quad x \in A$$

then

$$f' = f'' \quad \text{in } \mathbb{R}^2.$$

Theorem 3.1 shows the existence of the inverse operator H^{-1} on $H(\mathcal{D}_0)$. Once we embed $H(\mathcal{D}_0)$ with the sup norm, or even with an L_2-norm, the operator H turns out to be continuous [84]. Furthermore we have the following.

Theorem 3.2 – Let $\mathcal{D}_0'$ (or $\mathcal{D}_0''$) be the compact subset of $\mathcal{D}_0$ defined in Theorem 2.4 (or Theorem 2.5). Then H^{-1} is continuous over $H(\mathcal{D}_0')$ (or $H(\mathcal{D}_0'')$).

Remark 3.3 – The above results extend to the case in which the mean vector and the covariance matrix in Gaussian kernel N possibly vary on $\mathbb{R}^2$.

In order to correctly face this problem, we have to specify the general measurement equation given in the previous chapter, taking available information about motion into account. Recalling (2.4) and (2.27), with an obvious change in the integration variable, we have:

$$z_i(x) = \int_{R^2} N_{\Sigma_b(i)}(x-\xi) f_i(\xi) d\xi + n_i(x) = \int_{R^2} N_{\Sigma_b(i)}(x-\xi) f_0\left(\varphi_i^{-1}(\xi)\right) d\xi + n_i(x) =$$

$$= \int_{R^2} N_{\Sigma_b(i)}(x-\varphi(t_i,\eta)) \left| \det\left\{ \frac{\partial \varphi(t_i,\eta)}{\partial \eta} \right\} \right| f_0(\eta) d\eta + n_i(x) = \left(\overline{H}_i \cdot f_0\right)(x) + n_i(x) \qquad (3.1)$$

Note that in the case of affine $\varphi(t,\cdot)$ the kernel of $\overline{H}_i$ in (3.1), as a function of x, is proportional to a Gaussian density with mean $\varphi(t,\eta)$ and covariance matrix $\Sigma_b(i)$. In the particular case of rigid motion the kernel of $\overline{H}_i$ is a *bona fide* Gaussian density, while obviously (2.12) yields $\vartheta(t) = \frac{\pi}{2}$, $\forall t$. If moreover the motion reduces to a mere translation, $\overline{H}_i$ reduces to the convolution structure of (2.26) with the simple translation of the mean value $\bar{x}_i = \bar{x} + \left(\lambda_2(t_i) \quad \lambda_3(t_i)\right)^T$ in the $N_{\Sigma_b(i)}$ kernel.

Hence, the aim we want to pursue is estimating f_0, given full information about the motion operator φ. Such a problem turns out to be linear, but with very high dimensionality. This feature is connected to the dimensionality of a matrix to be inverted and might be consequently handled by splitting the inversion of a matrix in a number of inversions of suitable minors of the matrix itself. Another approach might consider partitioning of the unknown image in sub images in such a way that basically no two different elements in the partition affect each other. Since the regular parts $\left\{\gamma_k^f\right\}$ are mostly responsible for the influence among the different elements of the partition, the above requirement would call for the existence of discontinuities along the boundaries of the sub images.

A further approach could be to ask the operator H_i to have local character (finite support), which would allow the observation z_i at each point to spatially depend essentially on the grey level within a bounded neighbourhood of that point. This latter aim can be achieved by pre-processing the measured image z_i with a suitable wavelet, a linear operator which enhances the local character of the overall resulting output operator.

In this work we implement this third approach, showing also that the proposed pre-processing has a positive effect on the signal-to-noise ratio, enhancing the discontinuity with respect to the regular part of the signal.

§ 3.2 – Image pre-processing

As said in the previous section, our first need for edge identification and detection is to enhance the local character of the output operator. To that purpose, for each t_i, $i = 1,2,\ldots,M$, we propose to pre-process the available data z_i and compute the increments, over a given space shift, of their convolution by two classes of suitable kernels. Let us introduce two classes of linear operators defined by 2-dimensional wavelet-like kernels:

$$\tilde{z}_{i,w}^{(1)}(x) = \int_{R^2} F_{(i,w;1)}(\tau) z_i(x-\tau) d\tau, \qquad w = 1,2,\ldots,W \tag{3.2}$$

$$\tilde{z}_{i,w}^{(2)}(x) = \int_{R^2} F_{(i,w;2)}(\tau) z_i(x-\tau) d\tau, \qquad w = 1,2,\ldots,W \tag{3.3}$$

where:

$$F_{(i,w;s)}(\tau) = \left|\det\{K_i\}\right|^{-2} N_{\Sigma(i,w;s)}^{(s,s)}(\varsigma)\Big|_{\varsigma = K_i^{-1}\tau}, \quad s = 1,2 \tag{3.4}$$

$K_i = R(t_i)P(t_i)$ and $N_{\Sigma}^{(s,s)}$, $s = 1,2$, denotes the second derivative w.r.t. the *s*-th component of the argument of a 2-dimensional Gaussian kernel with zero mean and (positive semidefinite) covariance matrix Σ.

The wavelet behaviour of $F_{(i,w;s)}$ (and in particular the zero area property) is inherited by that one of $N_{\Sigma(i,w;s)}^{(s,s)}$.

Remark 3.1. It is worth pointing out that the proposed pre-processing, once we disregard the measurement noise, preserves the identifiability property given in Section 3.1. In fact, the relationship between the initial image f_0 and the pre-processing outputs $\tilde{z}_{i,w}^{(s)}$, $s = 1,2$, at any time instant t_i, as described by (3.1), (3.2), (3.3), is the composition of three affine continuous transformation, namely φ_t^{-1}, and the convolutions with kernels $N_{\Sigma_b(i)}$ and $F_{(i,w;s)}$. It follows that, as said in chapter 2, if f_0 belongs to any suitable compact subset of $\mathcal{D}_0$, then in the absence of additive noise, the inverse transformation from $\tilde{z}_{i,w}^{(s)}$ to f_0 exists and is

continuous. Thus the identification problem for f_0 is well posed in that it enjoys a robust solution, even in the case of a single measure and/or of a single wavelet to be used.

From (3.2), (3.3) and taking (2.28) into account we have:

$$\begin{aligned}
\widetilde{z}_{i,w}^{(s)}(x) &= \left|\det\{K_i\}\right|^{-2} \int_{R^2} N_{\Sigma(i,w;s)}^{(s,s)}(\varsigma)\Big|_{\varsigma=K_i^{-1}\tau} \int_{R^2} N_{\Sigma_b(i)}(x-\tau-\xi) f_i(\xi) d\xi d\tau + \widetilde{n}_{i,w}^{(s)}(x) = \\
&= \left|\det\{K_i\}\right|^{-1} \int_{R^2} N_{\Sigma(i,w;s)}^{(s,s)}(\varsigma) \int_{R^2} N_{\Sigma_b(i)}(x-K_i\varsigma-\xi) f_i(\xi) d\xi d\varsigma + \widetilde{n}_{i,w}^{(s)}(x) = \\
&= \left|\det\{K_i\}\right|^{-1} \int_{R^2} N_{\Sigma(i,w;s)}^{(s,s)}(\varsigma) \int_{R^2} N_{\Sigma_b'(i)}\left(K_i^{-1}(x-\xi)-\varsigma\right) f_i(\xi) d\xi d\varsigma + \widetilde{n}_{i,w}^{(s)}(x) = \\
&= \left|\det\{K_i\}\right|^{-1} \int_{R^2} N_{\widetilde{\Sigma}(i,w;s)}^{(s,s)}(\eta)\Big|_{\eta=K_i^{-1}(x-\xi)} f_i(\xi) d\xi + \widetilde{n}_{i,w}^{(s)}(x) = \\
&= \int_{R^2} N_{\widetilde{\Sigma}(i,w;s)}^{(s,s)}(\eta) f_i(x-K_i\eta) d\eta + \widetilde{n}_{i,w}^{(s)}(x) = \\
&= \int_{R^2} N_{\widetilde{\Sigma}(i,w;s)}^{(s,s)}(\eta) f_0\left(\varphi_i^{-1}(x)-\eta\right) d\eta + \widetilde{n}_{i,w}^{(s)}(x), \qquad s=1,2
\end{aligned} \tag{3.5}$$

where:

$$\widetilde{n}_{i,w}^{(s)}(x) = \int_{R^2} F_{(i,w;s)}(\tau) n_i(x-\tau) d\tau \tag{3.6}$$

$$\Sigma_b'(i) = K_i^{-1} \Sigma_b(i) K_i^{-T} \tag{3.7}$$

$$\widetilde{\Sigma}(i,w;s) = \Sigma(i,w;s) + \Sigma_b'(i) \tag{3.8}$$

Equation (3.5) shows how the convolution operations (3.2), (3.3) on the blurred noisy data at time t_i correspond to convolution operations directly acting on the original image f_0, apart from a shift which compensates the motion, and the influence on the additive noise.
We now choose $\Sigma(i,w;s)$ in such a way that (3.8) yields a diagonal (positive semidefinite) $\widetilde{\Sigma}(i,w;s)$ matrix. Let us introduce the notations:

$$\Sigma_b'(i) = \begin{pmatrix} \sigma_{b;11}'^2(i) & \sigma_{b;12}'(i) \\ \sigma_{b;12}'(i) & \sigma_{b;22}'^2(i) \end{pmatrix},$$

$$\Sigma(i,w;s)=\begin{pmatrix}\sigma_{11}^2(i,w;s) & \sigma_{12}(i,w;s)\\ \sigma_{12}(i,w;s) & \sigma_{22}^2(i,w;s)\end{pmatrix},$$

$$\widetilde{\Sigma}(i,w;s)=\begin{pmatrix}\widetilde{\sigma}_{11}^2(i,w;s) & \widetilde{\sigma}_{12}(i,w;s)\\ \widetilde{\sigma}_{12}(i,w;s) & \widetilde{\sigma}_{22}^2(i,w;s)\end{pmatrix}=\begin{pmatrix}\sigma_{11}^2(i,w;s)+\sigma_{b;11}'^2(i) & \sigma_{12}(i,w;s)+\sigma_{b;12}'(i)\\ \sigma_{12}(i,w;s)+\sigma_{b;12}'(i) & \sigma_{22}^2(i,w;s)+\sigma_{b;22}'^2(i)\end{pmatrix}.$$

Then $\widetilde{\Sigma}(i,w;s)$ will be diagonal as soon as we choose:

$$\sigma_{12}(i,w;s)=-\sigma_{b;12}'(i). \tag{3.9}$$

Furthermore, in order for $\Sigma(i,w;s)$ to be positive semidefinite, we must have:

$$\sigma_{11}^2(i,w;s)\geq 0 \tag{3.10}$$

$$\sigma_{11}^2(i,w;s)\sigma_{22}^2(i,w;s)\geq\sigma_{12}^2(i,w;s)=\sigma_{b;12}'^2(i) \tag{3.11}$$

Conditions (3.9) – (3.11) define how to select a suitable $\Sigma(i,w;s)$. In the general case this has to be nondiagonal, with sufficiently high product of diagonal entries.

Should $\Sigma_b(i)$ be proportional to the identity matrix, and the deformation matrix $P(i)$ be a diagonal matrix (which implies the motion to preserve orthogonality of the original coordinate axes), then (3.9) shows that $\Sigma_b'(i)$ is still diagonal. In this case for $\widetilde{\Sigma}(i,w;s)$ to be diagonal it is necessary and sufficient that $\Sigma(i,w;s)$ is such.

Moreover, conditions (3.9) – (3.11) allow the convolution operator in (3.5) to factorize, so as to separate in $\widetilde{z}$ the effects of discontinuities of f_0 along the two coordinate axes:

$$\widetilde{z}_{i,w}^{(1)}(x)=\int_{R^2}N^{(2)}_{\widetilde{\sigma}_{11}^2(i,w;1)}(\eta_1)N_{\widetilde{\sigma}_{22}^2(i,w;1)}(\eta_2)f_0\left(\varphi_i^{-1}(x)-\eta\right)d\eta+\widetilde{n}_{i,w}^{(s)}(x), \tag{3.12}$$

$$\widetilde{z}_{i,w}^{(2)}(x)=\int_{R^2}N_{\widetilde{\sigma}_{11}^2(i,w;1)}(\eta_1)N^{(2)}_{\widetilde{\sigma}_{22}^2(i,w;1)}(\eta_2)f_0\left(\varphi_i^{-1}(x)-\eta\right)d\eta+\widetilde{n}_{i,w}^{(s)}(x), \tag{3.13}$$

where N_{σ^2} is a 1-dimensional Gaussian kernel with zero mean and variance σ^2, and $N_{\sigma^2}^{(q)}$ is its $q-th$ derivative $(q=1,2)$.

A final step in the preprocessing procedure is evaluating increments of (3.12), (3.13) respectively over shifts $\Delta_i^{(1)}, \Delta_i^{(2)}$ given by:

$$\Delta_i^{(s)} = \kappa_i^{(s)}\alpha, \quad s = 1,2$$

with $\kappa_i^{(s)}$ defined by (2.10), (2.11) and α a positive number. Let us define the following total variations:

$$F_\alpha^{(1)}(\xi_1,\xi_2) = f_0(\xi_1+\alpha,\xi_2) - f_0(\xi_1,\xi_2)$$
$$F_\alpha^{(2)}(\xi_1,\xi_2) = f_0(\xi_1,\xi_2+\alpha) - f_0(\xi_1,\xi_2)$$

For $s=1$ we have from (3.12):

$$\begin{aligned}
&\tilde{z}_{i,w}^{(1)}\left(x+\Delta_i^{(1)}\right) - \tilde{z}_{i,w}^{(1)}(x) = \\
&= \int_{R^2} N^{(2)}_{\tilde{\sigma}_{11}^2(i,w;1)}(\eta_1) N_{\tilde{\sigma}_{22}^2(i,w;1)}(\eta_2)\left[f_0\left(\varphi_i^{-1}\left(x+\Delta_i^{(1)}\right)-\eta\right) - f_0\left(\varphi_i^{-1}(x)-\eta\right)\right]d\eta + \tilde{v}_{i,w}^{(1)}(x) = \\
&= \int_{R^2} N^{(2)}_{\tilde{\sigma}_{11}^2(i,w;1)}(\eta_1) N_{\tilde{\sigma}_{22}^2(i,w;1)}(\eta_2)\left[f_0\left(\varphi_i^{-1}(x)-\eta+\begin{pmatrix}\alpha\\0\end{pmatrix}\right) - f_0\left(\varphi_i^{-1}(x)-\eta\right)\right]d\eta + \tilde{v}_{i,w}^{(1)}(x) = \\
&= \int_{R^2} N^{(2)}_{\tilde{\sigma}_{11}^2(i,w;1)}\left(\varphi_i^{-1}(x)\big|_1 - \xi_1\right) N_{\tilde{\sigma}_{22}^2(i,w;1)}\left(\varphi_i^{-1}(x)\big|_2 - \xi_2\right) F_\alpha^{(1)}(\xi_1,\xi_2)\,d\xi_1 d\xi_2 + \tilde{v}_{i,w}^{(1)}(x) = \\
&= \int_{R^2} N^{(2)}_{\tilde{\sigma}_{11}^2(i,w;1)}\left(\varphi_i^{-1}(x)\big|_1 - \xi_1\right) N_{\tilde{\sigma}_{22}^2(i,w;1)}\left(\varphi_i^{-1}(x)\big|_2 - \xi_2\right)\left[\delta_\alpha^{(1)} f_0(\xi_1,\xi_2) + \partial_\alpha^{(1)} f_0(\xi_1,\xi_2)\right]d\xi_1 d\xi_2 + \\
&\qquad\qquad + \tilde{v}_{i,w}^{(1)}(x)
\end{aligned}$$

(3.14)

where $v\big|_l$ denotes the *l*-th component of vector v, and:

- $\tilde{v}_{i,w}^{(1)}(x) = \tilde{n}_{i,w}^{(1)}\left(x+\Delta_i^{(1)}\right) - \tilde{n}_{i,w}^{(1)}(x)$

  (3.15)
- $\delta_\alpha^{(1)} f_0(\xi_1,\xi_2)$ and $\partial_\alpha^{(1)} f_0(\xi_1,\xi_2)$ account respectively for the discontinuous and continuous variation of f when moving from (ξ_1,ξ_2) to $(\xi_1+\alpha,\xi_2)$ along a straight line. More precisely we have:

- $\delta_{\alpha}^{(1)}(\xi_1,\xi_2)=\sum^{*}\left[f_0\left(\xi_1+t^{+},\xi_2\right)-f_0\left(\xi_1+t^{-},\xi_2\right)\right]$

  (3.16) where $\sum^{*}$ sums over all $t\in(0,\alpha)$, such that $f_0\left(\xi_1+t^{+},\xi_2\right)-f_0\left(\xi_1+t^{-},\xi_2\right)\neq 0$

- $\partial_{\alpha}^{(1)}(\xi_1,\xi_2)=F_{\alpha}^{(1)}(\xi_1,\xi_2)-\delta_{\alpha}^{(1)}(\xi_1,\xi_2).$

  (3.17)

In a similar way, for $s=2$ (and with an obvious meaning of the notations) we get from (3.13):

$$\tilde{z}_{i,w}^{(2)}\left(x+\Delta_i^{(2)}\right)-\tilde{z}_{i,w}^{(2)}(x)=$$
$$=\int_{R^2}N_{\tilde{\sigma}_{11}^2(i,w;2)}\left(\varphi_i^{-1}(x)\big|_1-\xi_1\right)N^{(2)}_{\tilde{\sigma}_{22}^2(i,w;2)}\left(\varphi_i^{-1}(x)\big|_2-\xi_2\right)\left[\delta_{\alpha}^{(2)}f_0(\xi_1,\xi_2)+\partial_{\alpha}^{(2)}f_0(\xi_1,\xi_2)\right]d\xi_1 d\xi_2+$$
$$+\tilde{v}_{i,w}^{(2)}(x)$$

(3.18)

where:

$$\tilde{v}_{i,w}^{(2)}(x)=\tilde{n}_{i,w}^{(2)}\left(x+\Delta_i^{(2)}\right)-\tilde{n}_{i,w}^{(2)}(x) \tag{3.19}$$

§ 3.3 – Localization effect

To evidentiate the localization effect of the pre-processing we develop some simplified approximate versions of (3.14), (3.18) by introducing the following reasonable regularity assumptions on the original image *f* as compared to the pre-processing parameters.

1. The total variation $F_{\alpha}^{(s)}(\xi_1,\xi_2)$, $s=1,2$ behaves approximately as a constant w.r.t. ξ_{3-s} within the range around $\varphi_i^{-1}(x)\big|_{3-s}$ in which $N_{\tilde{\sigma}^2_{3-s,3-s}(i,w;s)}\left(\varphi_i^{-1}(x)\big|_{3-s}-\xi_{3-s}\right)$ is significantly above zero[6].

[6] Note that obviously $3-s$ denotes the component in $\mathbb{R}^2$ other than the s-th one

2. The continuous variation $\partial_\alpha^{(s)} f(\xi_1,\xi_2)$ behaves approximately as an affine function of ξ_s for $\xi_{3-s} \cong \varphi_i^{-1}(x)\big|_{3-s}$ and ξ_s within the range around $\varphi_i^{-1}(x)\big|_s$ in which $N^{(2)}_{\tilde{\sigma}^2_{s,s}(i,w;s)}\left(\varphi_i^{-1}(x)\big|_s - \xi_s\right)$ is significantly away from zero.
3. Within the range mentioned in 2. above, $f_0(\xi_1,\xi_2)$ undergoes at most one jump (located say in $\xi_s = \bar{\xi}_s$ and of size say $J^{(s)}\left(\varphi_i^{-1}(x)\right)$, with $J^{(s)}\left(\varphi_i^{-1}(x)\right) = 0$ in the case of no jump), when $\xi_{3-s} \cong \varphi_i^{-1}(x)\big|_{3-s}$ and ξ_s is increased by $\Delta_i^{(s)}$ around $\varphi_i^{-1}(x)\big|_s$.

Taking assumptions 1.-3. into account, equations (3.14), (3.18) are approximated by:

$$
\begin{aligned}
&\tilde{z}^{(s)}_{i,w}\left(x+\Delta_i^{(s)}\right) - \tilde{z}^{(s)}_{i,w}(x) \cong \\
&\cong \int_{\Re} N^{(2)}_{\tilde{\sigma}^2_{ss}(i,w;s)}\left(\varphi_i^{-1}(x)\big|_s - \xi_s\right)\left[\delta_\alpha^{(s)} f_0(\xi_1,\xi_2) + \partial_\alpha^{(s)} f_0(\xi_1,\xi_2)\right]\Big|_{\xi_{3-s}=\varphi_i^{-1}(x)\big|_{3-s}} d\xi_s + \tilde{v}^{(s)}_{i,w}(x) \cong \\
&\cong \int_{\Re} N^{(2)}_{\tilde{\sigma}^2_{ss}(i,w;s)}\left(\varphi_i^{-1}(x)\big|_s - \xi_s\right)\delta_\alpha^{(s)} f_0(\xi_1,\xi_2)\Big|_{\xi_{3-s}=\varphi_i^{-1}(x)\big|_{3-s}} d\xi_s + \tilde{v}^{(s)}_{i,w}(x) \cong \\
&\cong J^{(s)}\left(\varphi_i^{-1}(x)\right)\rho^{(s)}_{i,w}(x) + \tilde{v}^{(s)}_{i,w}(x) = \varsigma^{(s)}_{i,w}(x) \qquad\qquad s = 1,2
\end{aligned}
$$

(3.20)

where:

$$
\begin{aligned}
\rho^{(s)}_{i,w}(x) &= \int_{\bar{\xi}_s-\alpha}^{\bar{\xi}_s} N^{(2)}_{\tilde{\sigma}^2_{ss}(i,w;s)}\left(\varphi_i^{-1}(x)\big|_s - \xi_s\right) d\xi_s = \\
&= N^{(1)}_{\tilde{\sigma}^2_{ss}(i,w;s)}\left(\varphi_i^{-1}(x)\big|_s - \left(\bar{\xi}_s - \alpha\right)\right) - N^{(1)}_{\tilde{\sigma}^2_{ss}(i,w;s)}\left(\varphi_i^{-1}(x)\big|_s - \bar{\xi}_s\right)
\end{aligned}
\tag{3.21}
$$

With the notation $y = \varphi_i^{-1}(x)$, eq. (3.20) becomes:

$$
\varsigma^{(s)}_{i,w}(\varphi(t_i,y)) = J^{(s)}(y)\rho^{(s)}_{i,w}(\varphi(t_i,y)) + \tilde{v}^{(s)}_{i,w}(\varphi(t_i,y)) \tag{3.22}
$$

Equation (3.22) enlightens the localization effect obtained by the pre-processing, as well as the enhancement of discontinuities and the smoothing out of the regular parts in *f*.

In particular the difference $\tilde{z}^{(s)}_{i,w}\left(\varphi(t_i,y)+\Delta_i^{(s)}\right) - \tilde{z}^{(s)}_{i,w}(\varphi(t_i,y))$ yields a noisy evaluation of the jump size $J^{(s)}(y)$ of *f* around y along the direction s, s = 1, 2. In addition, the original noise

n_i is filtered out, in the sense that the signal-to-noise ratio can be improved by choosing the covariance matrix $\Sigma_{(i,w;s)}$ large enough, as it will be shown in the next section.

Note that assumptions 1, 2 essentially imply some regularity of the total variation $F_\alpha^{(s)}(\xi_1,\xi_2)$, when (ξ_1,ξ_2) moves orthogonally w.r.t. x_s, and of the continuous variation $\partial_\alpha^{(s)} f_0(\xi_1,\xi_2)$, when (ξ_1,ξ_2) moves along x_s.

On the other hand, assumption 3 essentially involves the behaviour of the discontinuity set C_f in that it prevents C_f from having branches "too close" each other.

As we said in previous chapter, for the purpose of our work, it suffices to consider the subset $\mathcal{D}_0'''$ of $\mathcal{D}_0$, introduced in Definition 2.6, for which assumption 3 is largely satisfied.

The grid *G*, therewith given, is thus constituted by two bundles of straight lines, respectively parallel to one or the other of the coordinate axes, and separated by a distance Δ. Under $\varphi(t,\cdot)$, *G* is transformed in a new possibly non orthogonal grid, G_t, again constituted by two bundles of straight lines, with directions respectively identified by $\kappa_1(t)$, $\kappa_2(t)$ as in (2.10), (2.11) and separated by distances respectively given by:

$$d_1(t) = \Delta\sqrt{\lambda_4^2(t) + \lambda_5^2(t)}$$

$$d_2(t) = \Delta\sqrt{\lambda_5^2(t) + \lambda_6^2(t)}$$

Obviously the angle which characterizes the grid G_t is equal to $\vartheta(t)$ given in (2.12).

Because of the very definition of $\mathcal{D}_0'''$, it turns out that, whenever $y \in C_f$, then either $J^{(1)}(y)$ or $J^{(2)}(y)$ or both are different from zero.

Thus in $\mathcal{D}_0'''$ the edge identification problem is well posed in that the pre-processing output $\tilde{z}_{i,w}^{(s)}\left(\varphi(t_i,y)+\Delta_i^{(s)}\right) - \tilde{z}_{i,w}^{(s)}(\varphi(t_i,y))$, $s=1,2$ is a noisy measurement of the discontinuity size around y, significantly non vanishing only at $y \in C_f$.

In case local preprocessing might use some prior knowledge of the image *f*, then the orthogonal axes x_s, $s=1,2$, can be chosen so as to guarantee non vanishing of either one of $\tilde{z}_{i,w}^{(s)}\left(\varphi(t_i,y)+\Delta_i^{(s)}\right) - \tilde{z}_{i,w}^{(s)}(\varphi(t_i,y))$, $s=1,2$, at any $y \in C_f$, even if angles in C_f are less than $\pi/2$.

§ 3.4 – Improving the signal-to-noise ratio

In order to evaluate the improvement of the signal-to-noise ratio, achievable by the pre-processing, it is necessary to investigate the statistical properties of the "equivalent" measurement noise $\tilde{v}_{i,w}^{(s)}(x),\ s=1,2$, as defined in (3.6), (3.15), (3.19). We have:

$$\tilde{v}_{i,w}^{(s)}(x)=\int_{\Re^2} F_{(i,w;s)}(\tau)\left[n_i\left(x+\Delta_i^{(s)}-\tau\right)-n_i(x-\tau)\right]d\tau=$$

$$=\left|\det\{K_i\}\right|^{-1}\int_{\Re^2} N_{\Sigma(i,w;s)}^{(s,s)}(\xi)\left[n_i\left(x+\Delta_i^{(s)}-K_i\xi\right)-n_i\left(x-K_i\xi\right)\right]d\xi \tag{3.23}$$

Thus, for any fixed x, t_i, w, it is evident that $\tilde{v}_{i,w}^{(s)}(x)$ is a zero mean Gaussian random variable. Furthermore, for any fixed x, and $\upsilon, w=1,2,\ldots,W$, $\tilde{v}_{i,w}^{(s)}(x)$ is independent of $\tilde{v}_{j,\upsilon}^{(s)}(x),\ \forall t_i\neq t_j$, while the covariance between $\tilde{v}_{i,w}^{(s)}(x)$ and $\tilde{v}_{i,\upsilon}^{(s)}(x)$ is given by:

$$\begin{aligned}\psi_{i;w,\upsilon}^{(s)} &= E[\tilde{v}_{i,w}^{(s)}(x)\tilde{v}_{i,\upsilon}^{(s)}(x)]=\\ &=\left|\det\{K_i\}\right|^{-2}\int_{\Re^2}\int_{\Re^2} N_{\Sigma(i,w;s)}^{(s,s)}(\xi)N_{\Sigma(i,\upsilon;s)}^{(s,s)}(\theta)\\ &\quad E\left[\left(n_i\left(x+\Delta_i^{(s)}-K_i\xi\right)-n_i\left(x-K_i\xi\right)\right)\left(n_i\left(x+\Delta_i^{(s)}-K_i\theta\right)-n_i\left(x-K_i\theta\right)\right)\right]d\xi d\theta=\\ &=2\left|\det\{K_i\}\right|^{-2}\sigma_{n_i}^2\int_{\Re^2} N_{\Sigma(i,w;s)}^{(s,s)}(\xi)N_{\Sigma(i,\upsilon;s)}^{(s,s)}(\xi)d\xi-\\ &\quad -2\left|\det\{K_i\}\right|^{-2}\sigma_{n_i}^2\int_{\Re^2} N_{\Sigma(i,w;s)}^{(s,s)}(\xi)N_{\Sigma(i,\upsilon;s)}^{(s,s)}\left(\xi-K_i^{-1}\Delta_i^{(s)}\right)d\xi=\\ &=2\left|\det\{K_i\}\right|^{-2}\sigma_{n_i}^2\left[N_{\Sigma(i,w;s)+\Sigma(i,\upsilon;s)}^{(s,s,s,s)}(0)-N_{\Sigma(i,w;s)+\Sigma(i,\upsilon;s)}^{(s,s,s,s)}\left(K_i^{-1}\Delta_i^{(s)}\right)\right]\end{aligned} \tag{3.24}$$

where the last equality requires nonsingularity of $\Sigma(i,w;s)+\Sigma(i,v;s)$. Note that (3.24) is independent of x.

With reference to eq. (3.22), the signal-to-noise ratio can be defined as:

$$r^{(s)}(i,w)=\frac{J^{(s)}(y)\rho_{i,w}^{(s)}(\varphi(t_i,y))}{\sqrt{\psi_{i;w,w}^{(s)}}}, \tag{3.25}$$

where now, in order to have the greatest sensitivity, $\rho_{i,w}^{(s)}(\varphi(t_i,y))$ should be derived as in (3.20), but considering the symmetric difference:

$$\tilde{z}_{i,w}^{(s)}\left(x+\Delta_i^{(s)}/2\right)-\tilde{z}_{i,w}^{(s)}\left(x-\Delta_i^{(s)}/2\right), \qquad \text{with } x=\varphi(t_i,y).$$

It results:

$$\rho_{i,w}^{(s)}(\varphi(t_i,y))=\int_{y_s-\alpha/2}^{y_s+\alpha/2}N_{\tilde{\sigma}_{ss}^2(i,w;s)}^{(2)}(y_s-\xi_s)d\xi_s=-2N_{\tilde{\sigma}_{ss}^2(i,w;s)}^{(1)}\left(\frac{\alpha}{2}\right)=\frac{\alpha\exp\left(-\dfrac{\alpha^2}{8\tilde{\sigma}_{ss}^2(i,w;s)}\right)}{\sqrt{2\pi}\tilde{\sigma}_{ss}^3(i,w;s)}. \tag{3.26}$$

For sake of simplicity, we now consider the case of Σ_b proportional to the identity matrix and of a motion with a deformation matrix $P(i)$ diagonal. In this case we may choose a diagonal $\Sigma_w(i,w;s)$ with $\sigma_{3-s,3-s}^2(i,w;s)=0$. Then from (3.24) we have:

$$\psi_{i;w,w}^{(s)}=2\left|\det\{K_i\}\right|^{-2}\sigma_{n_i}^2\left(N_{2\sigma_{ss}^2(i,w;s)}^{(4)}(0)-N_{2\sigma_{ss}^2(i,w;s)}^{(4)}(\alpha)\right)=$$

$$=\frac{\sqrt{2}\sigma_{n_i}^2\left[3-\exp\left(-\dfrac{\alpha^2}{2\sigma_{ss}^2(i,w;s)}\right)\left(3-6\dfrac{\alpha^2}{\sigma_{ss}^2(i,w;s)}+\dfrac{\alpha^4}{\sigma_{ss}^4(i,w;s)}\right)\right]}{\lambda_4^2(i)\lambda_6^2(i)\sqrt{\pi}\sigma_{ss}^5(i,w;s)}. \tag{3.27}$$

For $\varepsilon = \dfrac{\alpha}{\sigma_{ss}(i,w;s)}$, $\chi = \dfrac{\sigma_{ss}(i,w;s)}{\widetilde{\sigma}_{ss}(i,w;s)}$, (3.25) becomes:

$$r^{(s)}(i,w) = J^{(s)}(y)\frac{\varepsilon\chi^3 \exp\left(-\dfrac{\varepsilon^2}{8}\chi^2\right)\lambda_4(i)\lambda_6(i)\sigma_{ss}^{1/2}(i,w;s)}{\sqrt[4]{2\pi}\sqrt{2}\sigma_{n_i}\sqrt{3-\exp\left(-\dfrac{\varepsilon^2}{2}\right)\left(3-6\varepsilon^2+\varepsilon^4\right)}} \tag{3.28}$$

As $\sigma_{ss}(i,w;s)$ increases, ε decreases to zero and χ increases to 1. Thus an easy computation from (3.28) shows that:

$$r^{(s)}(i,w) = \left[\frac{J^{(s)}(y)\lambda_4(i)\lambda_6(i)}{\sqrt[4]{2\pi}\sqrt{15}\sigma_{n_i}} + o\left(\frac{1}{\sigma_{ss}(i,w;s)}\right)\right]\sigma_{ss}^{1/2}(i,w;s) \tag{3.29}$$

From (3.29) we draw the conclusion that, with respect to the original signal-to-noise ratio $\dfrac{J^{(s)}(y)}{\sigma_{n_i}}$, $r^{(s)}(i,w)$ increases with $\sigma_{ss}^{1/2}(i,w;s)$, with a rate about proportional to the product $\lambda_4(i)\lambda_6(i)$ (the mesh area of the deformed grid).

§ 3.5 – Filtering for discontinuities and edge detection

We now consider again equation (3.20) for $x = \varphi(t_i, y)$:

$$\varsigma_{i,w}^{(s)}(\varphi(t_i,y)) = J^{(s)}(y)\rho_{i,w}^{(s)}(\varphi(t_i,y)) + \widetilde{\nu}_{i,w}^{(s)}(\varphi(t_i,y)). \tag{3.30}$$

Let us introduce the *W*-dimensional vectors:

$$\varsigma_i^{(s)}(\varphi(t_i,y)) = \left(\varsigma_{i,w}^{(s)}(\varphi(t_i,y)),\ \ w = 1,2,\ldots,W \right)^T, \qquad i = 1,2,\ldots,M$$

$$\rho_i^{(s)}(\varphi(t_i,y)) = \left(\rho_{i,w}^{(s)}(\varphi(t_i,y)),\ \ w = 1,2,\ldots,W \right)^T, \qquad i = 1,2,\ldots,M$$

$$\tilde{\nu}_i^{(s)}(x) = \left(\tilde{\nu}_{i,w}^{(s)}(x),\ \ w = 1,2,\ldots,W \right)^T, \qquad i = 1,2,\ldots,M$$

It follows from (3.24) that $\tilde{\nu}_i^{(s)}(x)$ is Gaussian, with zero mean and covariance matrix given by:

$$\Psi_i^{(s)} = E\left[\tilde{\nu}_i^{(s)}(x)\ \tilde{\nu}_i^{(s)^T}(x) \right] = \left(\psi_{i;w,\upsilon}^{(s)},\ \ w,\upsilon = 1,2,\ldots,W \right),$$

which again turns out to be independent of x.

We further define the kW -dimensional vectors:

$$\varsigma^{(s),k}(y) = \left(\varsigma_i^{(s)^T}(\varphi(t_i,y)),\ \ i = 1,2,\ldots,k \right)^T,$$

$$\rho^{(s),k}(y) = \left(\rho_i^{(s)^T}(\varphi(t_i,y)),\ \ i = 1,2,\ldots,k \right)^T,$$

$$\tilde{\nu}^{(s),k}(y) = \left(\tilde{\nu}_i^{(s)^T}(\varphi(t_i,y)),\ \ i = 1,2,\ldots,k \right)^T.$$

Then the output relationship (3.30) when i, w range respectively in $\{1,2,\ldots,k\}$ and $\{1,2,\ldots,W\}$ becomes:

$$\varsigma^{(s),k}(y) = J^{(s)}(y)\rho^{(s),k}(y) + \tilde{\nu}^{(s),k}(y) \qquad s = 1,2 \qquad k = 1,2,\ldots,M \qquad (3.31)$$

with $\tilde{\nu}^{(s),k}(y)$ kW-dimensional Gaussian vector, with zero-mean and block-diagonal covariance matrix given by:

$$\Psi^{(s),k} = diag\left(\Psi_i^{(s)},\ \ i = 1,2,\ldots,k \right), \qquad k = 1,2,\ldots,M \tag{3.32}$$

independent of y.

For each k, (3.31) defines a linear Gaussian estimation problem for $J^{(s)}(y)$, $s = 1,2$, that is the problem of estimating the jump size of the original image at y in the direction s given all measurements up to time t_k.

As well known, its optimal (minimum variance) solution is given by the Markov estimate:

$$\hat{J}^{(s),k}(y) = \left[\rho^{(s),k\,T}(y) \left(\Psi^{(s),k} \right)^{-1} \rho^{(s),k}(y) \right]^{-1} \rho^{(s),k\,T}(y) \left(\Psi^{(s),k} \right)^{-1} \varsigma^{(s),k}(y) =$$

$$= \frac{\sum_{i=1}^{k} \rho_i^{(s)T}(\varphi(t_i, y)) \left(\Psi_i^{(s)} \right)^{-1} \varsigma_i^{(s)}(\varphi(t_i, y))}{\sum_{i=1}^{k} \rho_i^{(s)T}(\varphi(t_i, y)) \left(\Psi_i^{(s)} \right)^{-1} \rho_i^{(s)}(\varphi(t_i, y))}. \tag{3.33}$$

$\hat{J}^{(s),k}(y)$ is again Gaussian, with mean $J^{(s)}(y)$ and variance given by:

$$\sigma^2\left(\hat{J}^{(s),k}(y) \right) = \left[\sum_{i=1}^{k} \rho_i^{(s)T}(\varphi(t_i, y)) \left(\Psi_i^{(s)} \right)^{-1} \rho_i^{(s)}(\varphi(t_i, y)) \right]^{-1} \tag{3.34}$$

The estimation algorithm can be given an iterative structure, enhancing the innovation brought in by the last measurement. In this framework we have:

$$\hat{J}^{(s),k+1}(y) = \hat{J}^{(s),k}(y) + \sigma^2\left(\hat{J}^{(s),k+1}(y) \right) \rho_{k+1}^{(s)\,T}(\varphi(t_{k+1}, y)) \left(\Psi_{k+1}^{(s)} \right)^{-1}.$$

$$\cdot[\varsigma_{k+1}^{(s)}(\varphi(t_{k+1},y))-\rho_{k+1}^{(s)}(\varphi(t_{k+1},y))\hat{J}^{(s),k}(y)],\quad k=1,2,\ldots,M-1 \qquad (3.35)$$

$$\sigma^2\left(\hat{J}^{(s),k+1}(y)\right)=\left[\left(\sigma^2\left(\hat{J}^{(s),k}(y)\right)\right)^{-1}+\rho_{k+1}^{(s)\,T}(\varphi(t_{k+1},y))\Psi_{k+1}^{(s)}\rho_{k+1}^{(s)}(\varphi(t_{k+1},y))\right]^{-1}. \qquad (3.36)$$

We have to remark that the well behaving of the estimates $\hat{J}^{(s),k+1}(y)$, $s=1,2$, given by (3.35) and specifically their unbiasedness, strongly rely on the assumptions mentioned in section 3.3. In case *y* is such that these assumptions are not fully justified, one might expect the estimate accuracy to be somehow affected.

This procedure implements a total ignorance about existence of discontinuities at any given *y*, as well as stochastic independence between $J^{k+1}(y)$ and $J^{k+1}(y')$, $\forall y \neq y'$.

Should this be not the case, one might instead formulate a Bayesian estimation procedure, accounting for a priori distribution of the discontinuity set.

However, due to the linear character of the estimation procedures, the estimates $\hat{J}^{(1),k+1}(y)$, $\hat{J}^{(2),k+1}(y)$, as well as their modulus $\hat{J}^{k+1}(y)=\left(\left(\hat{J}^{(1),k+1}(y)\right)^2+\left(\hat{J}^{(2),k+1}(y)\right)^2\right)^{1/2}$ may still be granted a Gaussian distribution.

The edge detection problem is strongly connected to the one of jump size estimation. Indeed, for any given *y*, whenever we have the distribution of $\hat{J}^{k+1}(y)$, as well as its actual values, we may formulate and solve a hypothesis test with the null hypothesis $\mathcal{H}_0$ being:

$$\mathcal{H}_0=\left\{E\left\{\hat{J}^{k+1}(y)\right\}=0\right\}.$$

The problem can be solved by standard procedures for any fixed acceptance region and increasing confidence level.

Last, it is worth pointing out that strictly speaking the bijective assumption for $\varphi(t,\cdot)$ (Def. 2.7, b)) is only requested for $t=t_i \quad i=1,2,\ldots,M$. Moreover eq. (2.4) shows that existence of

the inverse mapping $y = \varphi_t^{-1}(x)$ may be waived at any x such that uniqueness of the value $f(y)$ for all possible $y : x = \varphi_t(y)$ is guaranteed.

This opens the way to an extension of this work to situations of applicative interest in which more than one object move on their own without overlapping each other on a constant background.

§ 3.6 – Application to simulated and real data

Several simulations have been done in order to establish real potentialities of the proposed methodology. The rationale behind experimentations was to increase step by step the complexity of the imaged scenario as far as the motion of objects was concerned, before considering in the last experiment a first application on real data coming from a camera acquisition.
We began with the case of null motion for a blurred noisy simulated image representing a four-corner star (figure 3.1). Both the variance of the additive noise and the covariance matrix of the gaussian blur kernel are supposed not to vary as the time increases ($\sigma_n^2(i) = 0.01$, $\Sigma_b(i) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, $i = 1,2,\ldots,M$). The image was additionally characterized by a regular sinusoidal grey level equal to $3\sin\left(\frac{x_1}{N}\pi\right)\sin\left(\frac{x_2}{N}\pi\right)$, where N is the dimension of the image along each direction and (x_1, x_2) are the coordinates of the generic pixel belonging to the image, accounting for possibly different enlightenment conditions in the scene. Such an image turned out to be useful also to test the behaviour of the algorithm when narrow cusps such as the corners of the star are concerned. Ten frames were considered, each with a different realization of the noise process. Three wavelets for the multiscale processing were chosen, with constant covariance matrices proportional to the identity matrix, with nonzero entries respectively equal to 0.5, 1.0 and 2.0. The output of the algorithm is reported in figure 3.2.

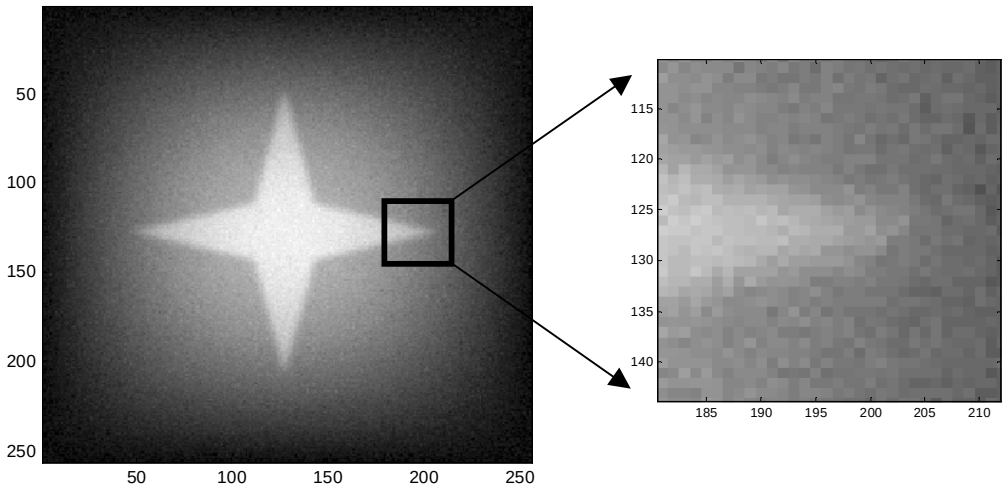


Figure 3.1 – One of the 10 frames considered in the first experiment and the zoom of a corner of the star to stress the level of image degradation.

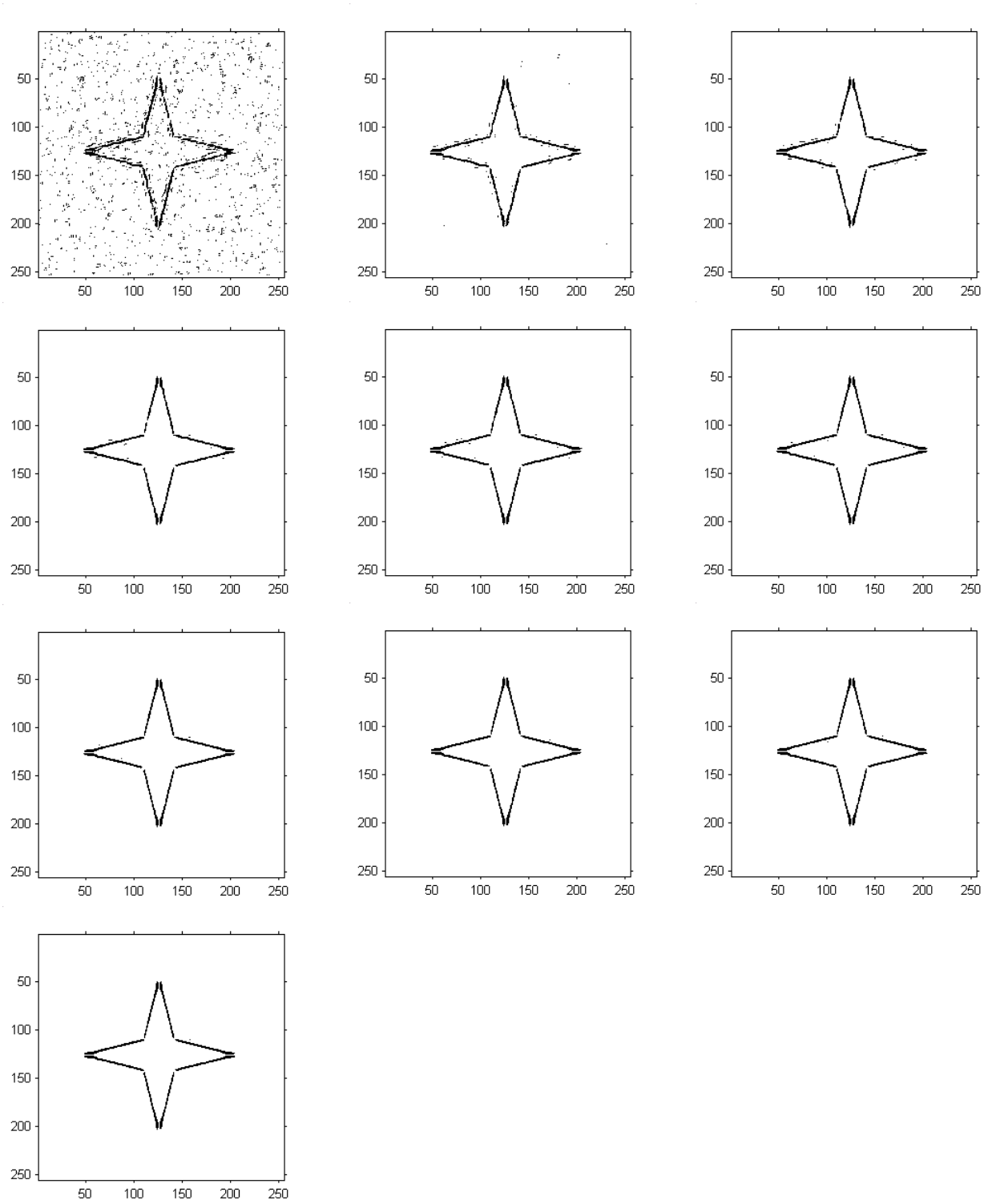


Figure 3.2 – Output of the edge detection algorithm for the chosen values of the wavelets.

It's worth to notice effectiveness and efficiency of the estimation procedure in removing false alarms and in identifying real edge points, as the acquired information increases. This was the main aim we were planning to achieve when we started to face this class of problems. This issue is further stressed by analysing the decreasing trend in normalized Hausdorff metric related to the distance between detected and real edge points, as well as the sudden shrink in the normalized number of misplaced edge points reported in figure 3.3. In particular we notice that both the two drop down an order of magnitude after very few images of the sequence (three in this case) are processed.

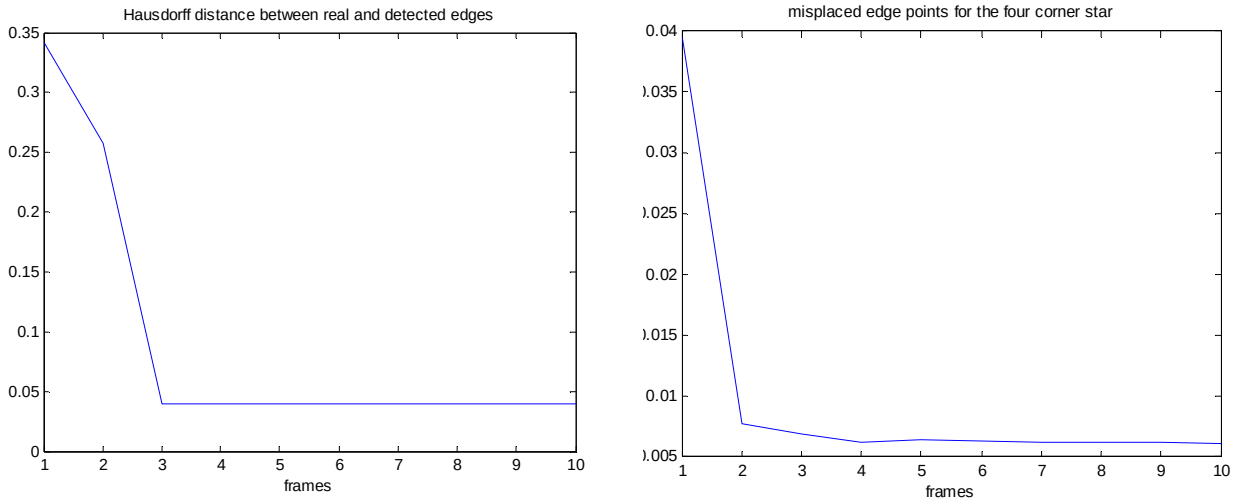


Figure 3.3 – Normalized Hausdorff metric between real and detected edge points and normalized number of misplaced edge points for the four corner star.

The successive experiments concerned the introduction of motion in the image. In particular we first considered a mere translation, simulating a circle whose center moves along a sinusoidal path in x direction, as reported in figure 3.4. The same level of degradation as the previous one was introduced on the generic frame, with the exception of the regular part which was suppressed in this case. The circle shape was chosen to verify that no privileged direction exists for edge detection. Results are reported in figure 3.5, where the detected edge is drawn for the initial image. Indeed we want to recall here that processing is made on the initial image, once the transformation due to known motion is compensated. Three wavelets were chosen, with same characteristics as the previous ones, but with nonzero entries respectively equal to 1.0, 2.0 and 4.0. We may again appreciate effectiveness and efficiency of the proposed methodology.

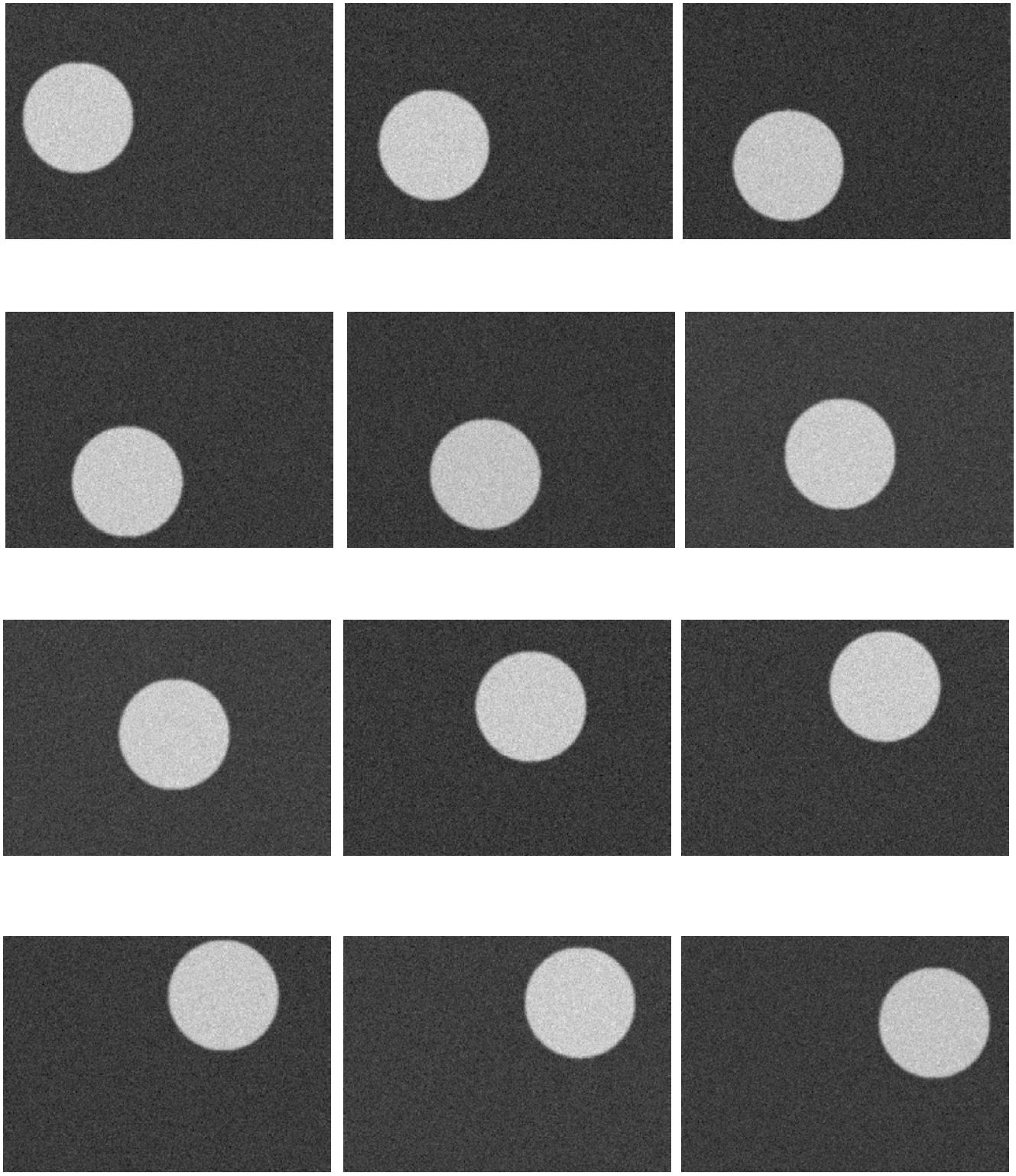


Figure 3.4 – The circle moving along a sinusoidal path.

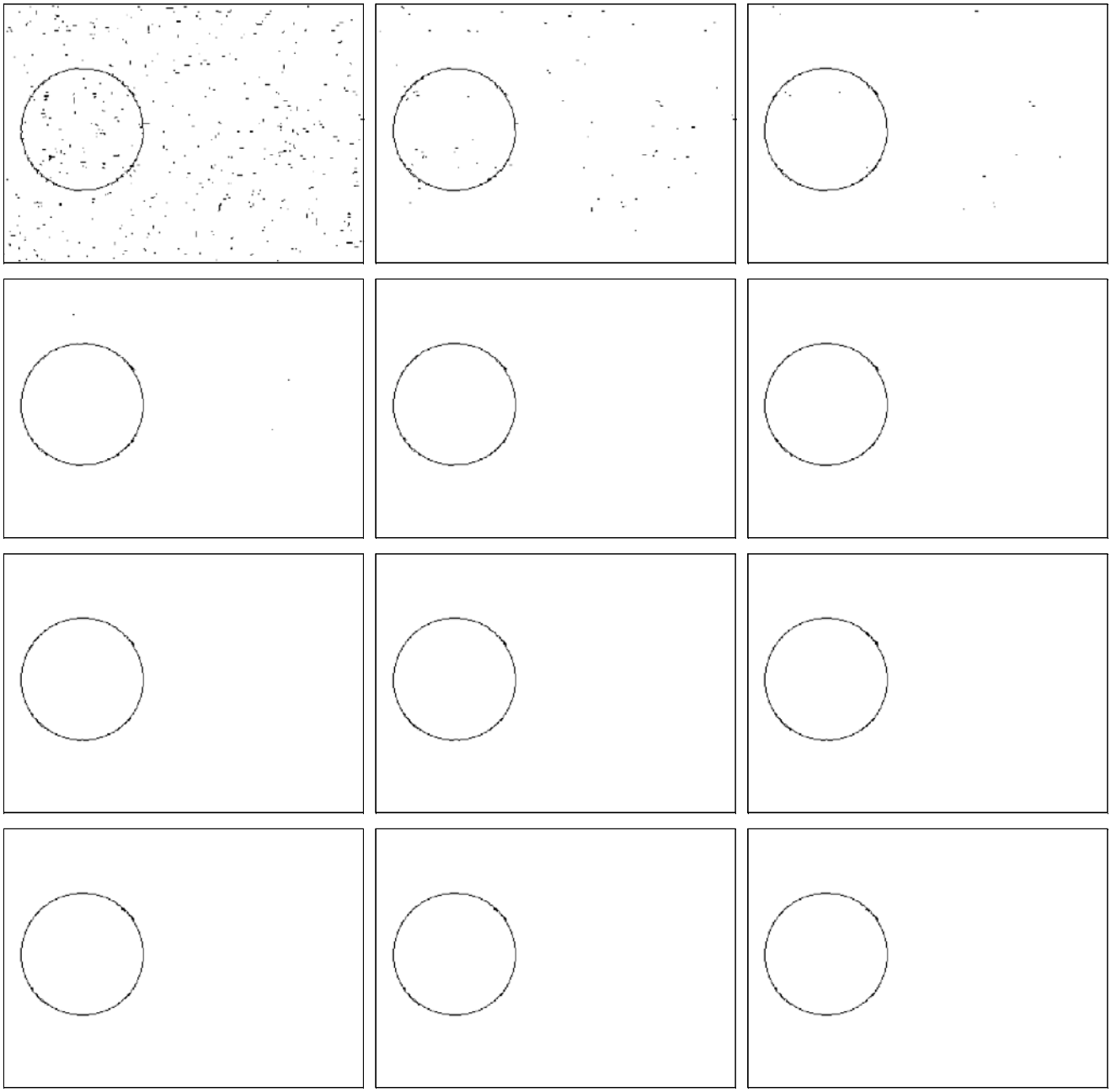


Figure 3.5 – Output of the algorithm.

Last simulated experiment was set up to handle deformations too. We considered once more the circle, with motion given by the composition of a sinusoidal translation along *x*-axis and of a deformation affecting the horizontal and the vertical axes by an oscillating gain. In particular we chose $P(t_i) = \begin{bmatrix} 1.1 & 0 \\ 0 & 0.9 \end{bmatrix}^{i-1}$, for $i \le M/3$ and $P(t_i) = \begin{bmatrix} 1.1 & 0 \\ 0 & 0.9 \end{bmatrix}^{M/3-1} \cdot \begin{bmatrix} 0.9 & 0 \\ 0 & 1.1 \end{bmatrix}^{i-M/3}$, otherwise. Degradation on the image, as in the first experiment, was supposed to be the result of a gaussian blur with covariance matrix constant in time and equal to $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, an additive noise with variance $\sigma_n^2(i) = 0.01$, $i = 1,2,\ldots,M$.

Also the same regular additive sinusoidal component was considered as in the first experiment. The twelve relevant frames are depicted in figure 3.6.
As to the output, we draw in figure 3.7 the detected edge for the generic frame. Wavelets used are the same as in the first experiment. This is possible because we suppose to know the motion and the detected edge points for the generic frame at time instant t_i can be obtained by simply propagating the detected edge points for the initial image. It's worth noticing once again the sudden decrease of both the metrics: the normalized Hausdorff distance between real and detected edge points goes from 38% down to 2.5% in six frames, while the normalized number of misplaced edge points goes from 8% down to 3% in four frames, as reported in figure 3.8.

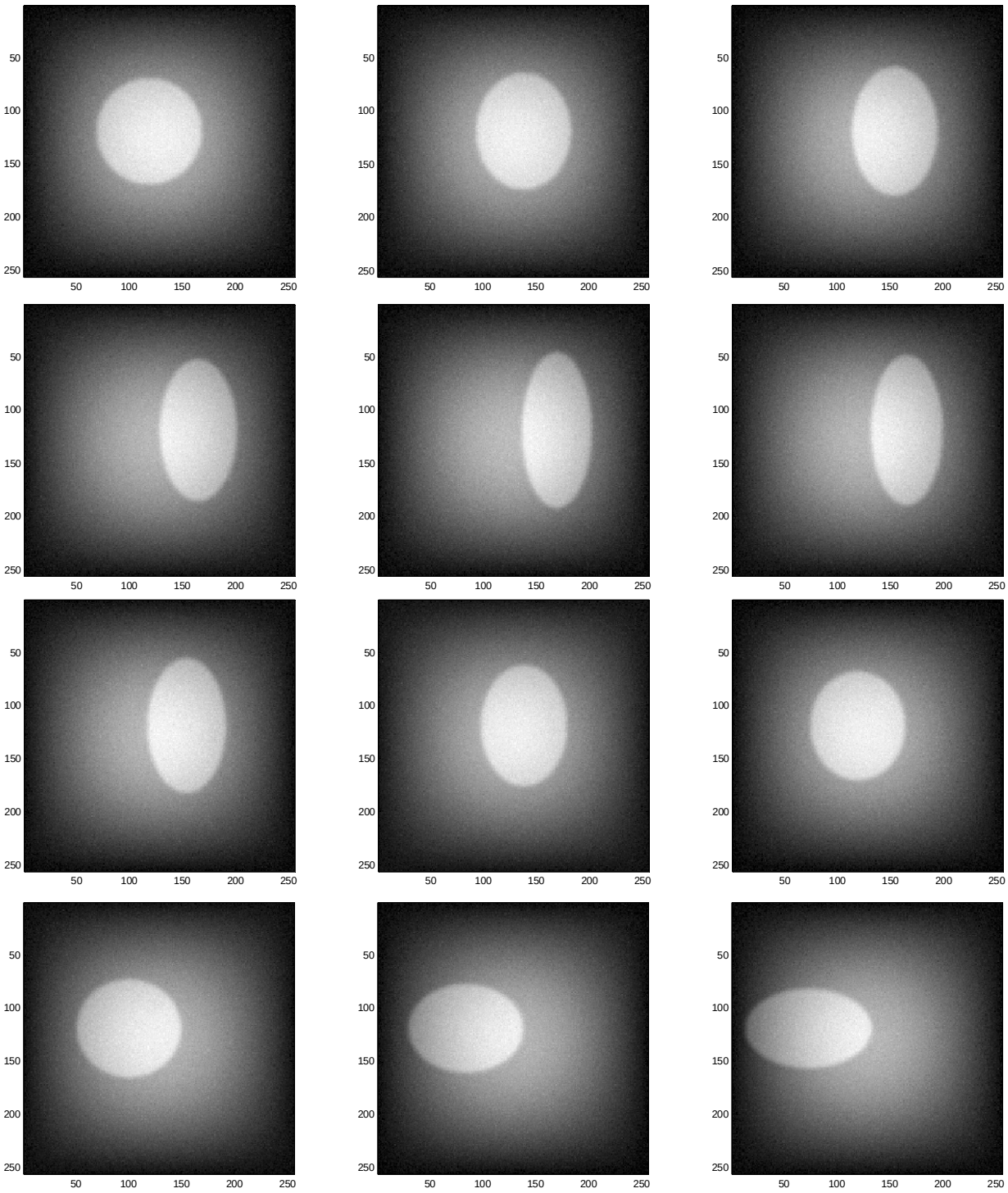


Figure 3.6 – The circle with sinusoidal deformation and translation.

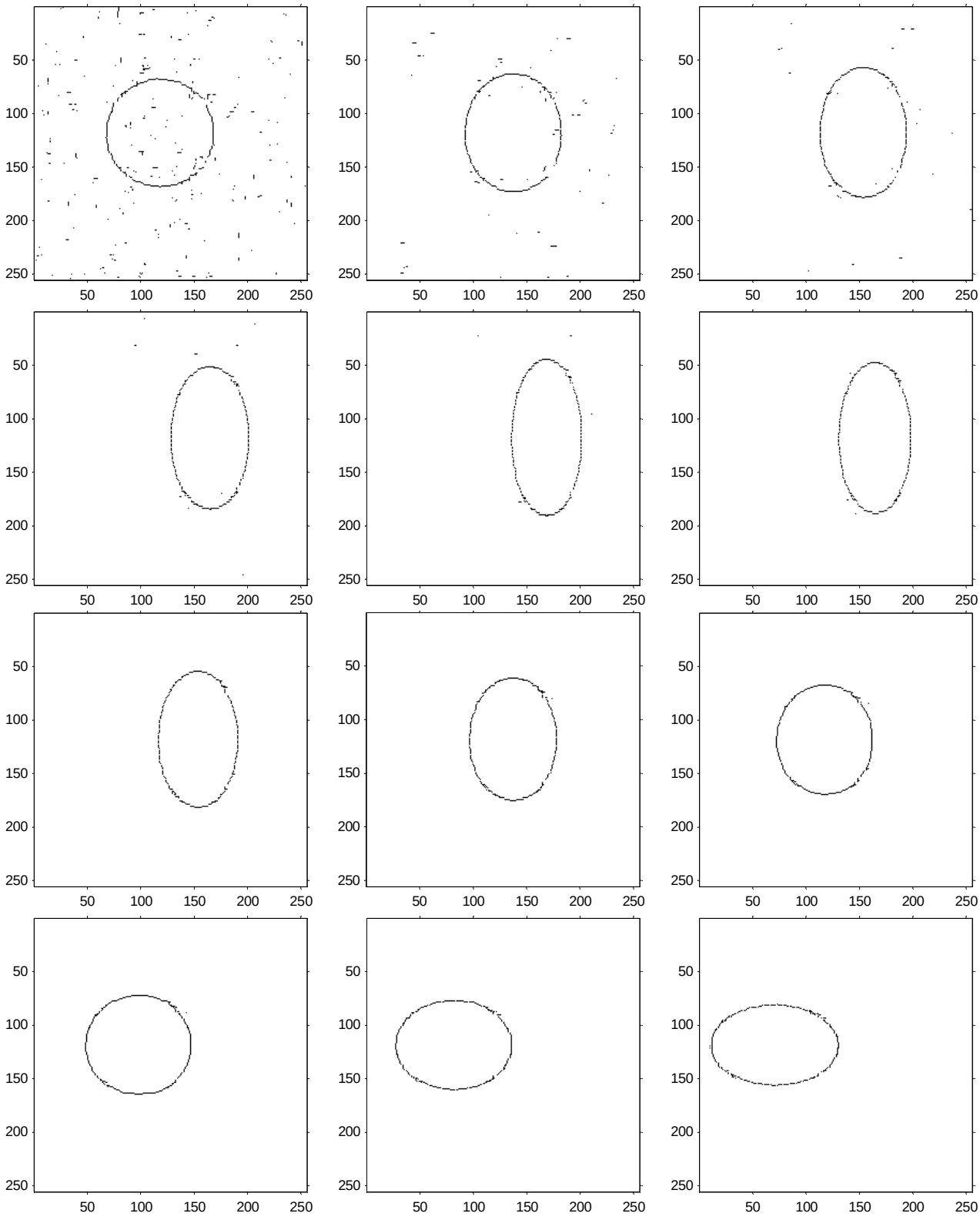


Figure 3.7 – Output of the algorithm, propagated along time by means of known motion.

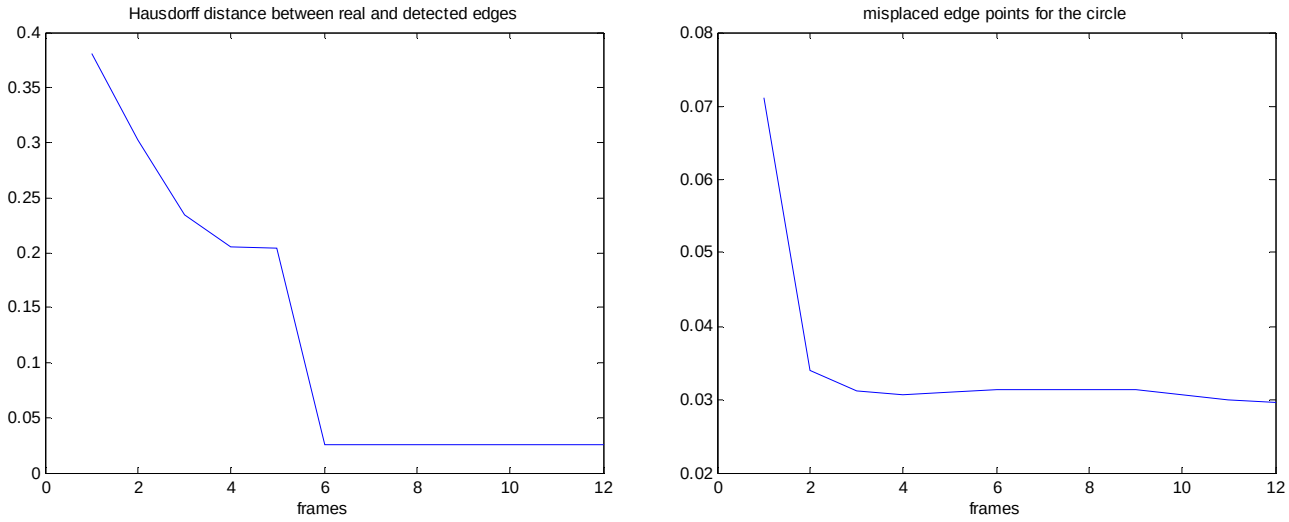


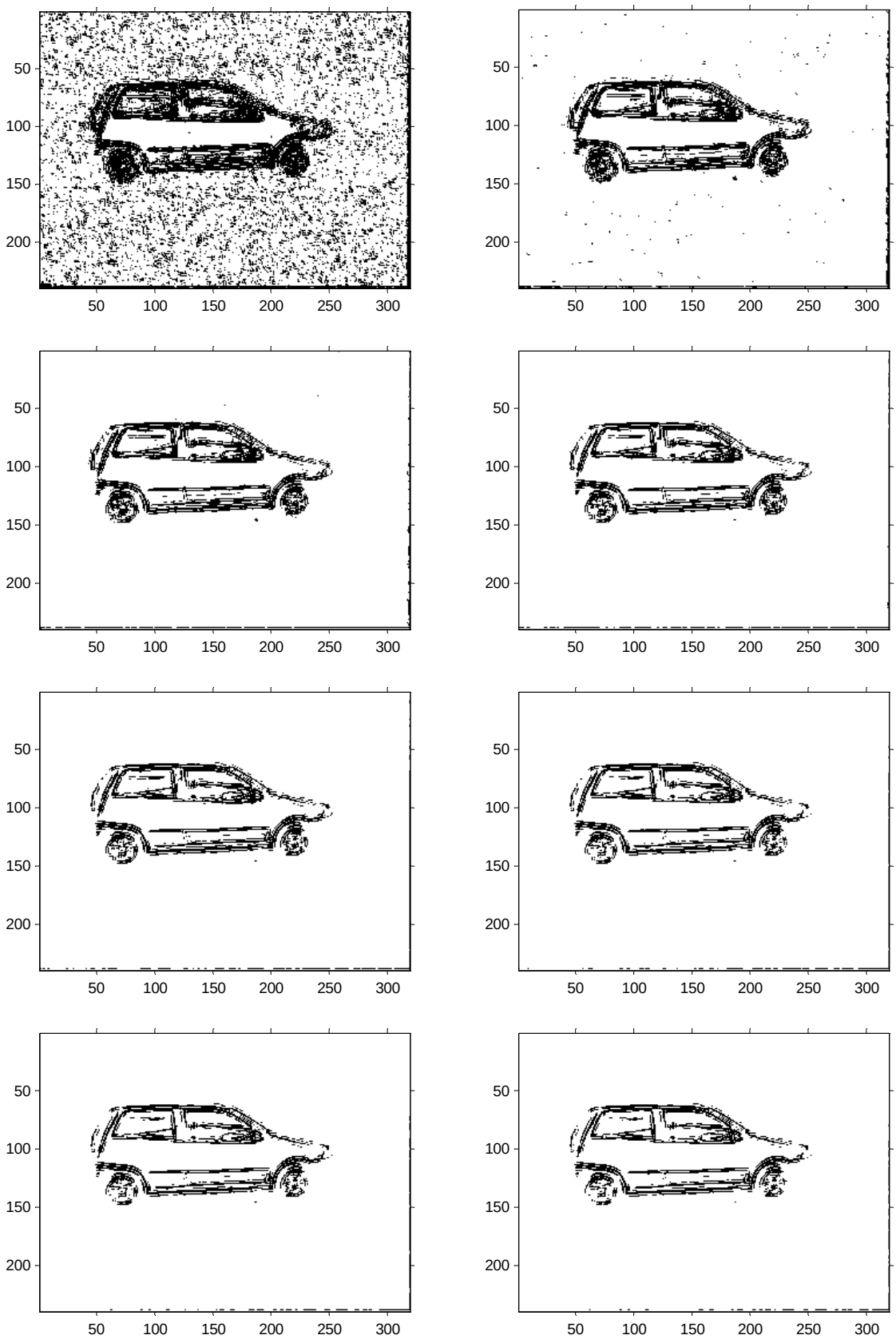
Figure 3.8 – Normalized Hausdorff metric between real and detected edge points and normalized number of misplaced edge points for the circle example.

The last experiment we want to show here concerns real acquired images. We shot a car moving on a black background using a common acquisition device, a web camera. One of the 10 frames we considered is shown in figure 3.9.

Figure 3.9 – The captured image for real data processing.

A preliminary task we have to accomplish in case real images are processed consists in estimating the degradation parameters (variance of the additive noise and covariance matrix of the gaussian blur) for the acquired images. To this purpose, we will introduce in Chapter 5 a relevant procedure. Nonetheless, we here suppose not to have such a procedure and to be able to give just a rough estimate of degradation parameters deduced by the characteristics of the web camera (in particular we assumed $\Sigma_b = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\sigma_n^2 = 0.0004$). This is because our aim is to investigate robustness property of the proposed methodology with respect to uncertainties in degradation parameters estimation. This would be a very important property in view of applications in real world contexts, because it would allow obtaining a suitable edge detection, even though ignoring the acquisition device characteristics. To this purpose such an experiment has been repeated several times with different values for the estimates of the image degradation parameters and the results obtained are very similar with each other. We present here (figure 3.10) output of one of these simulations. Wavelets have been chosen equal to 0.5, 0.75 and 1.0 respectively. The very high number of false alarms in the first

processed frame is due to the incorrectness of the degradation parameters estimation. Nevertheless, it's worth noting efficiency of the algorithm in drastically removing them, making the real edges emerge after few frames being processed.



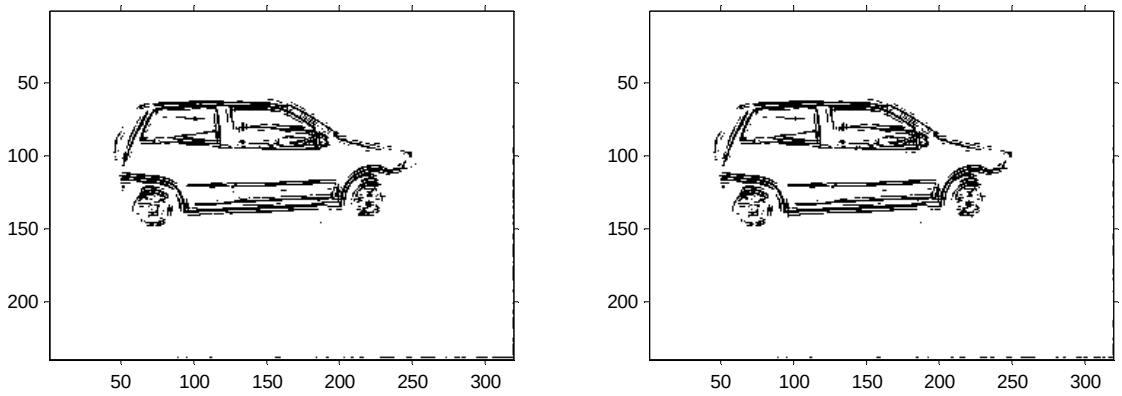


Figure 10 – Output of the algorithm when a real image sequence is considered.

4. Motion estimation from an image sequence

§ 4.1 – Problem set up

The problem we want to face here is the one of estimating the motion underlying a sequence of images, given observations characterized by blurring and additive noise, whose characteristics are known, as well as by the exact knowledge of the grey levels of the initial image. As we did in the previous chapter, we first cast the general framework given in Chapter 2 to the particular instance we are dealing with, by suitably specifying the measurement equation.

We assume that the grey level is measured at discrete times t_i, $i = 1,2,\ldots,M$, with $t_0 \le t_1 \le \ldots \le t_M \le t_f$.

With the already given notations, let $f_{t_i}(x)$ be the grey level of the image f at time t_i, $i = 1,2,\ldots,M$ and location $x \in \mathbb{R}^2$, and $z_i(u)$ the measured grey level at the same time and location $u \in \mathbb{R}^2$. As it is reasonable from an applicative point of view, in the following we shall assume f_{t_0} to have a compact support D in $\mathbb{R}^2$.

The grey level $f_{t_i}(x)$ can naturally be referred to the grey level in the original image f_{t_0} by:

$$f_{t_i}(x) = f_{t_0}\left(\varphi^{-1}(x; X(t_i))\right), \qquad x \in \mathbb{R}^2, \quad i = 1,2,\ldots,M \tag{4.1}$$

with $\varphi^{-1}(x; X(t))$ given by (2.25).

With the change of variable $x = \varphi(\xi; X(t_i))$, the measurement equation (2.28) becomes:

$$\begin{aligned} z_i(u) &= \int_{R^2} N_{\Sigma_b(t_i)}(u-x) f_{t_0}\left(\varphi^{-1}(x; X(t_i))\right) dx + n_i(u) = \\ &= \int_{R^2} N_{\Sigma_b(t_i)}(u - \varphi(\xi; X(t_i))) f_{t_0}(\xi) \left|\frac{\partial \varphi}{\partial \xi}(\xi; X(t_i))\right| d\xi + n_i(u) = \\ &= e^{X_9(t_i)} \int_{R^2} N_{\Sigma_b(t_i)}(u - \varphi(\xi; X(t_i))) f_{t_0}(\xi) d\xi + n_i(u), \end{aligned} \tag{4.2}$$

where we accounted for the Jacobian determinant $\left|\frac{\partial \varphi}{\partial \xi}(\xi; X(t_i))\right|$ to be constant in ξ and equal to:

$$\left|\frac{\partial \varphi}{\partial \xi}(\xi; X(t_i))\right| = |R(t_i)P(t_i)| = e^{X_9(t_i)} \tag{4.3}$$

as it easily follows from (2.6), (2.8), (2.15).

As said above, the problem is that one of on-line estimating the motion parameter vector $X(t)$, defined as in (2.19), for $t \in [t_0, t_f]$, with the dynamical model (2.18), given the original image $f_{t_0}(x)$, $x \in \mathbb{R}^2$ and the observations

$$z_\tau(u), \; u \in \mathbb{R}^2, \qquad t_\tau \in \{t_1,\ldots,t_M\} \cap [t_0, t] \tag{4.4}$$

measured up to time t, with $z_\tau(u)$ given by the measurement equation (4.2). This problem fits in the general framework of nonlinear filtering, due to the nonlinear character of the measurement equation.

§ 4.2 – Motion identifiability conditions

A preliminary issue with respect to the above problem concerns the question whether the motion is "identifiable", that is the question whether any two different motions may be distinguished from each other on the basis of the available observations in the absence of noise.
More precisely, for a given f_{t_0} and a fixed set of measurements $\{t_1,\ldots,t_M\}$, we define two motion parameter vectors X', X'' to be undistinguishable if they give rise to the same output (4.4) once we disregard the observation noise, that is, recalling (4.2):

$$e^{X'_9(t_\tau)}\int_{R^2} N_{\Sigma_b(t_\tau)}(u-\varphi(\xi;X'(t_\tau)))f_{t_0}(\xi)d\xi = e^{X''_9(t_\tau)}\int_{R^2} N_{\Sigma_b(t_\tau)}(u-\varphi(\xi;X''(t_\tau)))f_{t_0}(\xi)d\xi\,,$$
$$t_\tau\in\{t_1,\ldots,t_M\} \qquad (4.5)$$

The issue is relevant as it is clearly connected to the well posed ness property of the estimation problem.
The following sufficient identifiability results hold.

Proposition 4.1.
For a given f_{t_0} and for $t_\tau\in\{t_1,\ldots,t_M\}$ two motion parameter vectors X', X'' are undistinguishable if:

$$f_{t_0}\left(\varphi^{-1}(x;X'(t_i))\right)=f_{t_0}\left(\varphi^{-1}(x;X''(t_i))\right),\quad i=1,2,\ldots,M \qquad (4.6)$$

for almost all $x\in\mathbb{R}^2$.

Proof. By exploiting some previously established general results (Theorem 1 in [90]), X' and X'' are undistinguishable if, for each $t_\tau\in\{t_1,\ldots,t_M\}$, for any $x\in\mathbb{R}^2$ and for ξ', ξ'' such that

$$\varphi(\xi',X'(t_\tau))=\varphi(\xi'',X''(t_\tau))=x \qquad (4.7)$$

it is:

$$f_{t_0}(\xi')=f_{t_0}(\xi'') \qquad (4.8)$$

Solving (4.7) w.r.t. ξ', ξ'' yields (4.6). □

Given any two motion parameter vectors X', X'', respectively corresponding to the motion matrices (R',P',T'), (R'',P'',T'') according to the representation (2.7), we define the "difference" between X' and X'' as the motion represented by the parameter vector $\widetilde{X}$ corresponding to the matrices $(\widetilde{R},\widetilde{P},\widetilde{T})$, where:

$$\widetilde{T} = (R''P'')^{-1}[T' - T''],$$
$$\widetilde{R} = (R''P'')^{-1} R'P'\widetilde{P}^{-1}$$

and $\widetilde{P}$ is the "positive" definite square root of $(R''P'')^{-1} R'P'$.
In particular, if $X' = X''$, then $\widetilde{X}$ is the zero motion, corresponding to the matrices $\widetilde{R} = I$, $\widetilde{P} = I$, $\widetilde{T} = 0$, that is the absence of motion.

Proposition 4.2.
For a given f_{t_0} and for $t = t_M$ two motion parameter vectors X', X'' are undistinguishable if the motion parameter vector $\widetilde{X}$ and the zero motion are such.

Proof. It suffices to observe that, for $\chi = \varphi^{-1}(x; X'(t_i))$, equation (4.6) is equivalent to:

$$\begin{aligned} f_{t_0}(\chi) &= f_{t_0}\left(\varphi^{-1}(\varphi(\chi; X'(t_i))); X''(t_i)\right) = \\ &= f_{t_0}\left(\varphi^{-1}(R'(t_i)P'(t_i)\chi + T'(t_i)); X''(t_i)\right) = \\ &= f_{t_0}\left(\widetilde{R}(t_i)\widetilde{P}(t_i)\chi + \widetilde{T}(t_i)\right), \qquad i = 1,2,\ldots,M \end{aligned} \tag{4.9}$$

for almost all $\chi \in \mathbb{R}^2$.
A similar result clearly holds for the difference motion between X'' and X'. □

Corollary 4.3.
For a given image f_{t_0} and for $t = t_M$, the motion is identifiable if and only if no motion parameter vector $\overline{X}$ exists which is undistinguishable from the zero motion.

Proof. Immediate from above. □

There are noticeable instances of classes of images which might admit non identifiable motions for every $t = t_M$. Images with grey level contour lines invariant under classes of motion do not allow identification of those motions themselves. In particular:

1) images with contour lines parallel to a vector *T* do not allow identification of pure translation in the same direction;
2) images with circular contour lines with center in the origin do not allow identification of pure rotations;
3) images with radial or spyral-like contour lines do not allow identification of pure deformation $\widetilde{P}(t_i)$, with $\widetilde{P}(t_i) = K_i P$, or respectively of pure rotation-deformation $\widetilde{R}(t_i)\widetilde{P}(t_i)$, with $\widetilde{R}(t_i)\widetilde{P}(t_i) = K_i RP$ and the spyral-like contour lines invariant under *R P*;
4) constant grey level images do not allow identification of any motion whatsoever.

§ 4.3 – Filtering for motion estimation

§ 4.3.1 – General formulation

Let us suppose to have information available for time $t = t_i$ coming from measurements $z_j(u_k)$, where $u \in \mathbb{R}^2$, $j = 1,2,\ldots,i$ and $k = 1,2,\ldots,\overline{N}$, with $\overline{N} \le N$. Indeed, we are considering a spatial discretization of the image (pixels) besides the temporal one, due to the acquisition process. Further, we are assuming to consider only a subset of the image pixels with cardinality $\overline{N} \le N$. This choice seems reasonable since the dimensionality of the vector we are going to estimate (X) is significantly lower than the dimensionality of the available information (which is N, the number of pixels in the image) and it can be useful when implementing the algorithm for the filter, because it drastically reduces the computational load.
We now define $\overline{\mathcal{F}}_{t_i}^{z}$ the σ - algebra induced by this set of measurements. In order to obtain a suitable procedure for the motion parameter vector estimation, two steps are considered: prediction and updating.

As far as the prediction step is concerned, for $t_i \le t < t_{i+1}$, the motion parameter vector $X(t)$ evolves following the dynamical model (2.18). Consequently, its prediction estimate and the relevant covariance matrix for the estimation error are respectively given by:

$$\hat{X}(t \mid t_i) = E(X(t) \mid \overline{\mathcal{F}}_{t_i}^{z}), \tag{4.10}$$

$$\Psi_{\hat{e}}(t \mid t_i) = E\ [(\hat{X}(t \mid t_i) - X(t))(\hat{X}(t \mid t_i) - X(t))^T \mid \overline{\mathcal{F}}_{t_i}^{z}], \tag{4.11}$$

The following propagation equations hold:

$$d\hat{X}(t \mid t_i) = A(t)\hat{X}(t \mid t_i)dt + b(t)dt \tag{4.12}$$

$$d\Psi_{\hat{e}}(t \mid t_i) = A(t)\Psi_{\hat{e}}(t \mid t_i)A^T(t)dt + \Sigma(t)\Sigma^T(t)dt \tag{4.13}$$

with $A(t)$, $b(t)$ and $\Sigma(t)$ respectively defined in (2.20), (2.21) and (2.22). Initial conditions for (4.12) and (4.13) are respectively given by $\hat{X}(t_i \mid t_i)$ and $\Psi_{\hat{e}}(t_i \mid t_i)$.
In particular, should the probability distribution of $X(t_i)$ conditioned to $\overline{\mathcal{F}}_{t_i}^{z}$ be Gaussian, then the probability distribution of $X(t)$, $t_i \le t < t_{i+1}$, conditioned to $\overline{\mathcal{F}}_{t_i}^{z}$ would be Gaussian as well, so that it would be completely identified just by its mean value and the covariance matrix.

For $t = t_{i+1}$ we can proceed to updating distribution of $X(t)$ conditioned to $\overline{\mathcal{F}}_{t_i}^{z}$ by means of the Bayes formula:

$$p(X(t_{i+1}) \mid \overline{\mathcal{F}}_{t_{i+1}}^{z}) = p(X(t_{i+1}) \mid z_{i+1}, \overline{\mathcal{F}}_{t_i}^{z}) = \frac{p(z_{i+1} \mid X(t_{i+1}), \overline{\mathcal{F}}_{t_i}^{z})\, p(X(t_{i+1}) \mid \overline{\mathcal{F}}_{t_i}^{z})}{p(z_{i+1} \mid \overline{\mathcal{F}}_{t_i}^{z})}. \tag{4.14}$$

Hence, in order to update distribution $p(X(t_{i+1})|\overline{\mathcal{F}}_{t_i}^z)$ we should know both $p(z_{i+1}|X(t_{i+1}),\overline{\mathcal{F}}_{t_i}^z)$ and $p(z_{i+1}|\overline{\mathcal{F}}_{t_i}^z)$, which, generally speaking, can not be granted to be a finite dimensional problem.

With the purpose of obtaining a finite-dimensional filter, two ways will be proposed, which reduce the general formulation above to suitably approximated versions.

§ 4.3.2 – Gaussian approximation

The first approach consists in assuming for $p(X(t_i)|\overline{\mathcal{F}}_{t_i}^z)$ a Gaussian distribution, with mean value and covariance matrix respectively given by (4.10) and (4.11). As we said above, due to the linearity of the dynamical model of the motion, if $p(X(t)|\overline{\mathcal{F}}_{t_i}^z)$ is initialized with a Gaussian distribution, it keeps the same character and, as a consequence, $p(X(t_{i+1})|\overline{\mathcal{F}}_{t_i}^z)$ turns out to be Gaussian too. The distribution $p(X(t_{i+1})|\overline{\mathcal{F}}_{t_{i+1}}^z)$ can be then calculated by using (4.14). Even if this new distribution turns out to be non Gaussian any more, we can easily compute its mean value and its variance, by means of their very definitions. The approximation we consider here is to assume $p(X(t_{i+1})|\overline{\mathcal{F}}_{t_{i+1}}^z)$ as Gaussian with the above computed mean value and covariance matrix.
In particular, given $\overline{N} \le N$, let us introduce the vector

$$Z(t_{i+1}) = \begin{bmatrix} z_{i+1}(u_1) \\ \vdots \\ z_{i+1}(u_{\overline{N}}) \end{bmatrix}. \tag{4.15}$$

Thus, we can write the following vectorial measurement equation:

$$Z(t_{i+1}) = \mathcal{G}(X(t_{i+1})) + \nu(t_{i+1}), \tag{4.16}$$

where, recalling (4.2),

$$\mathcal{G}(X(t_{i+1})) = \begin{bmatrix} e^{X_9(t_{i+1})} \int_{R^2} N_{\Sigma_\delta(t_{i+1})}(u_1 - \varphi(\xi; X(t_{i+1})))f_{t_0}(\xi)d\xi \\ \vdots \\ e^{X_9(t_{i+1})} \int_{R^2} N_{\Sigma_\delta(t_{i+1})}(u_{\overline{N}} - \varphi(\xi; X(t_{i+1})))f_{t_0}(\xi)d\xi \end{bmatrix} \tag{4.17}$$

and

$$\nu(t_{i+1}) = \begin{bmatrix} n_{i+1}(u_1) \\ \vdots \\ n_{i+1}(u_{\overline{N}}) \end{bmatrix}. \tag{4.18}$$

We further define $\psi_\nu(t_{i+1})$ to be the covariance matrix of the noise vector $\nu(t_{i+1})$.

The first and second order momentums of the distribution $p(X(t_{i+1})|\overline{\mathcal{F}}^{z}_{t_{i+1}})$ can be then updated on the basis of their very definition:

$$\hat{X}(t_{i+1}\,|\,t_{i+1}) = E(X(t_{i+1})|\,\overline{\mathcal{F}}^{z}_{t_{i+1}}) = \\ = \frac{\int_{R^9} \xi \dfrac{e^{\left(-\frac{1}{2}(Z(t_{i+1})-\mathcal{C}_j(\xi))^T (Z(t_{i+1})-\mathcal{C}_j(\xi))\right)}}{(2\pi)^{N/2}|\psi_v(t_{i+1})|^{1/2}} \cdot \dfrac{e^{\left(-\frac{1}{2}\left(\xi-\hat{X}(t_{i+1}|t_i)\right)^T \psi(t_{i+1}|t_i)\left(\xi-\hat{X}(t_{i+1}|t_i)\right)\right)}}{(2\pi)^{9/2}|\psi(t_{i+1}\,|\,t_i)|^{1/2}} d\xi}{\int_{R^9} \dfrac{e^{\left(-\frac{1}{2}(Z(t_{i+1})-\mathcal{C}_j(\xi))^T (Z(t_{i+1})-\mathcal{C}_j(\xi))\right)}}{(2\pi)^{N/2}|\psi_v(t_{i+1})|^{1/2}} \cdot \dfrac{e^{\left(-\frac{1}{2}\left(\xi-\hat{X}(t_{i+1}|t_i)\right)^T \psi(t_{i+1}|t_i)\left(\xi-\hat{X}(t_{i+1}|t_i)\right)\right)}}{(2\pi)^{9/2}|\psi(t_{i+1}\,|\,t_i)|^{1/2}} d\xi}, \tag{4.19}$$

$$\Psi_{\hat{e}}(t_{i+1}\,|\,t_{i+1}) = \frac{\int_{R^9} \xi\xi^T \dfrac{e^{\left(-\frac{1}{2}(Z(t_{i+1})-\mathcal{C}_j(\xi))^T (Z(t_{i+1})-\mathcal{C}_j(\xi))\right)}}{(2\pi)^{N/2}|\psi_v(t_{i+1})|^{1/2}} \cdot \dfrac{e^{\left(-\frac{1}{2}\left(\xi-\hat{X}(t_{i+1}|t_i)\right)^T \psi(t_{i+1}|t_i)\left(\xi-\hat{X}(t_{i+1}|t_i)\right)\right)}}{(2\pi)^{9/2}|\psi(t_{i+1}\,|\,t_i)|^{1/2}} d\xi}{\int_{R^9} \dfrac{e^{\left(-\frac{1}{2}(Z(t_{i+1})-\mathcal{C}_j(\xi))^T (Z(t_{i+1})-\mathcal{C}_j(\xi))\right)}}{(2\pi)^{N/2}|\psi_v(t_{i+1})|^{1/2}} \cdot \dfrac{e^{\left(-\frac{1}{2}\left(\xi-\hat{X}(t_{i+1}|t_i)\right)^T \psi(t_{i+1}|t_i)\left(\xi-\hat{X}(t_{i+1}|t_i)\right)\right)}}{(2\pi)^{9/2}|\psi(t_{i+1}\,|\,t_i)|^{1/2}} d\xi} - \\ - E\left(X(t_{i+1})|\,\overline{\mathcal{F}}^{z}_{t_{i+1}}\right)E\left(X^T(t_{i+1})|\,\overline{\mathcal{F}}^{z}_{t_{i+1}}\right). \tag{4.20}$$

The new prediction step which follows is then obtained by the propagation of a Gaussian distribution with mean value given by (4.19) and covariance matrix given by (4.20).

Even if this method seems not to require hard approximations to work, it is worth to point out some reflection on the computational load it would imply.

Considering the only updating step, i.e. equations (4.19) and (4.20), we first notice that we should calculate three integrals on $\mathbb{R}^9$, which in turn implies a theoretical limit to do any calculation in a finite time. This problem can be easily overcome by suitably restricting this calculus to those regions where the two Gaussians into the integral are significantly away from zero. Furthermore, we already spoke about the possibility to only consider $\overline{N}$ pixels instead of the total number *N*, exploiting information redundancy, so that computational load is reduced.

Yet, there is another problem which cannot be overcome. Indeed, in order to compute the above integrals, we have to discretize the current differential variable (ξ). The more this discretization is fine, the better the discretized integral fits the original one. We now notice that we have to compute $\mathcal{C}_j(\xi)$, that is to say $\overline{N}$ blurring convolutions as in (4.17), for each different ξ we have to consider due to the above discretization in $\mathbb{R}^9$. Even if we considered a very rough discretization, say 10 intervals in each dimension for ξ, we would have to compute a very large number of convolution integrals (10^9 in this case), which is definitely unaffordable with the available technology. Unfortunately, non affordability considerations hold even if we might choose $\overline{N} << N$. A stochastic procedure to find suitable approximates for $\hat{X}(t_{i+1}\,|\,t_{i+1})$ and $\Psi_{\hat{e}}(t_{i+1}\,|\,t_{i+1})$ is currently being studied.

These considerations led us to study an alternative way of designing the filter.

§ 4.3.3 – Extended Kalman Filter

The other approach we considered is based on the well known Extended Kalman Filter algorithm. Assumption on $p(X(t_{i+1})|\overline{\mathcal{F}}_{t_i}^{z})$ to be Gaussian is preserved, so that the prediction step is once more implemented following equations (4.12), (4.13). The updating step is obtained by considering a linear approximation of the measurement equation (4.2). In particular, the expression of z_{i+1} is approximated to be linear with respect to $X(t_{i+1})$ in a neighbourhood of $\hat{X}(t_{i+1}|t_i)$, so that $p(X(t_{i+1})|\overline{\mathcal{F}}_{t_{i+1}}^{z})$ turns out to be Gaussian as well, with mean value and covariance matrix given by the Extended Kalman Filter equations.
In particular, from (4.2) we have:

$$z_{i+1}(u) = e^{X_9(t_{i+1})}\int_D N_{\Sigma_b(t_{i+1})}(u-\varphi(\xi;X(t_{i+1})))f_{t_0}(\xi)d\xi + n_{i+1}(u) \cong$$
$$\cong \alpha\left(u,\hat{X}(t_{i+1}|t_i)\right)+\beta\left(u,\hat{X}(t_{i+1}|t_i)\right)\left(X(t_{i+1})-\hat{X}(t_{i+1}|t_i)\right)+n_{i+1}(u), \qquad (4.21)$$

with:

$$\alpha\left(u,\hat{X}(t_{i+1}|t_i)\right)=z_{i+1}(u)\Big|_{X(t_{i+1})=\hat{X}(t_{i+1}|t_i)} = e^{\hat{X}_9(t_{i+1}|t_i)}\int_D N_{\Sigma_b(t_{i+1})}\left(u-\varphi\left(\xi;\hat{X}(t_{i+1}|t_i)\right)\right)f_{t_0}(\xi)d\xi, \quad (4.22)$$

$$\beta\left(u,\hat{X}(t_{i+1}|t_i)\right)=[\beta_1(u,\hat{X}(t_{i+1}|t_i)) \quad \cdots \quad \beta_9(u,\hat{X}(t_{i+1}|t_i))] =$$
$$= e^{\hat{X}_9(t_{i+1}|t_i)}\int_D N_{\Sigma_b(t_{i+1})}\left(u-\varphi\left(\xi;\hat{X}(t_{i+1}|t_i)\right)\right)\cdot$$
$$\cdot\left[\left(u-\varphi\left(\xi;\hat{X}(t_{i+1}|t_i)\right)\right)^T\Sigma_b^{-1}(t_{i+1})\frac{\partial\varphi(\xi;X(t_{i+1}))}{\partial X(t_{i+1})}\Big|_{X(t_{i+1})=\hat{X}(t_{i+1}|t_i)} + [0 \quad \cdots \quad 0 \quad 1]\right] f_{t_0}(\xi)d\xi \quad (4.23)$$

where, taking the affine model of motion (2.7) into account, we have:

$$\frac{\partial\varphi(\xi;X(t_{i+1}))}{\partial X(t_{i+1})}\Big|_{X(t_{i+1})=\hat{X}(t_{i+1}|t_i)} = \left[\left(\frac{\partial R(t_{i+1})}{\partial X(t_{i+1})}P(t_{i+1})+R(t_{i+1})\frac{\partial P(t_{i+1})}{\partial X(t_{i+1})}\right)\xi+\frac{\partial T(t_{i+1})}{\partial X(t_{i+1})}\right]\Bigg|_{X(t_{i+1})=\hat{X}(t_{i+1}|t_i)}. \quad (4.24)$$

With easy calculations one gets:

$$\frac{\partial\varphi(\xi;X(t_{i+1}))}{\partial X(t_{i+1})}\Big|_{X(t_{i+1})=\hat{X}(t_{i+1}|t_i)} = \left[\begin{pmatrix}-sinX_1 & -\cos X_1\\ \cos X_1 & -sinX_1\end{pmatrix}P(t_{i+1})\xi \quad \begin{matrix}1&0&0&0&0\\0&1&0&0&0\end{matrix} \quad \cdots\right.$$
$$\left. \cdots \quad R(t_{i+1})\begin{pmatrix}e^{X_7} & 0\\ 0 & -\dfrac{e^{X_9}+X_8^2}{e^{X_7}}\end{pmatrix}\xi \quad R(t_{i+1})\begin{pmatrix}0 & 1\\ 1 & +\dfrac{2X_8}{e^{X_7}}\end{pmatrix}\xi \quad R(t_{i+1})\begin{pmatrix}0 & 0\\ 0 & \dfrac{e^{X_9}}{e^{X_7}}\end{pmatrix}\xi\right]\Bigg|_{X(t_{i+1})=\hat{X}(t_{i+1}|t_i)}$$

Let us now introduce the $\overline{N}$-dimensional vector

$$a\left(\hat{X}(t_{i+1}\mid t_i)\right)=\begin{bmatrix}\alpha\left(u_1,\hat{X}(t_{i+1}\mid t_i)\right)\\ \vdots\\ \alpha\left(u_{\overline{N}},\hat{X}(t_{i+1}\mid t_i)\right)\end{bmatrix}$$

and the $\overline{N}\times 9$ matrix

$$B\left(\hat{X}(t_{i+1}\mid t_i)\right)=\begin{bmatrix}\beta\left(u_1,\hat{X}(t_{i+1}\mid t_i)\right)\\ \vdots\\ \beta\left(u_{\overline{N}},\hat{X}(t_{i+1}\mid t_i)\right)\end{bmatrix}.$$

The approximation we introduced in (4.21) reflects on the global measurement equation (4.16), which is in turn approximated as follows:

$$Z(t_{i+1})\cong a\left(\hat{X}(t_{i+1}\mid t_i)\right)+B\left(\hat{X}(t_{i+1}\mid t_i)\right)\left(X(t_{i+1})-\hat{X}(t_{i+1}\mid t_i)\right)+\nu(t_{i+1}). \tag{4.25}$$

This linear measurement equation can now be used to write the equations for the updating step of the estimation procedure. We have:

$$\hat{X}(t_{i+1}\mid t_{i+1})=\hat{X}(t_{i+1}\mid t_i)+K(t_{i+1},\hat{X}(t_{i+1}\mid t_i))[Z(t_{i+1})-a(\hat{X}(t_{i+1}\mid t_i))] \tag{4.26}$$

$$\Psi_{\hat{e}}^{-1}(t_{i+1}\mid t_{i+1})=\psi^{-1}(t_{i+1}\mid t_i)+B^T\left(\hat{X}(t_{i+1}\mid t_i)\right)\psi_\nu^{-1}(t_{i+1})B\left(\hat{X}(t_{i+1}\mid t_i)\right) \tag{4.27}$$

where, according to the Schur's Lemma, the gain K of the filter is given by:

$$K(t_{i+1},\hat{X}(t_{i+1}\mid t_i))=\psi(t_{i+1}\mid t_i)B^T\left(\hat{X}(t_{i+1}\mid t_i)\right)\left[B\left(\hat{X}(t_{i+1}\mid t_i)\right)\psi(t_{i+1}\mid t_i)B^T\left(\hat{X}(t_{i+1}\mid t_i)\right)+\psi_\nu(t_{i+1})\right]^{-1}=$$

$$=\psi(t_{i+1}\mid t_i)B^T\left(\hat{X}(t_{i+1}\mid t_i)\right)\cdot$$

$$\cdot\left[\psi_\nu^{-1}(t_{i+1})-\psi_\nu^{-1}(t_{i+1})B\left(\hat{X}(t_{i+1}\mid t_i)\right)\psi(t_{i+1}\mid t_{i+1})B^T\left(\hat{X}(t_{i+1}\mid t_i)\right)\psi_\nu^{-1}(t_{i+1})\right]. \tag{4.28}$$

Implementation of the Extended Kalman Filter algorithm is currently being completed and tests both on simulated and real data have been planned.

5. Application to pupillometric data

§ 5.1 – Introduction

The pupil of human eyes fluctuates either to adapt the amount of light to the retina or when a subject is gazing at a fixed object in absence of light stimulation and/or visual accommodation. Multiple nervous paths and structures are involved in the control of the iris muscles [91]; thus, the measure of pupil diameter dynamics has several clinical applications.
Pupil fluctuation seems to be a rather sensitive indicator of mental activity [92]; the reflex pupillary constriction to light accounts for functional abnormality in Alzheimer's disease [93]; moreover in the analysis of sleep disorder and in the analysis of pupil response to light stimuli important studies show relation with the variations of the pupil size [94]; another relevant applicative field is the study of correlation between the variation of the pupil diameter and the effects of consuming medicines as well as drugs and alcohol [95]. Recent studies show some correlation between the human fatigue during Visual Display Terminal (VDT) tasks and variation in the pupil diameter [96]; finally, in video-based systems the determination of pupil centre is significant for calculating eye orientation [97].
A crucial point for many of the previous applications is represented by a reliable estimation of the pupil edge. To this purpose, most of studies of the pupil assume a circular shape and calculate only its diameter or its area. Scarce attention has been paid to develop algorithms for automated estimation of other types of shape, accounting for more accurate analysis, such as departures from circularity and concentricity [98]. The first attempt to characterize the pupil as an ellipse has been made by Kristek [99] who measured the longest and shortest diameters. Recently great importance has been devoted to the relation between some pathologies and departure from circularity, since this feature accounts for anisotropies of structure and/or innervations; furthermore, it has been suggested that pupils sometimes ovalize with increasing age [100].
In the following we assume an elliptic shape for the pupil and the aim we want to pursue is to identify some relevant parameters, such as the location of centre of the pupil, the length of the two semi axes and the amplitude of the rotational angle.
We assume to perform measures by a suitable CCD based device, called pupillometer. Gaussian blurring and additive white noise are supposed to corrupt the measured image, so that the first step of our analysis consists in identifying the parameters which describe this degradation. We then detect the edge points of the pupil by means of a suitable specification of the multiscale processing introduced in Chapter 3, with restriction to a single image with null motion. Last, we propose an optimal procedure to detect the above morphological parameters. The proposed procedure turns out to be very reliable and the produced results are good with a short execution time, which is an important characteristic too in view of performing such a processing over a sequence of frames.
The extension to sequence processing is also implemented, following the general approach described in Chapter 3, in order to reduce the amount of information previously needed to apply the procedure to a single frame. In particular, the lack of exact knowledge about degradation parameters is assumed and encouraging results are presented at the end of this chapter. As already said in Chapter 3, we think that this is one of the main potentialities of the proposed methodology of image sequence processing, which should be exploited in the future. In fact, it's worth to point out that generally all the methods implemented in literature to deal with real images, require knowledge of the degradation parameters introduced by the acquisition process, typically the blur and the additive noise and the relevant procedure to identify them usually constitutes a necessary preliminary step that has to be performed before processing the images [101]-[107]. Through recursive filtering of the image sequence,

instead, we get rid of this first step and we are able anyhow to obtain good edge estimates just with few frames being processed.

§ 5.2 – Characterization of the pupillometer

The pupillometer was first introduced in [108]. It consists of a micro infrared CCD camera mounted on a light helmet. The camera is equipped with an infrared filter to eliminate reflexes from the natural light source in the room; the proper illumination is provided by a near infrared diode (780 nm). A video capture board (IMAQ PCI 1408, National Instruments) set with a frame rate of 8.33 frames/s and a resolution of 768×576 pixels, 256 grey levels is used. An example of measured image is shown in figure 5.1. With such a frame rate, in the framework of the application we are dealing with, it is possible to disregard the presence of motion blur; furthermore the set up of the acquisition process is designed in such a way that the region of interest (i.e. the pupil) turns out to be not too close to the limits of the image, so that we can also ignore the presence of distortion due to boundary effects.

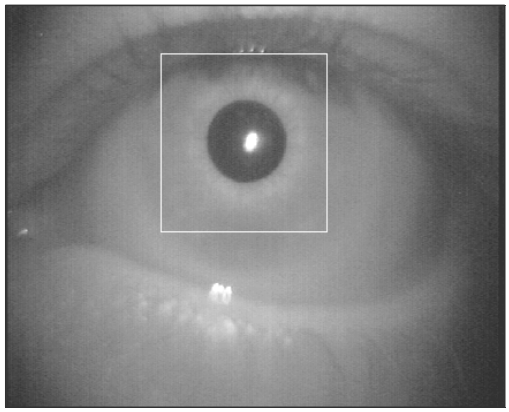


Figure 5.1 – A sample of the frames captured by pupillometer

In general, the characterization of the blur effects introduced by an optical imaging system, as the CCD camera, can be a very difficult task. If one should have to construct a model based upon physical optics (white box model), it would be necessary the exact knowledge of such parameters as depth of the imaged objects, lens aberrations and so forth [100]. This kind of study, of course, often results impractical and unfeasible, so that usually in image processing a blind estimation method is introduced, that's to say that the blur is estimated directly from the degraded images, after a suitable modelling, independent from the acquisition system (black box model).
For our purpose we chose to model the measured image following the framework introduced in Chapter 2, considering Gaussian blur and additive white noise. As we said, we here suppose that no extra degradations significantly affect the acquired image, so that our model turns out to be coherent with the methodology we proposed in previous chapters, as well as efficient for future more complex applications (i.e. analysis of temporal sequence of frames).
The model we assume for the measured data z is a discretized version of (2.28):

$$z(i,j) = y(i,j) + n(i,j) = \sum_{(h,k)\in F} f(h,k) N_{\Sigma_d}(i-h, j-k) + n(i,j), \quad (i,j) \in F \tag{5.1}$$

where y is the blurred signal, described as the convolution of the original grey level f with a zero mean 2-D Gaussian kernel N_{Σ_d}, n is an additive white Gaussian noise, with zero mean

and covariance matrix $\sigma_n^2 I$, and F denote the set of vertexes of the grid G (cfr. Definition 2.6). The covariance matrix Σ_d will be assumed to be diagonal, with $\Sigma_d = diag(\sigma_d^2)$, and $\sigma_d^2 > 0$, so that the following factorisation holds: $N_{\Sigma_d}(i,j) = N_{\sigma_d^2}(i) N_{\sigma_d^2}(j)$, $(i,j) \in F$; this assumption is introduced for sake of simplicity but does not essentially affect the results.

§ 5.3 – Estimation of degradation parameters

The edge detection procedure for a single image with null motion can be obtained by a suitable particularization of the general methodology described in Chapter 3. A preliminary task to be accomplished, as we said, is providing knowledge about the image degradation parameters, which in the case of our model are the variances of the Gaussian blur σ_d^2 and of the additive noise σ_n^2. To this purpose, we propose here a simple procedure which only needs measured data, instead of characteristics of the measurement process and of the acquisition device.
We consider a target image, consisting in our case of a 2×2 checkerboard (figure 5.2), shot from the CCD camera in the same conditions in which we will then capture the pupil image.
From the analysis of this image we note that, probably due to the automatic gain control of the camera, the level of additive noise *n* weakly depends on the value of the signal *f*; hence, we consider *S* homogeneous regions Λ_k of the target image and we compute the estimation of σ_n^2. Assuming the mean value of $n(i,j)$ to be zero for each Λ_k, we have:

$$\bar{z}_k = \frac{1}{N_k} \sum_{l=1}^{N_k} z\left(x_l^{(k)}\right) \cong y^{(k)} = f^{(k)},$$

where $x_l^{(k)}$ are N_k points in the homogeneous region Λ_k away from Λ_k's boundaries and $f^{(k)}$ is the corresponding constant value of the grey level. Thus, an estimate of the variance is given by:

$$\hat{\sigma}_n^2 = \frac{1}{S} \sum_{i=1}^{S} \left[\sum_{l=1}^{N_k} \left[z\left(x_l^{(k)}\right) - f^{(k)} \right]^2 / N_k \right]. \tag{5.2}$$

To estimate the variance of the blur σ_d^2 we consider once more the target image. If we consider a line *L* of this image, either horizontal or vertical, it can be characterized by an orthogonal grey level jump of amplitude *M* at the abscissa $\tau = \tau_0$. Thus, a local 1-*D* model for the grey level along *L* is $z_L(\tau) = c + M\delta_{-1}(\tau - \tau_0)$, $c > 0$. Then the model of the measure along *L* turns out to be:

$$(z_L * N_{\sigma_d^2})(\tau) = \{[c + M\delta_{-1}(\bullet - \tau_0)] * N_{\sigma_d^2}\}(\tau) + n(\tau), \qquad \tau \in L.$$

where the symbol $*$ denotes convolution.
Finally, the parameters $(c, M, \tau_0, \sigma_d^2)$ can be numerically determined minimizing a suitable integral norm of the difference between the above 1-*D* model $z_L * N_{\sigma_d^2}$ and the measured data ς:

$$\min_{(c,M,\tau_0,\sigma_d^2)} \left\| \varsigma - z_L * N_{\sigma_d^2} \right\|, \tag{5.3}$$

Several tests both on simulated and real data showed very good matching between the model and the measure. It is worth noting that in order to calibrate the camera we already need a target image to be shot so that the proposed estimation procedure does not require any further acquisition.

§ 5.4 – Single frame processing

§ 5.4.1 – Edge detection for a single frame

A preliminary approach to edge detection for the image sequence coming from the pupillometer consists in analysing each different frame separately. To do that, we have to cast the general framework given in Chapter 3 to the particular case of a single image with null motion. Moreover, dealing with real images acquired by a digital device, also discretization of measured data must be taken into account. The consequent particularization of the proposed approach for edge detection is sketched hereafter.

Due to the assumption about the presence of the edges on the grid *G,* we separate the analyses along the two coordinate directions. More precisely, let us denote by $\theta^{(1)}(i,j)$:

$$\theta^{(1)}(i,j) = f(i,j) - f(i,j-1), \qquad (i,j) \in F,\ (i,j-1) \in F \tag{5.4}$$

the grey level variation at the pixel (i,j) along the horizontal axis. We denote by N_{σ^2} the 1-D Gaussian with zero mean and variance σ^2 and with $N_{\sigma^2}^{(k)}$ its *k*-th derivative. From equation (5.1) we have:

$$z(i,j) - z(i,j-1) = \sum_{(h,k) \in F} N_{\sigma_d^2}(i-h) N_{\sigma_d^2}(j-k) \theta^{(1)}(h,k) + \eta(i,j), \qquad (i,j) \in F \tag{5.5}$$

where $\eta(i,j) = n(i,j) - n(i,j-1)$.

Pre-processing of the observed data *z* is obtained by convoluting them along the grid axes by kernels $N_{\sigma_w^2}^{(2)}$, $w = 1,2,...,W$. This convolution is sensitive to discontinuities locally orthogonal to the convolution direction; therefore, along the horizontal side of the grid, for different choices of the wavelets, from (3.7) we have:

$$\tilde{z}_w^1(i,j) = \sum_k z(i,k) N_{\sigma_w^2}^{(2)}(j-k) \qquad (i,j) \in F, \qquad w = 1,2,\ldots,W \tag{5.6}$$

For $(i,j) \in F$, reproducing calculations in (3.19), from (5.5) it follows:

$$\tilde{z}_w^{(1)}(i,j) - \tilde{z}_w^{(1)}(i,j-1) = \sum_{(h,k) \in F} N_{\sigma_d^2}(i-h) N_{\sigma_d^2+\sigma_w^2}^{(2)}(j-k) \theta^{(1)}(h,k) + \tilde{\eta}_w^{(1)}(i,j)$$

where

$$\tilde{\eta}_w^{(1)}(i,j) = \sum_k N_{\sigma_w^2}^{(2)}(j-k)\eta(i,k).$$

Now we exploit the fact that $\theta^{(1)}(h,k)$ does not significantly vary with respect to *h*, for *h* in the range $(i-3\sigma_d, i+3\sigma_d)$, where $N_{\sigma_d^2}(i-h)$ is significantly different from zero (cfr. Hypothesis 1, Section 3.3). Furthermore, observing that $\theta^{(1)}(i,k)$ is zero inside the grid and it is equal to $\theta^{(1)}(i,j)$ on the only value of k where $N^{(2)}_{\sigma_d^2+\sigma_w^2}(j-k)$ is significantly non-zero, that is for $k=j$, we approximate:

$$\tilde{z}_w^{(1)}(i,j) - \tilde{z}_w^{(1)}(i,j-1) \cong \sum_k N^{(2)}_{\sigma_d^2+\sigma_w^2}(j-k)\theta^{(1)}(i,k) + \tilde{\eta}_w^{(1)}(i,j) \cong N^{(2)}_{\sigma_d^2+\sigma_w^2}(0)\theta^{(1)}(i,j) + \tilde{\eta}_w^{(1)}(i,j),$$
$$w = 1,2,\ldots,W \qquad (5.7)$$

Defined the following *W*-size vectors:

$$q^{(1)}(i,j) = \begin{pmatrix} \tilde{z}_1^{(1)}(i,j) - \tilde{z}_1^{(1)}(i,j-1) \\ \vdots \qquad\qquad \vdots \\ \tilde{z}_W^{(1)}(i,j) - \tilde{z}_W^{(1)}(i,j) \end{pmatrix}, \qquad u = \begin{pmatrix} N^{(2)}_{\sigma_d^2+\sigma_1^2}(0) \\ \vdots \\ N^{(2)}_{\sigma_d^2+\sigma_W^2}(0) \end{pmatrix}, \qquad \tilde{\eta}^{(1)}(i,j) = \begin{pmatrix} \tilde{\eta}_1^{(1)}(i,j) \\ \vdots \\ \tilde{\eta}_W^{(1)}(i,j) \end{pmatrix}$$

equation (5.7) becomes:

$$q^{(1)}(i,j) = u \cdot \theta^{(1)}(i,j) + \tilde{\eta}^{(1)}(i,j) \qquad (5.8)$$

We note that $\tilde{\eta}^{(1)}(i,j)$ is a Gaussian noise vector, with zero mean and covariance matrix:

$$\Psi = E[\tilde{\eta}^{(1)}(i,j)\tilde{\eta}^{(1)^T}(i,j)] = \{\psi_{st}\},$$

where, with some computation and recalling (3.32):

$$\psi_{st} = E[\tilde{\eta}_s^{(1)}(i,j)\tilde{\eta}_t^{(1)}(i,j)] = 2\sigma_n^2\left(N^{(4)}_{\sigma_s^2+\sigma_t^2}(0) - N^{(4)}_{\sigma_s^2+\sigma_t^2}(1)\right), \qquad s,t = 1,2,\ldots,W$$

From (5.8) we get the Markov estimate of $\theta^{(1)}(i,j)$, [109]:

$$\hat{\theta}^{(1)}(i,j) = \frac{u^T\Psi^{-1}q^{(1)}(i,j)}{u^T\Psi^{-1}u} = \theta^{(1)}(i,j) + \frac{u^T\Psi^{-1}\tilde{\eta}^{(1)}(i,j)}{u^T\Psi^{-1}u} \qquad (5.9)$$

$\hat{\theta}^{(1)}(i,j)$ is a Gaussian variable with mean $\theta^{(1)}(i,j)$ (unbiased estimate) and with variance $[\sigma^{(1)}(i,j)]^2 = \left(u^T\Psi^{-1}u\right)^{-1}$, independent of (i,j) (minimum variance estimate).

With obvious meaning of notation, we obtain for the estimate of the discontinuous variation $\theta^{(2)}(i,j)$ of *f* along the vertical side of the grid and its variance:

$$\hat{\theta}^{(2)}(i,j)=\frac{u^T\Psi^{-1}q^{(2)}(i,j)}{u^T\Psi^{-1}u}, \qquad [\sigma^{(2)}(i,j)]^2=\left(u^T\Psi^{-1}u\right)^{-1}.$$

Knowledge of the parameters σ_d^2 and σ_n^2 is supposed in calculating the vector u and the matrix Ψ respectively.
Now at each point $(i,j)\in F$ we have the estimates of the possible discontinuities, along the horizontal and the vertical direction. As we said in Section 3.5, information about the presence of an edge can be obtained if the size of its estimate is sufficiently larger than the standard deviation of this estimate; then, in order to decide if the estimated jump on a vertex of the grid represents a true discontinuity or a false alarm, we use a statistical test of simple hypothesis ([109]). Obviously, this test is possible because of the Gaussian nature of the obtained estimations. We define two hypotheses:

$\mathcal{H}_0$: $\theta^{(1)}(i,j)=0$;
$\mathcal{H}_1$: $\left|\theta^{(1)}(i,j)\right|\geq\Xi$, where $\Xi>0$ is a fixed threshold.

On the basis of the information obtained by the estimation procedure, if $\hat{\theta}^{(1)}(i,j)\notin[-\Xi,\Xi]$ we reject Hypothesis $\mathcal{H}_0$, otherwise we reject Hypothesis $\mathcal{H}_1$; the threshold Ξ is chosen in such a way that the probability of rejecting hypothesis H_0 when it is true is equal to some fixed $\varepsilon\%$.
In the same way we define an analogous test on $\hat{\theta}^{(2)}(i,j)$.

The described procedure for edge detection requires the choice of the number of wavelets to use and of their variances. To that purpose there is no general rule, even if some advice can be emphasized. Generally speaking, a too small variance could filter the noise not enough, while, a too high variance could smooth the signal too much, vanishing the effect of enhancing the contribution of the discontinuous component over the regular part. Moreover, it is useful that the variances of the wavelets be not too similar with each other, in order to enjoy advantages of the multiscale approach. About the number of wavelets to employ, experience suggests using three or four different wavelets is enough, as it has been observed that a larger number does not lead to significant improvements.

The detected edge points, with the probability fixed in the hypothesis test, are close to the true edge, which is to say that their maximum distance from the edge is less than the length Δ of the grid. It is possible to choose the size of the grid Δ as small as the pixel resolution allows; the possibility of considering a bigger grid size can allow a minor computational cost, at the expense of a lower number of identified points.

§ 5.4.2 – Clustering and optimal fitting

As already pointed out in the Introduction of this chapter, it is reasonable to assume an elliptic shape for the pupil. The ellipse can be completely characterized by five parameters, the two coordinates of the centre, the length of the two semi axes and the rotational angle, describing its orientation in the plane. Our aim will be therefore to identify these five quantities starting from the previously detected edge points. In particular, identification time is an important issue to be considered in order to study fluctuations of parameters over a sequence of frames.
While processing the detected edge points, we first notice the presence of two distinct clusters of points, one concerning the real edge between the pupil and the iris and the other deriving

from the reflection of the camera light diode within the pupil. Since we cannot eliminate this white spot within the pupil in any artificial way during acquisition, we have to cluster the detected points, in order to discard the second group and to fit only the first. The simplest heuristic to accomplish this task turned out to be the introduction of a threshold after a polar coordinates description of the figure. In fact, considering the polar coordinates reference system (phase vs. radius) centred in the diode light area (which is in turn quickly detectable considering its grey average value), the two elliptic groups of points belong to two well separate distributions along two horizontal lines; then by mean of a threshold, we can easily isolate the only group of points we need to process in the following.
Successively, in order to evaluate the morphological parameters of the ellipse we perform an optimal interpolation of the selected points, by minimizing the sum of the absolute values of the fitting residues:

$$\min_{(a,b,u,v,\alpha)} \sum_{l=1}^{P} \Big| \left(a^2 \sin^2\alpha + b^2 \cos^2\alpha\right)(m_l - u)^2 + \\ + \left(a^2 \cos^2\alpha + b^2 \sin^2\alpha\right)(n_l - v)^2 + \\ + \left(b^2 - a^2\right)\sin 2\alpha (m_l - u)(n_l - v) - a^2 b^2 \Big| \tag{5.10}$$

where (m_l, n_l), $l = 1, \ldots, P$ are the coordinates of the P previously selected points on F, a and b are the semi axes of the ellipse, (u, v) are the coordinates of its centre and α is the rotation angle. The numerical procedure used to find the minimum is *fminsearch*, a standard Matlab routine, which implements the Nelder-Mead method.
In this way, what we find is the optimal estimates of the pupil's characteristic parameters, which are the quantities of interest in all the applications mentioned in Section 5.1.

§ 5.4.3 – Results of single frame processing

The whole algorithm presented above is made up of three steps:

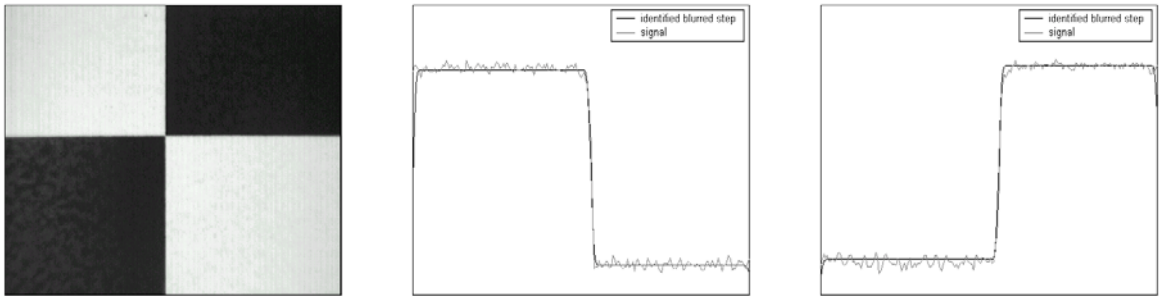
- degradation parameters estimation;
- edge points detection;
- clustering and optimal fitting.

Preliminary to application on real data, it's worth to test them on suitable simulated data.

In order to validate the procedure for the estimation of the degradation parameters, we consider a target image (a 2×2 checkerboard) shot by the camera (figure 5.2(A)). This image has been chosen also to investigate if the camera presents isotropic blurring effects (i.e. the same blur both on horizontal and vertical directions and when crossing from white to black and from black to white). In Table 5.1 we report the average values of the blur estimations.

BLUR ESTIMATION	Vertical	Horizontal
White to Black	1.5777	1.4148
Black to White	1.5093	1.5324

Table 5.1 – Blur estimation analysis on the target checkerboard image shot by the camera



(A) (B) (C)

Figure 5.2 - Identification of parameters of degradation for the target image: (A) the image, horizontal line: white-to-black (B) and black-to-white (C) blur identification.

We notice that there is no considerable difference on the estimated blur value neither when considering horizontal or vertical directions, nor when considering white to black or black to white crossing. Figures 5.2(B) and 5.2(C) represent a horizontal line when passing from white to black and from black to white respectively, together with the reconstructed step; it is worth noting the good reconstruction performed by the estimated model of the step; similar results have been obtained for the vertical direction.
The estimated value of additive noise variance is about 3.87, its mean value is zero and we notice that the grey level does not significantly affect the value of these estimates.

As far as the edge detection procedure is concerned, we here present a highly critical simulated benchmark, to stress efficiency and effectiveness of the proposed approach. In particular, we consider the image of figure 5.3. It is a black ellipse, centred at (128,128) and with semi axes 90 and 100, respectively, with zero rotational angle, in a white background. We blurred the image with a Gaussian kernel with variance $\sigma_d^2 = 4$ and with added Gaussian noise with zero mean and variance $\sigma_n^2 = 0.03$; we also added a regular sinusoidal part, such as in the experiments reported in Chapter 3, standing for reflection of light on the pupil's concavity.

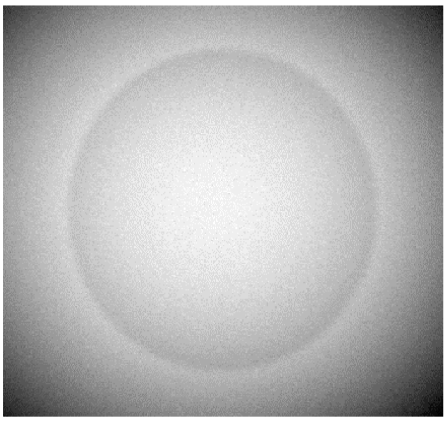


Figure 5.3 - Benchmark for testing the edge detection procedure

We chose a coordinate grid with $\Delta = 1$ and three wavelets with variances $\sigma_w^2 = 0.7;\ 2.5;\ 4$. Probability of false alarms was set to 1%. In figure 5.4(A) results of the edge detection procedure are shown. Figures 5.4(B)-5.4(F) show the best results obtained considering some well-known edge detectors, when facing the same image: Sobel, Canny, Laplacian of Gaussian, Roberts, Prewitt, respectively. For this comparison we used the Matlab Toolbox of Edge Detection. We notice that our algorithm produces more reliable results with respect to the presence of critical degradation of the images. This is a consequence of better exploiting the information we have about the model of the image.

This quality can be very useful in real image processing, where, for example, the presence of shadows and/or flat background can be regarded as additive regular signals. In figure E we present the results of the optimal fitting described in Section 5.4; the estimated values for the two semi axes are 89.81 and 100.12 respectively (pixels), the estimated rotational angle is 0.0001 rads and the coordinates of the estimated centre are (127.7, 129.4), which are estimates very close to the corresponding true values.

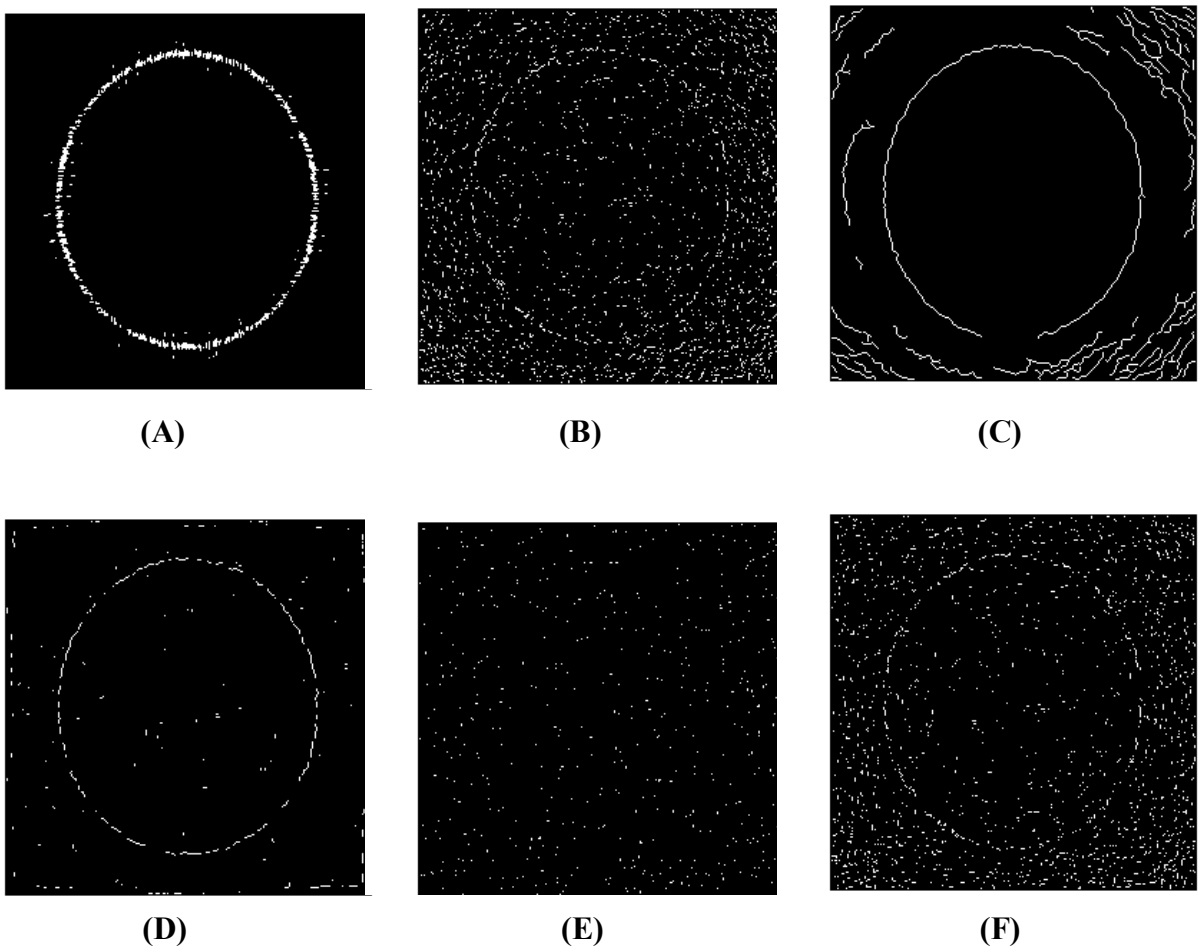


Figure 5.4 – Comparisons between the results obtained considering the proposed algorithm and some well known edge detectors when facing the target image. (A) the proposed algorithm; (B) Sobel; (C) Canny; (D) Laplacian of Gaussian; (E) Roberts; (F) Prewitt.

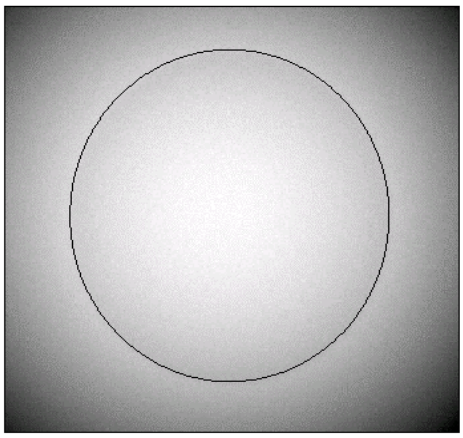
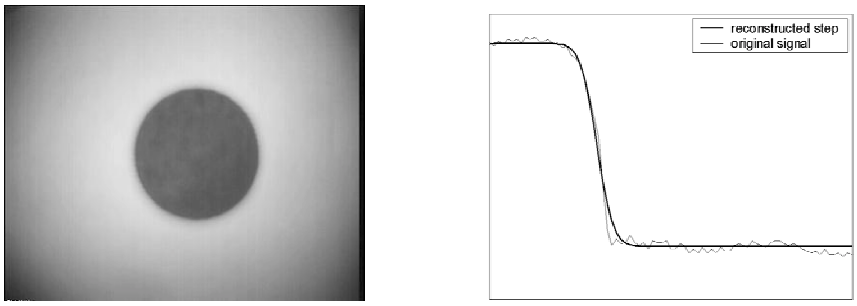


Figure 5.5 – The original image and the detected optimal ellipse.

We now consider real pupil data. The experimental set up has been already described in Section 5.2. The measured pupil image we consider is the one shown in figure A. We are going to process the image in the white square of size 256×256. As we said, in order to estimate the degradation parameters, we need a target to be shot in the same conditions in which we capture the pupil image (figure 5.6(A)). We estimate for the target image $\hat{\sigma}_d^2 = 2.90$ and $\hat{\sigma}_n^2 = 5.97$; in figure 5.6(B) a proper line extracted from the image is plotted together with the estimated blurred step.



(A) **(B)**

Figure 5.6 – The target image for the pupillometer and results of degradation parameter estimation

Using these values we apply the edge detection procedure; we consider three wavelets of variances 0.9, 1.8 and 3.6 respectively and we choose a probability value for false alarm error equal to 1%. Result is shown in figure 5.7. Comparisons with other methods can be found in [108].

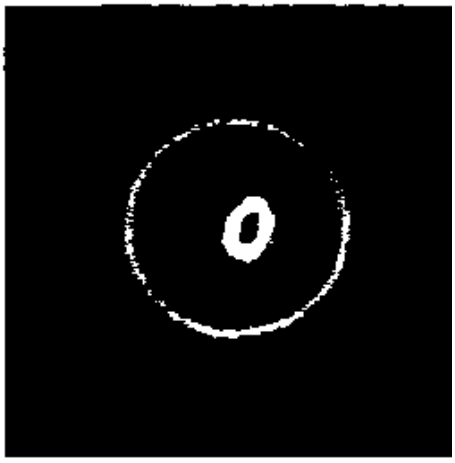


Figure 5.7 – Detected edge points

In order to obtain a correct estimation of the characteristic parameters of the pupil we have now to discard all the points detected as a consequence of the reflex of the spot of light due to the camera. To do this, we perform the clustering described in Section 5.4. We notice we have a very small number of false alarms (i.e. points not belonging to any contour at all) and the clustering procedure results to be as simple as efficient. Hence we interpolate the remaining points by optimal fitting them with an elliptic shape. In figure 5.8 result is shown. The estimated values for the semi axes of the ellipse (in pixel units) are 73.21 and 70.86.

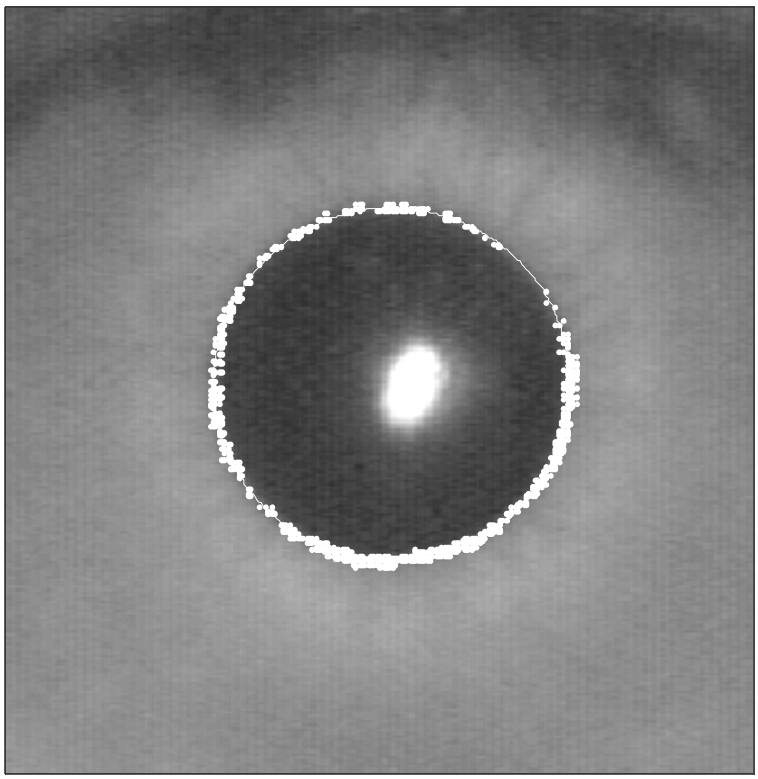


Figure 5.8 – Final result: optimal fitting of the detected edge points selected after clustering.

Many tests have been made on different pupils shot in different experimental conditions; the results we achieved always proved the effectiveness of the proposed approach for the

degradation characterization as well as the edge detection and the parametric reconstruction of the image of a human pupil.
It is worth noting that the proposed procedure, due to the clustering and discarding policy, enjoys very good fitting with short identification time. More precisely, we ran all the experiments on a personal computer Athlon XP 1700 and the complete identification procedure on average lasted 3.2 seconds. This is to say that large improvements can still be obtained by a computational point of view, if we had a more suitable computer, with a faster CPU. Such an efficiency is an important issue because the future interest could be to estimate the fluctuation of the pupil size over the acquisition of a time sequence of many frames. In [108] this processing is extended to a whole sequence of frames, each one considered independent and separately processed from the others and results are shown about the fluctuation in time of the two semi axes of the ellipse fitting the pupil.

§ 5.5 – Frame sequence processing

As said, our aim is not to apply an edge identification procedure to each frame separately, but to consider information coming from knowledge of scene motion, in order to relate the different images of the sequence with each other. Knowledge about the motion of the objects in the scene can be regarded as an additional information we ask for. Consequently, we can reasonably hope to be able to discard some other information, otherwise necessary, such as, for example, the degradation of the acquired images, given by blurring covariance matrix and additive white Gaussian noise statistical propeties.
A useful experiment would be then to suppose to have not the correct estimation for these parameters, as we already did in the last experiment presented in Section 3.
Unfortunately, in the case of the pupillometer application, a key issue for applying the complete procedure to the acquired sequence, is not satisfied. In fact, as we said in the introduction of this Chapter, human pupil continuously fluctuates (expanding or reducing) and moves in a random way, so that a model for its motion, if any, can be only stochastic. Considering such a characteristic, the only way we have to apply our methodology is to find some frames in which the pupil did not move nor fluctuate. If they exist, we can give null motion in input to our algorithm and see what happens. These considerations led us to design the following experiment.
First, we found three subsequent frames corresponding to zero motion for the pupil. Then, we did not perform any degradation parameter estimation, nor target image identification. We just gave reasonable values for the covariance matrix of the Gaussian blur kernel and for the variance of noise. In particular we set $\Sigma_b = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ and $\sigma_n^2 = 0.0004$, which are clearly wrong. Wavelets were chosen equal to $\sigma_w^2 = [0.6 \quad 1.2 \quad 2.4]$. Frames considered, together with obtained results, are reported in figure 5.9.

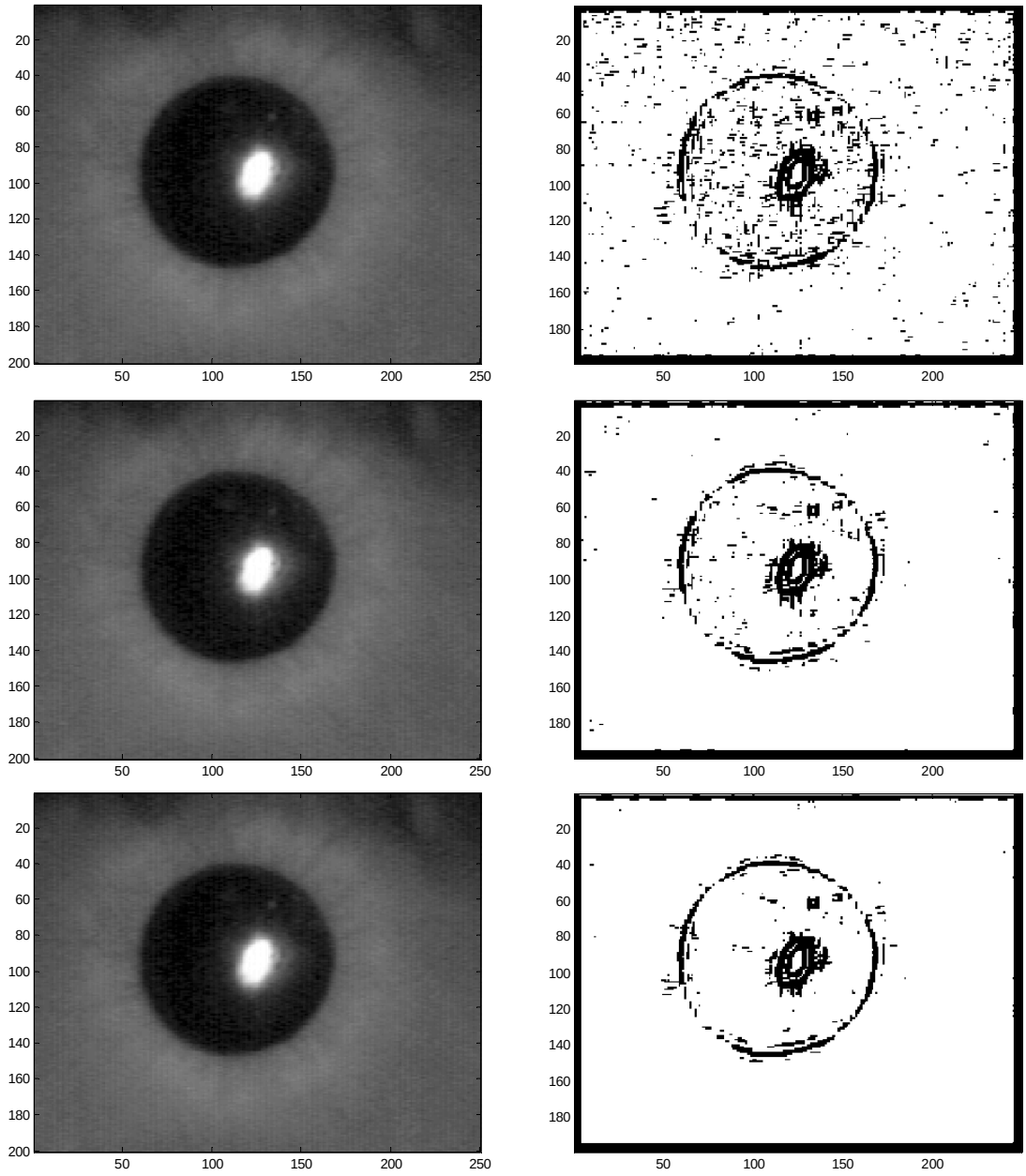


Figure 5.9 – Analysing the sequence of pupil images with wrong information about degradation.

We notice that the initial number of false alarms, due to wrong information about degradation parameters, suddenly decreases. The following steps we would have to accomplish concerns clustering of the detected edge points, in order to distinguish the ones due to pupil contour from the ones due to light reflection, and optimal fitting to identify the morphological parameters.

Even if on a difficult benchmark and with such few frames, the proposed methodology succeeds in detecting real edges of the image. Once more, as we asserted in Chapter 3, this experiment shows the potentiality of our algorithm in efficiently obtaining a suitable estimate of edges in the image sequence, allowing us to avoid any preliminary task of blur and noise identification.

Several experiments have been carried out to prove robustness of the algorithm with respect to uncertainties in degradation parameter estimation and all the results we obtained confirm the above considerations.

Concluding remarks

The problem of image sequence analysis has been faced in this Thesis.

Since its birth, this problem received much attention by the scientific community due to either theoretical aspects or applicative relevance. It's worth to notice that also the case of single image processing, as well as the case of processing a sequence of frames with null motion, fit in the general set up of image sequence analysis. In Chapter 1 we presented a review of the main achievements in these fields.
In this work, images are supposed to be acquired by a measurement device which introduces a level of degradation. Furthermore, the imaged scene undergoes a transformation, consequent to objects motion and/or deformation as well as camera motion.

Definition of admissible motion and model of the degradation introduced by the measurement process are given in Chapter 2. A general mathematical framework is therewith introduced to deal with both the problems of reconstructing edges of the images from the sequence and of identifying the motion parameters vector. From a methodological point of view, this unified framework represents a step forward with respect to the existent theoretical achievements, because it allows looking at different problems as particular instances of the same set up.

The first problem we dealt with (Chapter 3) was edge detection of objects in the sequence given exact knowledge about the underlying motion pattern. We proposed a recursive estimation procedure, made up of two steps: a preliminary pre-processing of the measured image and the proper edge identification procedure. The pre-processing was aimed to localize the detection problem and to improve the signal-to-noise ratio. The estimation problem turned out to be linear, so that it was accomplished by an algorithm which implements recursive linear quadratic estimation techniques and successive hypothesis testing. The procedure was tested against simulated and real data with outstanding results. Effectiveness of the proposed algorithm is largely due to the exploitation of the relationship introduced by the known motion between the various frames. Robustness of the estimation procedure with respect to degradation parameters estimation was also stressed as a qualifying feature of the overall methodology.

The second problem we faced (Chapter 4) was motion parameter identification and estimation when the initial uncorrupted image is known. Differently from the previous case, the estimation problem turned out to be non linear, in that the measurement equation is such. Moreover, we had no hypothesis on the probability distribution we dealt with, so that we couldn't assure the resulting filtering equations to be finite dimensional. As a matter of fact, to make the problem feasible for numerical solution, we presented two different ways to suitably approximate the general formulation: the first one was based on a Gaussian approximation, the second one was based on a linearization operation.

In the last chapter of the Thesis (Chapter 5) we presented an application of the above procedure of edge detection to a biomedical applicative problem. Images coming from a pupillometer were considered. We started with considering a single pupil image. Preliminary to edge detection we presented a suitable procedure to estimate the degradation parameters introduced by the measurement device. Preliminary to edge detection was the degradation parameters estimation and to this aim a procedure was introduced to reconstruct both blurring and noise characteristic parameters. The above procedure for edge detection was then implemented in the case of a single image, with good results, also due to a suitable fitting of the detected edge points with an elliptic model.

A remarkable improvement was obtained by processing a sequence of pupil images, acquired from the pupillometer over sufficiently short time instants, so as to justify the hypothesis of null motion. In particular, this experiment confirmed robustness of results with respect to inaccurate estimates of the parameters which describe degradation introduced by the pupillometer.

Possible future developments of this work can be planned following three main lines.

- Identification and solution of applicative problems relevant to the edge identification procedure with known motion. Particular attention will be devoted to biomedical problems.
- Characterization and implementation of suitable approximations to reduce computational complexity, with reference to the motion identification procedure with knowledge about the initial image. This will lead also in this case to possible applications to real contexts.
- Integration of the two faced problems in a unique framework, by setting up a general identification procedure capable of simultaneously estimating contours and motion in an image sequence.

Bibliography

[1] D. Ziou, S. Tabbone, *Edge Detection Techniques – An Overview*, Journal of Pattern Analysis and Artificial Intelligence, 8 (4), 1998.

[2] V. Torre and T. Poggio, *On Edge Detection*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 8 (2): 147-163, 1986.

[3] T. Poggio and V. Torre, *Ill-Posed Problems and Regularization Analysis in Early Vision.* Artificial Intelligence Lab. Memo, No. 773, MIT, 1984.

[4] T. Poggio and V. Torre, *A Regularized Solution to Edge Detection.* Artificial Intelligence Lab. Memo, No. 833, MIT, 1985.

[5] A. Graps, *An Introduction to Wavelets*, IEEE Computational Sciences and Engineering, 2 (2): 50-61, 1995.

[6] I. Pitas and A. N. Venetsanopoulos, *Edge Detector Based on Nonlinear Filters*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(4): 538-550, 1986.

[7] J. F. Canny, *A Computational Approach to Edge Detection*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 8 (6): 679-698, 1986.

[8] S. Dellepiane, D. D. Guisto, and G. Vernazza, *Automatic Parameter Computation for Edge Detection by the Zero-Crossing Method*, in 12th Conference on Signal Processing and Images, pp. 617-620, 1989.

[9] A. Rosenfeld and M. Thurston, *Edge and Curve Detection for Visual Scene Analysis*, IEEE Transactions on Computer, 20 (5): 562-569, 1971.

[10] D. Marr and E. C. Hildreth, *Theory of Edge Detection*, Proceedings of the Royal Society of London, B207, pp. 187-217, 1980.

[11] L. Wu and Z. Xie, *Scaling Theorems for Zero-Crossings*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 12 (1): 46-54, 1990.

[12] D. J. Williams and M. Shah, *Edge Contours Using Multiple Scales*, Computer Vision, Graphics and Image Processing, 51: 256-274, 1990.

[13] F. Bergholm, *Edge Focusing*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 9 (6): 726-741, 1987.

[14] P. Perona and J. Malik, *Scale-Space and Edge Detection Using Anisotropic Diffusion*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 12 (7): 629-639, 1990.

[15] Y. Lu and R. C. Jain, *Reasoning about Edges in Scale Space*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 14 (4): 450-468, 1992.

[16] W. H. H. J. Lunscher and M. P. Beddoes, *Optimal Edge Detector Design I: Parameter Selection and Noise Effects*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 8 (2): 164-177, 1986.

[17] M. Gokmen and C. C. Li, *Edge Detection and Surface Reconstruction Using Refined Regularization*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 15 (5): 492-499, 1993.

[18] J. Babaud, A. P. Witkin, M. Baudin, and R. O. Duda, *Uniqueness of the Gaussian Kernel for Scale-Space Filtering*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 8 (1): 26-33, 1986.

[19] I. E. Abdou, *Quantitative Methods of Edge Detection*, Technical Report no. 830, Image Processing Institute, University of Southern California, 1978.

[20] S. Venkatesh and L. J. Kitchen, *Edge Evaluation Using Necessary Components*, CVGIP: Graphical Models and Image Processing, 54 (1): 23-30, 1992.

[21] L. Kitchen and A. Rosenfeld, *Edge Evaluation Using Local Edge Coherence*, IEEE Transactions Systems, Man and Cybernetics, SMC-11 (9): 597-605, 1981.

[22] T. Kanungo, M. Y. Jaisimba, J. Palmer, and R. M. Haralick, *A Methodology for Quantitative Performance Evaluation of Detection Algorithms*, IEEE Transactions on Image Processing, 12 (4): 1667-1674, 1995.

[23] I. Pitas, *Digital Image Processing Algorithms and Applications*, John Wiley and Sons, 2000.

[24] J. M. S. Prewitt, *Object Enhancement and Extraction*, in Picture Processing and Psychopictorics, B. S. Lipkin et al., Eds. Academic Press, New York, 1970.

[25] M. Basu, *Gaussian-Based Edge-Detection Methods — A Survey*, IEEE Transactions on Systems, Man and Cybernetics — Part C: Applications and Reviews, 32 (3): 252-260, 2002.

[26] S. Mallat, S. Zhong, *Characterization of Signals from Multiscale Edges*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 14 (7), 1992.

[27] M. H. Hueckel, *A Local Visual Operator Which Recognizes Edges and Lines*, J. Assoc. Comput. Mach., 20 (4): 634-647, 1973.

[28] S. K. Nayar, S. Baker and H. Murase, *Parametric Feature Detection*, in Proceedings of IEEE, International Conference on Computer Vision and Pattern Recognition, pp. 471-477, San Francisco, 1996.

[29] R. M. Haralick, *Digital Step Edges from Zero Crossing of Second Directional Derivatives*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 6 (1): 58-68, 1984.

[30] H. Kalviainen, P. Hirvonen, L. Xu, E. Oja, *Probabilistic and non-Probabilistic Hough Transform: Overview and Comparisons*, Image and Vision Computing, 13 (4): 239-252, 1995.

[31] D. Ziou, *Line Detection Using an Optimal IIR Filter*, Pattern Recognition, 24 (6): 465-478, 1991.

[32] K. N. V. L. N. Koundinya and B. Chanda, *Detecting Lines in Gray Level Images Using Search Techniques*, Signal Processing, 37: 287-299, 1994.

[33] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision*. International Thompson Computer Press, London, 1993.

[34] M. Levy, *New theoretical approach to relaxation, application to edge detection*, in Proceedings of 9th International Conference on Pattern Recognition, pp. 208-212, Rome, 1988.

[35] R. M. Haralick, L. G. Shapiro, *SURVEY: Image Segmentation Techniques*, Computer Vision, Graphics and Image Processing, 29: 100-132, 1985.

[36] C. Yian-Leng, L. Xiaobo, *Fast Image Region Growing*, Image and Vision Computing, 13 (7): 559-571, 1995.

[37] M. Kass, A. Witkin, and D. Terzopoulos, *Snakes: active contour models*, International Journal of Computer Vision, 1 (4): 321-331, 1988.

[38] A. A. Amini, T. E. Weymouth, and R. C. Jain, *Using Dynamic Programming for Solving Variational Problems in Vision*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 12: 855-867, 1990.

[39] T.S.Huang, *Image sequence analysis*, Springer Series in Information Sciences, Springer-Verlag, 1981.

[40] C.Stiller, J.Conrad, *Estimating motion in image sequences*, *IEEE Signal Processing Magazine*, 70-91, July 1999

[41] B. K. P. Horn, B. G. Schunck, *Determining Optical Flow*, Artificial Intelligence, 17: 185-203, 1981.

[42] A. Mitiche, P. Bouthemy, *Computation and Analysis of Image Motion: A Synopsis of Current Problems and Methods*, International Journal of Computer Vision, 19 (1): 29-55, 1996.

[43] R. C. Jain, *Dynamic scene analysis*, Progress in Pattern Recognition 2, L. Kanal and A. Rosenfeld (Eds.), north Holland, 125-167, 1985.

[44] S. D. Blostein, T. S. Huang, *Detecting small, moving objects in image sequences using sequential hypothesis testing*, IEEE Transactions on Signal Processing, 39 (7): 1161-1629, 1991.

[45] J. Wiklund, G. H. Granlund, *Image sequence analysis for object tracking*, Proceedings of 5th Scandinavian Conference on Image Analysis, 641-648, 1987.

[46] K. P. Karmann, A. Brandt, *Moving Object Recognition using an Adaptive Background Memory*, Proceedings of 3rd International Workshop on Time-varying Image Processing and Moving Object Recognition, 289-296, 1989.

[47] J. M. Letang, V. Rebuffel, P. Bouthemy, *Motion detection robust to perturbations: A statistical regularization and temporal integration framework*, Proceedings of 4th International Conference on Computer Vision, 21-30, 1993.

[48] S. Ullman, *The Interpretation of Visual Motion*, Cambridge, MA: MIT Press, 1979.

[49] J. W. Roach, J. K. Aggarwal, *Determining the movement of objects from a sequence of images*, IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-2 (6): 554-562, 1980.

[50] B. G. Schunck, *Image flow: fundamentals and algorithms*, in *Motion understanding: Robot and Human Vision*, W. N. Martin and J. K. Aggarwal (Eds.), Norwell, MA: Kluwer Academic Publishers, 1988.

[51] J. K. Aggarwal, N. Nandhakumar, *On the Computation of Motion from Sequences of Images – A Review*, Proceedings of the IEEE, 76 (8): 917-935, August 1988.

[52] W. B. Thompson, T. G. Pong, *Detecting Moving Objects*, International Journal of Computer Vision, 4: 39-57, 1990.

[53] R. C. Nelson, *Qualitative detection of motion by a moving observer*, International Journal of Computer Vision, 7 (1): 33-46, 1991.

[54] H. H. Nagel, G. Socher, H. Kollnig, M. Otte, *Motion boundary detection in image sequences by local stochastic tests*, Proceedings of 3rd European Conference on Computer Vision, 305-315, 1994.

[55] S. Peleg, H. Rom, *Motion-based Segmentation*, Proceedings of 10th International Conference on Pattern Recognition, 109-113, 1990.

[56] K. Wohn, A. M. Waxman, *The analytic structure of image flows: Deformation and segmentation*, Computer Vision, Graphics and Image Processing, 49: 127-151, 1990.

[57] P. H. S. Torr, D. W. Murray, *Statistical detection of independent movement from a moving camera*, Image and Vision Computing Journal, 11 (4): 180-187, 1993.

[58] T. Jebara, A. Azarbayejani, A. Pentland, *3D Structure from 2D Motion*, IEEE Signal Processing Magazine, 16 (3), 1999.

[59] A. L. Yuille, N. M. Grywacz, *A mathematical analysis of the motion coherence theory*, International Journal of Computer Vision, 3: 155-175, 1989.

[60] A. J. Patti, A. M. T. M. I. Sezan, *Image sequence restoration and de-interlacing by motion compensated Kalman filtering*, Proceedings of SPIE, 1903: 59-70, 1991.

[61] S. J. Ko, Y. H. Lee, *Nonlinear spatio-temporal noise suppression techniques with applications in image sequence processing*, in IEEE International Symposium on Circuits and Systems, 5: 662-665, 1991.

[62] A. K. Katsaggelos, R. P. Kleihorst, S. N. Efstratiadis, R. L. Lagendijk, *Adaptive image sequence noise filtering methods*, Proceedings of SPIE, 1606: 716-727, 1991.

[63] M. Elad, A. Feuer, *Superresolution Restoration of an Image Sequence: Adaptive Filtering Approach*, IEEE Transactions on Image Processing, 8 (3): 387-395, 1999.

[64] W. S. Lee, D. C. Slaughter, D. K. Giles, *Development of a Machine Vision System for Weed Control Using Precision Chemical Application*, Proceedings of International Conference on Agricultural Machinery Engineering, 3: 802-811, 1996.

[65] M. A. Shahin, S. J. Symons, *A machine vision system for grading lentils*, Canadian Biosystems Engineering, 43: 7.7-7.14, 2001.

[66] S. Iwamoto, D. M. Checkley Jr., M. M. Trivedi, *REFLICS: Real-time flow imaging and classification system*, Machine Vision and Visual Applications, 13: 1-13, 2001.

[67] T. Steinherz, E. Rivlin, and N. Intrator, *Offline cursive script recognition – A survey*, International Journal on Document Analysis and Recognition, 2: 90-110, 1999.

[68] V. S. Nalwa, *Automatic on-line Signature Verification*, Proceedings of the IEEE, 85 (2): 215-239, 1997.

[69] K. Rohr, *Towards model-based recognition of human movements in image sequences,* Computer Vision Graphics and Image Processing, 59(1):94-115, 1994.

[70] Y. Cui and J. Weng, *Hand segmentation using learning-based prediction and verification for hand-sign recognition*, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 88-93, 1996.

[71] A. K. Jain, A. Ross, S. Pankanti, *A Prototype Hand Geometry-based Verification System*, Proceedings of 2nd International Conference on Audio- and Video-based Biometric Person Authentication, 166-171, 1999.

[72] K.-N. Kim, R. S. Ramakrishna, *Vision-based eye-gaze tracking for human computer interface*, Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, 2: 324-329, 1999.

[73] M. Saedan, M. H. Ang Jr., *3D Vision-Based Control ON An Industrial Robot*, Proceedings of IASTED International Conference on Robotics and Applications, 152-157, 2001.

[74] M. Vincze, A. Pichler, G. Biegelbauer, K. Hausler, H. Andersen, O. Madsen, M. Kristiansen, *Automatic Robotic Spray Painting of Low Volume High Variant Parts*, Proceedings of 33rd International Symposium on Robotics, 2002.

[75] O. G. Cula, K. J. Dana, *Image-based Skin Analysis*, Proceedings of European Conference on Computer Vision, 2002.

[76] S. Loncaric, M. Subasic, E. Sorantin, *3D image analysis of abdominal aortic aneurysm*, Proceedings of SPIE, 4684: 1681-1690, 2002.

[77] S. Vilas Boas Jardim, M. A. T. Figuereido, *Automatic Analysis of Fetal Echographic Images*, Proceedings of 12th Portuguese Conference on Pattern Recognition, 2002.

[78] A. Materka, P. Cichy, J. Tuliszkiewicz, *Texture Analysis of X-Ray Images for Detection of Changes in Bone Mass and Structure*, Proceedings of the Workshop on Texture Analysis in Machine Vision, 1999.

[79] G. M. Siouris, G. Chen, J. Wang, *Tracking an Incoming Ballistic Missile Using an Extended Interval Kalman Filter*, IEEE Transactions on Aerospace and Electronic Systems, 33 (1): 232-240, 1997.

[80] P. I. Corke, *Visual Control of Robots: high-performance visual servoing*, Research Studies Press, John Wiley & Sons inc., 1996.

[81] D. P. Mukherjee, S. T. Acton, *Cloud Tracking by Scale Space Classification*, IEEE Transactions on Geoscience, and Remote Sensing, 40 (2): 405-415, 2002.

[82] V. Nair, J. J. Clark, *Automated Visual Surveillance Using Hidden Markov Models*, Proceedings of International Conference on Vision Interface, 88-93, 2002.

[83] P. Smith, M. Shah, N. da Vitoria Lobo, *Monitoring Head/Eye Motion for Driver Alertness with One Camera*, Proceedings of 15th IEEE International Conference on Pattern Recognition, 4: 636-642, 2000.

[84] C. Bruni, A. De Santis, D. Iacoviello, G. Koch, *Modeling for Edge Detection Problems in Blurred Noisy Images*, IEEE Transactions on Image Processing, 10 (10): 1447-1453, 2001

[85] C. Bruni, A. De Santis, G. Koch, *Convergence and Compactness Properties for a Family of 2-D Discontinuous Signals*, Acta Applicandae Mathematicae, 62: 131-153, 2000.

[86] F. R. Gantmacher, *Matrix Theory*, Chelsea Pub. Co, New York, 1960.

[87] A. Sommerfeld, *Mechanics of Deformable Bodies*, Lectures on Theoretical Physics, Vol. II, Academic Press, 1950.

[88] J.K.Aggarwal, Q.Cai, W.Liao, B.Sabata, *Articulated and Elastic Non-rigid Motion: A Review*, Proceedings of the Workshop on Motion of Non Rigid and Articulated Objects, 2-14, 1994.

[89] C. Bruni, R. Bruni, A. De Santis, D. Iacoviello, G. Koch, *Global Optimal Image Reconstruction from Blurred Noisy Data by a Bayesian Approach*, Journal of Optimization Theory and Applications, 115 (1): 67-96, 2002.

[90] C. Bruni, G. Koch, *Identifiability of Continuous Mixtures of Unknown Gaussian Distributions*, Annals of Probability, 13 (4): 1341-1357, 1985.

[91] L. W. Stark, *The pupil as a paradigm for neurological control systems*, IEEE Transactions on Biomedical Engineering, 31, 919-924, 1984.

[92] Hunter, Acta Neuropathologica, 68, 53-58, 1985.

[93] J.Richman, R.S.Noriega, *The sensitivity and specificity of infrared pupillometry measurements in identifying drug impairment in county probation program* , Optometry and Vision Science, 79, 2002.

[94] A.M.Oroujeh, W.O'Neill, A.P.Keegan, S.L.Merritt, *Using recursive parameter estimation for sleep disorder discrimination*, Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 3: 1928-1931, 1995.

[95] A.Uetake, A. Murata, M.Otsuka, Y.Takasawa, *Evaluation of visual fatigue during VDT tasks*, IEEE International Conference on Systems, Man and Cybernetics, 2: 1277-1282, 2000.

[96] E.H.Hess, J.M. Polt, *Pupil size in relation to mental activity during simple problem solving*, Science, 140, 1190-1192, 1964.

[97] D. Zhu, S.T. Moore, T. Raphan, *Robust pupil center detection using a curvature algorithm*, Computer Methods and Programs in Biomedicine, 59: 145-157, 1999.

[98] H. J. Wyatt, *The form of human pupil*, Vision Research, 35: 2021-2036, 1995.

[99] A. Kristek, *The physiological pupillogram of the human eye*, Czechoslovakian Journal of Ophthalmology, 21: 28-35, 1965.

[100] I.E.Lowenfeld, *Pupillary changes related to age*, in H.S. Thompson, R. Daroff, L.Frisen (Eds.), Topics in neurophthalmology, 124-150, 1979.

[101] J.R. Price, T.F.Gee, K.W.Jr.Tobin, *Blur estimation in limited-control environments*, Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 3: 1679-1672, 2001.

[102] H.-C.Lee, *Review of image-blur models in photographic system using the principles of optics*, Optical Engineering, 29: 405-421, 1990.

[103] S. Chardon, B. Vozel, K. Chehdi, *A comparative study between parametric blur estimation methods*, Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 6: 3233-3236, 1999.

[104] J.Biemond, R.L.Lagendij, M. Mersereau, *Iterative methods for image deblurring*, Proceedings of the IEEE, 78: 856-883, 1990.

[105] A.Bosco, M.Mancuso, *Adaptive filtering for image denoising*, International Conference on Consumer Electronics, 208-209, 2001.

[106] M.Ibrahim Sezan, G. Pavlovic, A.Murat Tekalp, A.Tanju Erdem, *On modelling the focus blur in image restoration*, Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing 4: 2485-2488, 1991.

[107] M.Subbarao, Tse-Chung Wei, G.Surya, *Focused image recovery from two defocused images recorded with different camera settings*, IEEE Transactions on Image Processing 4: 1613-1628, 1995.

[108] D.Iacoviello, M.Lucchetti, G.Calcagnini, F.Censi, *Pupil edge detection and morphological identification from blurred noisy images*, Proceedings of 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2003.

[109] D. Iacoviello, M. Lucchetti, *Parametric characterization of the form of the human pupil from blurred noisy images*, Computer Methods and Programs in Medicine, to appear, 2005.