UNIVERSITÀ DEGLI STUDI DI PADOVA

UNIVERSITA' DEGLI STUDI DI PADOVA

Dipartimento di Ingegneria Industriale DII

Dipartimento di Dipartimento di Tecnica e Gestione dei Sistemi Industriali (DTG)

Corso di Laurea Magistrale in Ingegneria Meccanica

Caratterizzazione Termodinamica della Propagazione di una Cricca Tramite l'Uso di una Termo-Camera

(Thermodynamic Characterization of Crack Growth using Infrared Camera)

Berto Filippo

Vinogradov Alexei

Romanin Luca, 1128577

Anno Accademico 2016/2017

Desidero ringraziare il prof. Berto F. per i consigli, critiche e il tempo dedicatomi nel corso dello svolgimento di questo lavoro. Nondimeno l'accoglienza dimostrata durante la permanenza all'NTNU

Un sentito ringraziamento, inoltre, al prof. Vinogradov A. per l'aiuto nel risolvere i numerosi inconvenienti nati durante questo percorso e per la fiducia riposta

Ai miei Genitori che mi hanno sempre sostenuto per raggiungere questo traguardo

1 TABLE OF CONTENTS

1 Table of Contents

1 Table of Contents

1 Table of Contents

2 INTRODUCTION

The assessment of the mechanical performance of engineering materials subjected to high cycle fatigue requires very long and costly testing procedures. Therefore, there is a strong drive from industry to find alternative ways of evaluating high cycle fatigue performances.

Under deformation the structural evolution is observed at all scale levels, and its development leads to irreversible deformation and failure. These processes are accompanied by accumulation and dissipation of energy. Most fatigue crack propagation studies have been conducted from either a continuum or micro-mechanics point of view. An energy approach could represent a solution.

Recent developments in IR-camera permits to take the problem from an energetic point of view.

The scope of the work is to perform a crack growth test according to ASTM E647 on a Compact Specimen and correlate energy production with the growth rate. The heat is composed by a thermo-elastic coupling term and by mechanical dissipation term. Energy dissipation in metallic materials subjected to mechanical deformation is a result of the movement of dislocations, deriving from plasticity at the crack tip, and generates a monotonously increasing variation in temperature during the test.

Experimental data are contaminated by some noise, either because of the data acquisition process and because of naturally occurring phenomena (such as material inhomogeneity). A first pre-processing step in analysing such datasets is denoising, that is, estimating the unknown signal of interest from the available noisy data.

Two ways for denoising have been performed. The first involved the use of classical spatial filtering and then filtering in the frequency domain. The second is a novel approach which employees wavelets.

A MATLAB framework has been developed to automatize the process of data processing that is quite long. In this way new functionalities and changes of parameters can easily be made.

Numerical transient simulations have also been performed to try to replicate the temperature variation. Both the case with only the thermo-elastic effect (linear material model) and with plastic heat dissipation (multilinear kinematic model) have been considered. Any discrepancies with real data might tell us what model is more accurate.

3 ENERGY APPROACHES

Most fatigue crack propagation studies have been conducted from either a continuum or micro-mechanics point of view. An energy approach could represent a solution to this problem. In fact, Griffith hypothesized in 1920 that:

> *The "theorem of minimum potential energy" may be extended so as to be capable of predicting the breaking loads of elastic solids, if account is taken of the increase of surface energy which occurs during the formation of cracks.*

And in 1950 Freudenthal proclaimed that:

> *The study of mechanical behaviour and properties at different levels of aggregation is possible only in terms of a concept which on all levels has the same meaning. This concept is energy.*

Numerous energy criteria for fatigue life estimation based on hysteresis loss have been proposed [1]–[4]. But assuming that plastic-strain hysteresis is a measure of fatigue damage doesn't provide a good fit for experimental data [1] or it doesn't account for different stress ratio R [5].

In [6] it has been proposed that the dissipated energy per saturated cycle could be considered as a material constant expressing the necessary energy to propagate a crack by the distance equivalent to the size of the process zone, and proposed a way in which the lifetime of the specimen was related to a crack propagation law driven by dissipated energy. The limit of this theory is that only a fraction of the dissipated energy contributes to damage and it's difficult to establish a precise law for the fraction itself.

in [7] a more general energy balance is proposed: at the Global level of description the rate of internal energy change in the near crack tip region may be written as:

$$\dot{U} = \int T_i \, \dot{u}_i dS + \dot{Q}$$

where dS is an incremental portion of the crack tip bounding surface, T_i is the stress and $\dot{u}_i$ the rate of displacement. It simply requires that the rate of change in the internal energy of the volume, V, bounded by S be equal to the rate of work done by the surface tractions acting on S plus the rate at which heat is added to V. Body forces and inertial effects are assumed negligible in this analysis.

The internal energy expression at the local level of description is based on the atomic material model illustrated in the figure 2-2. This lattice model is a simplification of the polycrystalline, polyphase materials usually encountered in practice (change of phase is neglected).

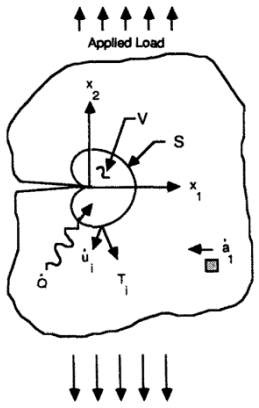


Figure 3-1 Crack tip viewed at global level of description

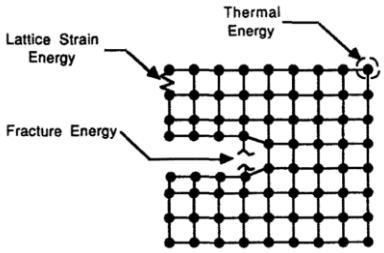


Figure 3-2 Schematic illustration of simple material model

The internal energy in a small mass element of material can be partitioned as follows:

> **Internal energy** = Lattice strain en. + Fracture en. + Thermal energy due to "internal friction" within mass element + Thermal energy due to "thermal communication" with surroundings

This fracture energy may be regarded as representing the decrease in lattice strain energy associated with the formation of lattice imperfections but in reality other phenomena contribute to fracture energy.

In [7] an analytical expression for every type of energy is substituted after some assumption and a fatigue crack propagation rate is calculated but no experimental validation has been performed. The beauty of this approach is that contributions to the internal energy of a material system due to chemical reactions and thermal (creep) effects may theoretically be included in the energy balance.

Another implementation of the energy approach is the one made by Klingbeil [5]. During fatigue crack extension the change in total potential energy per cycle must balance the plastic work per cycle W_p.

$$\frac{d\Pi}{dN} = \frac{dW_p}{N} = \frac{dW_p}{da}\frac{da}{dN}$$

Under the hypothesis that $\frac{dW_p}{da}$ corresponds to the Critical Energy Release Rate G_c (monotonic fracture toughness) of the material. The equation can be rewritten as:

$$\frac{da}{dN} = \frac{1}{G_c}\frac{dW_p}{dN}$$

The critical energy release rate G_c is related to the fracture toughness K_c in plane strain condition as:

$$G_c = \frac{K_C^2(1-\nu^2)}{E}$$

The values for $\frac{dW_p}{dN}$ were, by the author, calculated numerically and found to be proportional to ΔK^4. The crack growth can so be modelled from the fracture toughness and the values of $\frac{dW_p}{dN}$ calculated numerically. The simplicity of the model doesn't permit to account for R-ratio effects but the mean crack growth behaviour is well predicted.

An energy approach to fatigue life seems to be possible but a local approach seems more suitable. That's the case of the work of Fedorova, Bannikov and Plekhov [3].

Using thermal camera images of a growing crack and calculated the plastic work at the fatigue crack tip, from the Hutchinson, Rice and Rosengren solution, a local parameter of stored energy is defined.

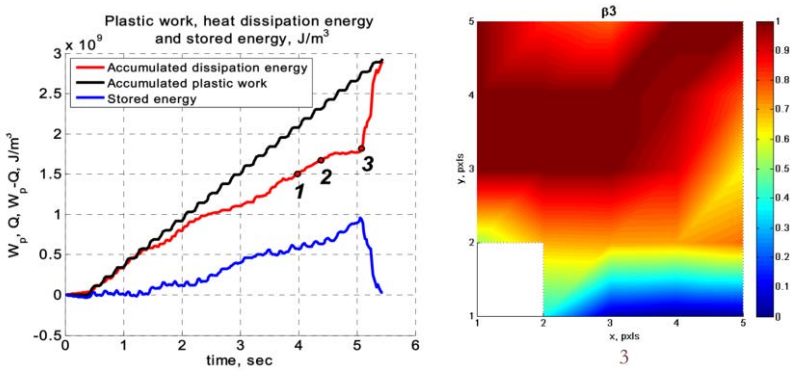


Figure 3-1 Experimental curves of evolution of heat source value, plastic work and stored energy. The accumulated plastic energy (black line) after point 3 should be corrected because the HRR solution is unable to describe the behaviour. On the right the distribution of the storage energy coefficient near the crack at the point 3 of the graph. The white pixel is the crack tip. It can be seen that on the crack surface the stored energy is zero.

At the last moments before fracture, the heat dissipation energy increases and reaches the value of plastic work. At this time the rate of plastic work is less than the rate of heat dissipation energy, this means that the stored energy goes then to zero. It has been assumed by the authors that the damage accumulation mechanism is changed, and the material approaches the fracture stage, where the role of macroscopic displacements is essential, and the energy dissipation increases significantly.

They propose the value of stored energy in the material as a failure criterion based on thermodynamic laws.

Vshivkov, Iziumova and Plekhov [8] using a heat flux sensor based on the Seeback effect to measure the heat dissipation on the crack. They discovered a linear relationship between the rate of fatigue crack propagation and the product of the rate of energy dissipation by the length of the crack. The developed method can be used to calculate the value of the dissipated energy at any moment of the material lifetime. No storage energy calculation has been performed however.

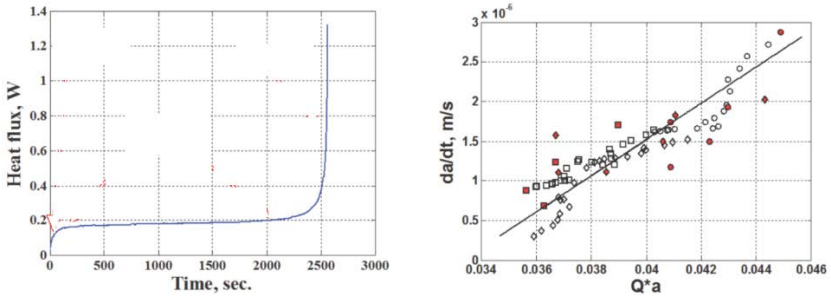


Figure 3-3 On the left, the characteristic dependence of the power of heat flux from the crack tip. On the right, the relation between fatique crack rate and the fatique crack length multiplied by power of heat flow

3.1 Self Heating

Temperature in metallic materials defines the degree of movement of the atoms constituting the lattice (Atkins, 1994). The deformation of this lattice structure can involve a variation of the atomic energy:

- The increase or the decrease of the atomic distance due to elastic straining (respectively tension or compression) causes variation of the kinetic energy of the atoms. This leads to the thermo-elastic effect where, at a macroscopic scale, the specimen temperature decreases when tested in tension and increases when tested in compression;
- Internal friction (Caillard and Martin, 2003) is a term that gathers a number of dissipative phenomena attributed to the movements of atoms from their equilibrium position. This is caused either by some local defect in the lattice (diffusion of interstitial atoms in the Snoek effect, substitution in the Zener effect) or by the movements of dislocations. The fast energy release occurring in the lattice when the atoms go back to their equilibrium positions generates an acoustic wave that increases the kinetic energy of the atoms, therefore increasing the temperature.

Nevertheless, there are several obstacles to the movements of dislocations that require a sufficient stress level, such as the resistance of the atomic lattice, the dislocation network, the grain boundaries, the polycrystal and/or external obstacles. When the stress level increases, more dislocations are created, eventually leading to failure.

The heat source is a priori composed of three parts:

- **Thermo-elastic** coupling term, reversible in terms of thermodynamics, which is reflected as first approximation as a harmonic variation of temperature in phase with the loading;
- **Latent heat** source due to phase transformation (if present, like exothermic A to M transformation in shape memory alloys)
- **Mechanical dissipation** (associated to an irreversibility in the phase transformation or in the martensite reorientation process, and to plasticity). Energy dissipation in metallic materials subjected to mechanical deformation is a result of the movement of dislocations, among other less important mechanisms, generate a monotonously increasing variation in temperature during a fatigue test.

All these three terms are included into the heat source field $Q(x,y,z,t)$ inside the three-dimensional Heat Equation.

$$Q(x,y,z,t) + k\left(\frac{\partial^2 T(x,y,z,t)}{\partial x^2} + \frac{\partial^2 T(x,y,z,t)}{\partial y^2} + \frac{\partial^2 T(x,y,z,t)}{\partial z^2}\right) = \rho C \frac{\partial T(x,y,z,t)}{\partial t}$$

The heat equation has to be averaged on the thickness dimension of the specimen because only the surface temperature is available from experimental data. The temperature at the surface is so assumed to be representative of the entire thickness.

$$Q(x,y,t) = \rho C \frac{\partial T(x,y,t)}{\partial t} - k\left(\frac{\partial^2 T(x,y,t)}{\partial x^2} + \frac{\partial^2 T(x,y,t)}{\partial y^2}\right)$$

Heat transfer effect needs to be considered, but neglecting the heat transfer with air is reasonable because of the large heat capacity of steels.

3.2 Analytical Models for Heat Generation

Various authors have proposed an analytical model for the heat generation at the crack tip. One of the first have been Rice and Levy in 1969 [9]. Using a non hardening plasticity model, neglecting all thermomechanical coupling in affecting the near tip temperature field and assuming the plastic work rate from these models as the heat-generation rate for the temperature calculation.

The crack is assumed to be embedded in a specimen finite in thickness and with infinite extension. The problem is then reduced to a 2 dimensional one.

Proceeding from the work of Carslaw and Jaeger (1959) and computing the Dugdale plastic work rate in plane strain for a stationary crack (the Dugdale model represents yielding by essentially viewing the crack as longer by a length equal to the plastic zone size), the crack temperature rise is given by Rice and Levy for a stationary crack.

After loading from times 0 to t:

$$\theta(t) = \int_0^t \left(\iint_{A_p(\tau)} \frac{q(r,\theta,\tau)}{\rho C} e^{-\frac{r^2}{4a^2(t-\tau)}} r \, d\theta \, dr \right) \frac{1}{4\pi a^2 (t-\tau)} d\tau$$

Where q is the plastic work rate.

A simple law of the dissipated power per unit length of crack front is supposed by Ranc, Palin-Luc and Paris [10] to be proportional to the surface area of the reverse cyclic plastic zone and the loading frequency f

$$q = f\varepsilon = f\beta r_R^2$$

Where in the plane stress the radius of the plastic zone is $r_R = \frac{\Delta K_I^2}{8\pi\sigma_0^2}$ with ε the dissipated energy per unit length of crack front during one cycle, r_R the radius of the reverse cyclic plastic zone, β a material dependent proportionality factor and σ_0 the cyclic yield stress of the material.

The dissipated power per unit length of crack front is therefore proportional to the variation of the stress intensity factor to the power four for a slow-moving crack.

$$q = q_0 \Delta K^4$$

A constant heat source can be considered only if the Pèclet number (which compares the characteristic time of thermal diffusion with the characteristic time associated to the heat source velocity).

$$Pe = \frac{L\,v}{a}$$

where 'L' is the characteristic length of crack propagation, 'v' the crack velocity and 'a' the thermal diffusivity. For a crack length of around 1 mm, a crack velocity of 0.1 mm/s and a thermal diffusivity typical of steel, Pèclet number remains small compared to unity and therefore the heat source can also be considered as motionless.

Hypothesizing a point heat source and at time zero an homogeneous temperature T_0 of the plate. Between time t = 0 and time t, the temperature variation field $\theta(r,t) = T(r,t)-T_0$ can be expressed by:

$$\theta(r,t) = \frac{1}{4\pi\lambda}\int_0^t q\, e^{-\frac{r^2}{4a(t-t')}}\,\frac{1}{t-t'}\,dt'$$

This simple law predicts that temperature increases abruptly when the radius tends to zero. The inconvenience is that q has to be derived experimentally.

Loos and Brotzen [11] extended this model and they predicted the crack tip temperature as a function of stress-intensity amplitude, cyclic loading frequency, time under load and material parameters. Moreover the plastic work rate is defined inside the plastic zone and zero elsewhere, the heat source isn't anymore a point but a distributed one. As before the deformation is assumed to be perfectly elastic plastic under small scale yielding and following the Irwin equations. More important the particular shape of the plastic zone is of little consequence.

A load with a sinusoidal law is hypothesized and phenomena like elastic relaxation of crack-tip material immediately after load reversals and crack closure due to plasticity are neglected.

$$\theta(x,y,t_f) = \theta(0,0,t_f) = \frac{\pi m K_I^2\sqrt{t_f}}{64\,G\,t_1\sqrt{\lambda\rho C}}\int_1^\infty \beta_p e^{-\frac{\alpha\beta_p^2}{2}}\left(I_0\left(\frac{\alpha\beta_p^2}{2}\right)+I_1\left(\frac{\alpha\beta_p^2}{2}\right)\right) sin\left(\pi\, frac\left(\frac{t_f}{t_1}\left(1-\frac{1}{\alpha}\right)\right)\right)\left[1 - cos\left(\pi\, frac\left(\frac{t_f}{t_1}\left(1-\frac{1}{\alpha}\right)\right)\right)\right]^3\frac{d\alpha}{\alpha}$$

Where $\beta_p = \frac{R_p}{2a\sqrt{t_f}}$, $R_p$is the plastic radius, $m = \frac{3(1-\nu)}{4\sqrt{2}(2+\pi)}$, $\alpha = \frac{t_f}{t_f-t}$ is the variable of integration. t_1 is the time for half load cycle and t_f the time considered at the end.

The stationary crack model assumes that crack propagation is negligible during the time t_f, so that little generated heat is left behind by the advancing crack. If this condition is not satisfied, the actual crack-tip temperature would be overestimated by the stationary crack model.

The previous equation present numerical problems during the numeric integration, so usually in literature the plastic zone is modelled as a heat point source, time invariant. The temperature raise has a singularity that isn't physically possible. The difference is shown in the next figure.

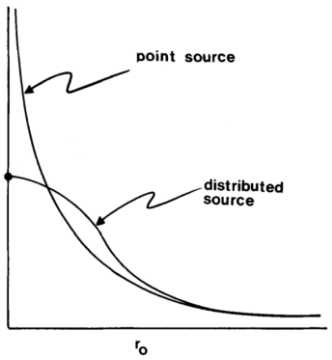


Figure 3-4 Temperature versus distance from crack tip. The point source equation has a singularity

In reality, the specimen is finite and heat eventually flows into the environment. Equilibrium is reached when heat leaving the free surface is equal to heat generated at the crack tip. But, long before equilibrium is reached, the specimen behaves essentially as if it were infinite. Moreover, the high frequency of cycles permits to neglect convection and radiation heat fluxes.

3.2.1 Thermo-elastic effect: analytical derivation

Since the continuum is considered to deform elastically, it will be possible to employ the concepts of reversibility to describe its behaviour. It is assumed that a reversible transformation always exists between two fixed states.

Therefore, state functions will be used to characterize the system during a process, regardless of whether thermodynamic equilibrium exists or not at that instant in reality, assuming that a locally reversible process is going on.

From the thermodynamic point of view, in order to investigate the mechanical and thermal behaviour, the state of an elastic solid can be described by assigning the strain tensor and temperature as state variables, assuming that only small displacements are present.

Internal energy u, may be written as:

$$u = u(\varepsilon_{ij}, T)$$

Assuming the important characteristic of state functions that infinitesimal increments are exact differentials:

$$du = \left(\frac{du}{d\varepsilon_{ij}}\right)_T d\varepsilon_{ij} + \left(\frac{du}{dT}\right)_{\varepsilon_{ij}} dT$$

From the first thermodynamical law with the terms written per unit mass:

$$du = \delta w + \delta q$$

u is the internal energy and w and q are the work and heat respectively exchanged between the system and the external surroundings (d is for exact differentials and δ for infinitesimal differences). q is

positive when transferred from the external surroundings to the system and the work positive when done on the system by forces external to it.

Entropy s can be defined with reference to a reversibile process as

$$(ds)_{reversible} = \frac{(\delta q)_{reversible}}{T}$$

But when considering a more general irreversible process, there is always a dissipation of energy and, even in an adiabatic process (δ $q_{reversible}$=0), there is always an increase in the entropy of the system.

$$ds = \frac{\delta q}{T} \geq \frac{(\delta q)_{reversible}}{T}$$

If a continuum in equilibrium is considered which undergoes deformations in a quasi-static way (such that there are no inertial forces acting on it and the kinetic energy is constant), and the continuum is assumed to have small deformations compared with the body dimensions, such that any higher-order terms for work are neglected, then, from the principle of virtual work, the work done by external forces on the system is equal to the strain energy gained by the continuum. Therefore, considering the system represented by an infinitesimal element of unit volume, the work exchanged with the external surroundings is given by the strain energy density:

$$\delta w_v = \sigma_{ij} \cdot d\varepsilon_{ij} \text{ (per unit of volume)}$$

where ε_{ij} is the small elastic strain tensor, σ_{ij} is the stress tensor and the temperature is considered as constant. For application to a unit mass, it is necessary to multiply by the specific volume v=1/ρ where ρ is the density.

$$\delta w \quad = \delta w_v \cdot v = \frac{\sigma_{ij} \cdot d\varepsilon_{ij}}{\rho}$$

Merging this equation with the first and second thermodynamic laws we obtain:

$$du = \delta w + \delta q = \frac{\sigma_{ij} \cdot d\varepsilon_{ij}}{\rho} + T \cdot ds$$

The Helmholtz thermodynamic potential, per unit mass, has to be introduced:

$$H = u - Ts$$

That can be written in its differential form

$$dH = du - T \cdot ds - s \cdot dT$$

Substituting du and for local reversible changes (ds=0):

$$dH = \frac{\sigma_{ij} \cdot d\varepsilon_{ij}}{\rho} - s \cdot dT$$

The Helmholtz thermodynamic potential is a function of strain and temperature and is an exact differential:

$$dH = \left(\frac{dH}{d\varepsilon_{ij}}\right)_T d\varepsilon_{ij} + \left(\frac{dH}{dT}\right)_{\varepsilon_{ij}} dT$$

Comparing the last two equations we can derive that:

$$\sigma_{ij} = \rho \frac{dH}{d\varepsilon_{ij}}$$

$$s = -\frac{dH}{dT}$$

At the same time the entropy s per unit mass is a state function, and the following equation can be written:

$$ds = \left(\frac{ds}{d\varepsilon_{ij}}\right)_T d\varepsilon_{ij} + \left(\frac{ds}{dT}\right)_{\varepsilon_{ij}} dT$$

The specific heat per unit mass at zero strain (for constant-volume transformations) is defined as:

$$c_v = \left(\frac{\delta q}{dT}\right)_{\varepsilon_{ij}} = \left(\frac{T \cdot ds}{dT}\right)_{\varepsilon_{ij}} = T\left(\frac{\partial s}{\partial T}\right)_{\varepsilon_{ij}} = -T\left(\frac{\partial^2 H}{\partial T^2}\right)_{\varepsilon_{ij}}$$

Merging all equations, we obtain:

$$ds = -\frac{\partial^2 H}{\partial \varepsilon_{ij}\, \partial T} d\varepsilon_{ij} - \left(\frac{\partial^2 H}{\partial T^2}\right)_{\varepsilon_{ij}} dT$$

And then:

$$ds = -\frac{1}{\rho}\frac{\partial \sigma_{ij}}{\partial T} d\varepsilon_{ij} - c_v \frac{dT}{T}$$

Knowing that $ds = \frac{\delta q}{T}$ we obtain:

$$\frac{\delta q}{T} = -\frac{1}{\rho}\frac{\partial \sigma_{ij}}{\partial T} d\varepsilon_{ij} - c_v \frac{dT}{T}$$

From which:

$$dT = \frac{T}{\rho}\frac{\partial \sigma_{ij}}{\partial T} d\varepsilon_{ij} + \frac{\delta q}{c_v}$$

The term representing the partial derivative of the stress tensor with respect to temperature. This term can be developed using at this stage the stress-strain-temperature relation (eq. 63 in [12]) which is valid for homogeneous isotropic materials, and assuming that the Lamè elastic parameters are independent of temperature. Thus:

$$\frac{\partial \sigma_{ij}}{\partial T} = -\gamma \cdot \delta_{ij}$$

Where δ_{ij} is 1 if i=j, 0 otherwise. γ is a function of Young's modulus, Poisson's ratio and the coefficient of linear expansion as defined in eq. 68 in [12]. The product $\delta_{ij} \cdot d\varepsilon_{ij}$ is the first strain invariant or cubic dilatation $d\varepsilon_{kk}$,therefore:

$$ds = \frac{\gamma}{\rho}\delta_{ij}d\varepsilon_{ij} + c_v\frac{dT}{T} = \frac{\gamma}{\rho}d\varepsilon_{kk} + c_v\frac{dT}{T}$$

Integrating and setting $s_v = 0$ at the starting conditions, when $\varepsilon_{ij} = \left(\varepsilon_{ij}\right)_0$ and $T = T_0$ we obtain:

$$s = \frac{\gamma}{\rho}\Delta\varepsilon_{kk} + c_v \log\frac{T}{T_0} = \frac{\gamma}{\rho}\Delta\varepsilon_{kk} + c_v \log\left(1 + \frac{\Delta T}{T_0}\right)$$

Considering $\Delta T \ll T_0$ and expanding the logarithm term into an infinite power series in which only the first term is retained:

$$s = \frac{\gamma}{\rho}\Delta\varepsilon_{kk} + c_v\frac{\Delta T}{T_0}$$

For and adiabatic process, s=0, so:

$$\frac{\gamma}{\rho}\Delta\varepsilon_{kk} = -c_v\frac{\Delta T}{T_0}; \qquad \Delta\varepsilon_{kk} = -c_v\rho\frac{\Delta T}{\gamma T_0}$$

Using again the stress-strain-temperature relationship and the relationship between the linear expansion coefficients and the Lamè constants (eq. 67 in [13]):

$$\Delta\varepsilon_{kk} = \frac{1}{2\mu + 3\lambda}(\Delta\sigma_{kk} + 3\gamma\Delta T) = \frac{\alpha}{\gamma}(\Delta\sigma_{kk} + 3\gamma\Delta T)$$

Merging the last two equations:

$$\Delta\sigma_{kk} = -\Delta T\left(\frac{c_v\rho}{\gamma T_0} + 3\gamma\right)$$

Substituting the relationship between the elastic and the Lamè constants found (eq. 68 in [12]), is found that:

$$\Delta\sigma_{kk} = -\Delta T\frac{\rho}{\alpha T_0}\left[c_v + \frac{3E\alpha^2 T_0}{\rho(1-2\nu)}\right]$$

The term is square brackets in the last equation is the specific heat capacity at constant pressure c_p. We can finally obtain:

$$\Delta T = -T_0\frac{\alpha}{c_p\rho}\Delta\sigma_{kk}$$

A "revised higher-order theory of thermo-elastic effect" also exists in literature to account when the mean stress has influence on the thermo-elastic signal response [14].

4 IR-LIGHT AND THERMOGRAPHY

At temperatures higher than 0 K, the absolute zero temperature, each body emits heat as thermal radiation. The intensity of this radiation depends on wavelength and the body's temperature.

4.1 IR LIGHT

Infrared (IR) light is electromagnetic radiation with a wavelength longer than that of visible light, measured from the nominal edge of visible red light at 0.7 micrometres, and extending conventionally to 300 micrometres. These wavelengths correspond to a frequency range of approximately 430 to 1 THz, and include most of the thermal radiation emitted by objects near room temperature.

The infrared spectrum is then divided in a sub scheme based on the response of various detectors (other types of subdivision exist): near, short, mid, long and very long wave infrared. Near infrared is the region closest to the visible spectrum.

Infrared cameras operate in the Short Wave (3-5 µm) or Long Wave (8-14 µm) infrared spectrum where the atmosphere has the maximum possible transmission.

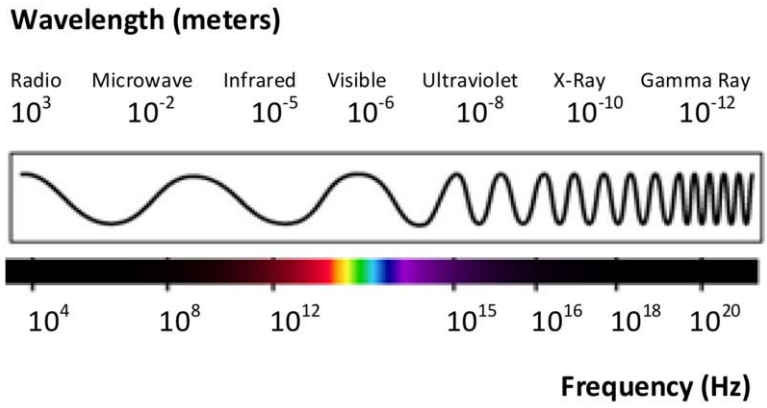


Figure 4-1 Classification of the electromagnetic radiations according to their wavelength and frequency

4.2 THE BLACK BODY AND THERMAL EMISSIVITY

A black body in thermal equilibrium (that is, at a constant temperature) emits electromagnetic radiation called black-body radiation. The radiation is emitted according to Planck's law, meaning that it has a spectrum that is determined by the temperature alone, not by the body's shape or composition.

A black body in thermal equilibrium has two notable properties:

- It is an ideal emitter: at every frequency, it emits as much energy as any other black body at the same temperature.
- It is a diffuse emitter: the energy is radiated isotropically, independent of direction.

Real materials emit energy at a fraction—called the emissivity—of black-body energy levels. By definition, a black body in thermal equilibrium has an emissivity of ε = 1.0. A source with lower emissivity independent of frequency often is referred to as a grey body.

The Spectral radiant Emittance, for a black body, is given by Plank's law:

$$q(\lambda, T) = \frac{2\pi hc^2}{\lambda^5 \left(e^{\frac{hc}{\lambda K_b T}} - 1 \right)}, \left[\frac{W}{m^2 \mu m} \right]$$

Where c is the speed of light, h is the Planck constant and K_b the Boltzman constant. Integrating over all the spectrum

$$q(T) = \sigma_0 T^4, \left[\frac{W}{m^2} \right]$$

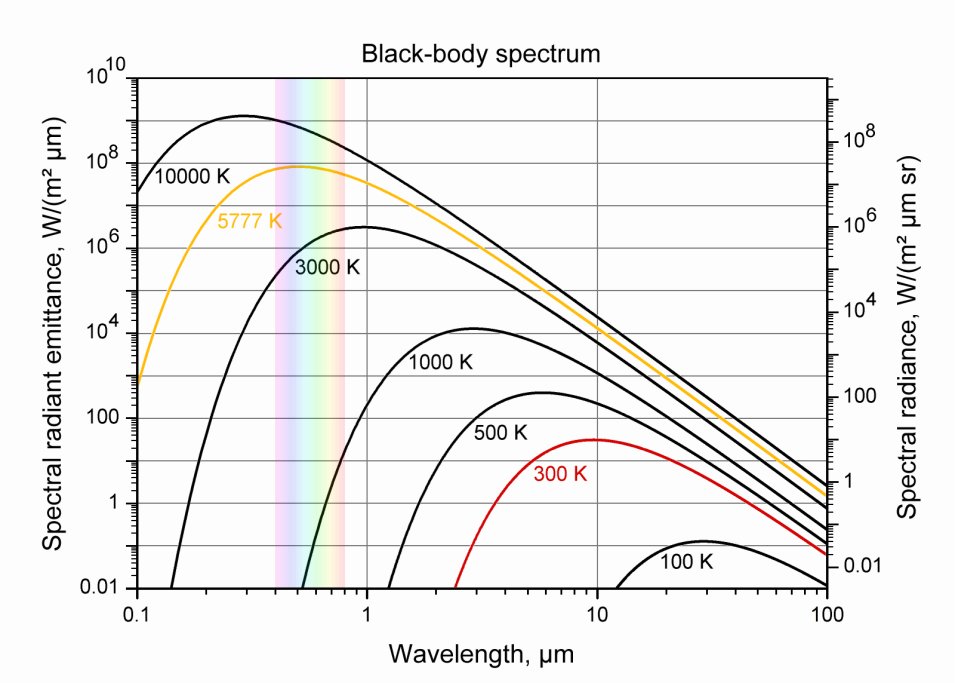


Figure 4-2B Spectrum for a Black Body. A normal ambient temperature (300K ca., the red line) emits light only in the infrared

4.2.1 Emissivity

The most important feature of a surface that affects the amount of energy radiating from it in stationary thermal conditions (fixed temperature) is its emissivity. If a surface whose temperature is to be

measured had the properties of a black body, the radiant emittance for fixed temperature and wavelength could be determined from Planck law [15].

However, under real conditions Planck law determines only a limiting (maximum) estimate of the thermal flux density. This is a consequence of the fact that all physical bodies have limited absorbing capacity: they do not satisfy Planck postulate referring to a perfect black body.

The emissivity ε of a body over the full radiation range, called the total emissivity, is the ratio of full-range radiant exitance $q(T)$ of that body to full-range radiant exitance $q_{black}(T)$ of a black body at the same temperature:

$$\varepsilon = \frac{q(T)}{q_{black}(T)}$$

Monochromatic emissivityε_λ is the ratio of monochromatic radiant emittance $q(\lambda, T)$ of a body at a given wavelength λ to monochromatic radiant exitance $q_{black}(\lambda, T)$ of a black body at the same wavelength, the same temperature and observed at the same angle:

$$\varepsilon_\lambda = \frac{q(\lambda, T)}{q_{black}(\lambda, T)}$$

Because we measure only IR-light, emissivity is considered a constant for every length wave. Moreover to increase the emissivity (and decrease the reflections) of the metal specimen, a high emissivity paint is applied.

The concept of Dissipative Body is a body has to be introduced. It's a grey body whose emissivity is independent of angle of observation. Its surface satisfies the conditions of Lambert's law (so it is called a Lambertian surface). Similarly, we can define a reflective body as a body whose reflectance is independent of angle of observation.

Remember that the emissivity coefficient ε a of a high-temperature object depends on angle of observation. As a reference, only when the surface temperature is much higher than the camera detector temperature, for cameras with cooled detectors.

The emissivity of a perfectly smooth metal surface as a function of wavelength λ (the relationship holds for λ > 2 μm)

$$\varepsilon = k\sqrt{\frac{\rho}{\lambda}}$$

where k=0.365$\Omega^{-½}$ is a constant coefficient and ρ the resistivity (Ωm). The emissivity of a real metal surface as a function of wavelength λ (Sala 1993) is:

$$\varepsilon = \frac{1}{b_1\sqrt{\lambda} + b_2}$$

where b_1 [$\mu m^{-0.5}$] and b_2 [0] are constant coefficients.

4.3 Non Destructive Evaluation

A type of non-destructive evaluation (NDE) is thermography. It consists of measuring the temperature field at surface of the object of study for detecting operating problems. Two main types, depending on the origin of heating, of thermography, exist with many variants [16]:

- Active thermography involves the use of an external heating and then measure the thermal response of the sample.
- Passive thermography where the thermal response of the sample, deriving from internal reasons, is measured

In our case the heating is produced by plastic heat dissipation so it's of the last type.

4.4 Thermographic camera

A thermographic camera (also called an infrared camera or thermal imaging camera) is a device that forms an image using infrared radiation, like a common camera that forms an image using visible light. Instead of the 400–700 nanometre range of the visible light camera, infrared cameras operate in wavelengths as long as 14 μm. Their use is called thermography.

The main steps in acquiring the temperature distribution are:

- The collection of photons in the case of a photon detector, or collection of heat energy with a thermal detector (such as a microbolometer).
- The detector produces, because of the collected energy, a signal voltage that results in a digital count through the system's A/D converter.
- Digital counts are transformed into radiance values after the camera is properly calibrated.
- Radiance values are converted to temperature using the known or measured emissivity of the target object from the calibrated camera's electronics.

Because the atmosphere has two bands of good transmission in the infrared range, most detectors and infrared (IR) cameras are divided in a natural way into:

- Short-wave (SW) devices, between 2-5 μm
- Long-wave (LW) devices, between 8-14 μm

Another classification follows from the detector type:

- With cooled detectors. Cooling can be done either with a dewar of liquid nitrogen, electrically with a thermoelectric cooler, or mechanically using a miniature Stirling cooler
- Non-cooled detectors, operating in the ambient temperature. It could be a camera with a microbolometric array of non-cooled detectors or arrays of pyroelectric detectors

4.4.1 Focal Plane Array

Radiation arriving at the detector is converted into electrical signals proportional to the radiant exitance of individual points of the image. Two-dimensional arrays of infrared detectors have been fabricated

and combined with charge-coupled device (CCD) electronics for read out, the CCD output from the detector array is converted into a digital signal and normalization procedures to correct for different pixel gains are performed numerically in real time. These arrays are called Focal Plane Arrays and they allow infrared cameras to be used like normal video cameras. High frame rates are achievable with these systems.

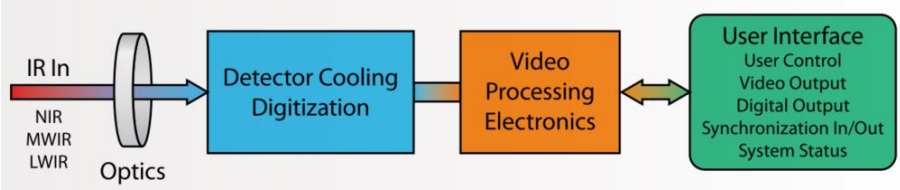


Figure 4-3 Scheme of the entire informations' flow from the object to the PC

The majority of commercial IR cameras have a microbolometer type detector, mainly because of cost considerations. A microbolometer responds to radiant energy in a way that causes a change of state in the bulk material (i.e., the bolometer effect). Generally, microbolometers do not require cooling, which allows compact camera designs that are relatively low in cost. Other characteristics of microbolometers are:

- Relatively low sensitivity
- Broad (flat) response curve
- Slow response time (time constant ~12ms)

Quantum detectors are instead used for more demanding applications like ours. They operate on the basis of an intrinsic photoelectric effect. These materials respond to IR by absorbing photons that elevate the material's electrons to a higher energy state, causing a change in conductivity, voltage, or current. By cooling these detectors to cryogenic temperatures, they can be very sensitive to the IR that is focused on them. They also react very quickly to changes in IR levels (i.e., temperatures), having a constant response time on the order of 1µs.

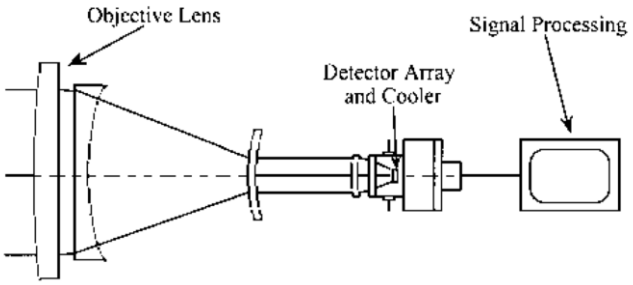


Figure 4-4 Focalplane array schematic of a IR-camera

4.4.2 Microcontrollers

In most video cameras, the microcontroller processing the measured signals is a very important component of the measurement process. An algorithm determines how temperature data are obtained from the detector's signals. Signal processing can be divided into the followings stages:

- Detection of infrared radiation in the array detector.
- Array calibration or mapping (i.e. linearization and temperature compensation of signals from single pixel).
- Processing of compensated signals by the camera measurement algorithm according to a suitable measurement model.

For most of the analysis, the important feature is the temporal dependence at certain pixel positions. Image processing algorithms can be used on the images to provide noise filtering or averaging.

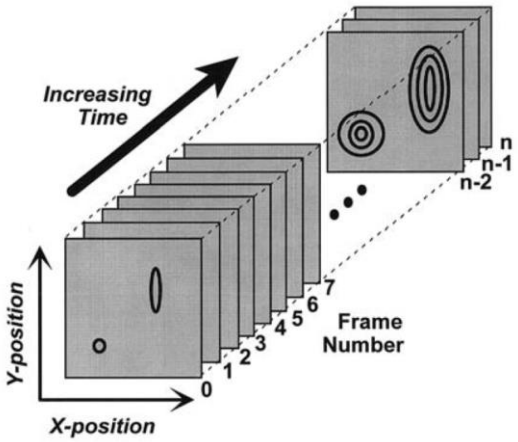


Figure 4-5 Schematic for data format used in thermography

4.4.3 Performance Measure

4.4.3.1 *Noise Equivalent Temperature Difference (NETD)*

It's the difference between the temperature of the observed object and the ambient temperature that generates a signal level equal to the noise level. It is also called the temperature resolution (it's not the accuracy).

NETD is defined as the ratio of the RMS noise voltage U_{noise} to the voltage increment ΔU_s generated by the difference in temperature between the measurement area of a technical black body (or test body) T_{ob} and background temperature T_o, divided by this difference:

$$NETD = \frac{U_{noise}}{\frac{\Delta U_s}{T_{ob} - T_o}} = \frac{T_{ob} - T_b}{\frac{\Delta U_s}{U_{noise}}}, [K]$$

NETD is defined as the minimum increment of the temperature difference that can be distinguished by a point detector for a given amplifier bandwidth.

The temperature resolution for different types of cameras has the following typical values:

- 10–30 mK – for cameras with QWIP detectors, designed for research and development applications;

- 50–100 mK – for measurement cameras;
- > 200 mK – for imaging cameras.

A typical NETD (noise equivalent temperature) for an InSb camera is in the order of 10–20 mK.

4.4.3.2 *Field of View (FOV)*

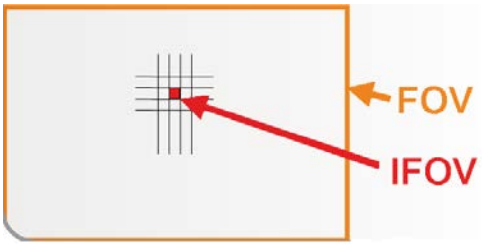
This determines the area that can be observed from a given distance d using the optics installed on the camera. This parameter determines the spatial (geometrical) resolution of a measurement IR camera. FOV is defined in meters and determines resolution in both the horizontal (H) and vertical (V) directions.

Field of view (FOV) versus distance d for 24° × 18° optics

d, m	0.50	1.0	2.0	5.0	10	30	100
H, m	0.20	0.41	0.81	2.0	4.1	12	41
V, m	0.15	0.31	0.62	1.5	3.1	9.3	31

4.4.3.3 *Instantaneous Field of View (Spatial Resolution)*

This determines the FOV of a single pixel in an array. From a practical point of view, it should be called the 'minimal field of view'. It is the second parameter determining the spatial (geometrical) resolution of a measurement IR camera. It is usually referred to as the spatial resolution and expressed in milliradians.



It is proportional to the distance of the object and depends on applied optics and on the number of pixel in the array.

An object situated closer to the camera fully irradiates one or more detectors in the array. When the object is warmer than the background, its temperature will be underrated if no detector is fully irradiated. On contrary when the background is warmer, the object temperature will be overrated.

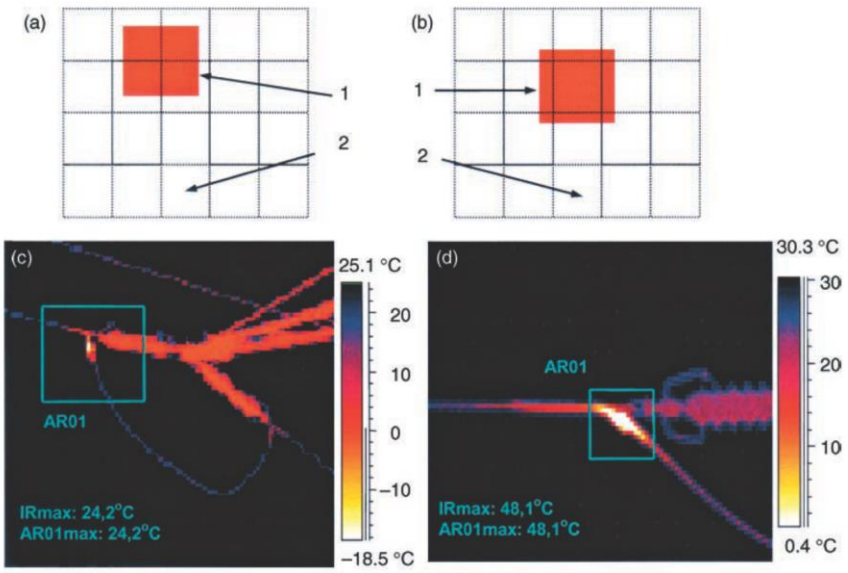


Figure 4-6 In the b) image, the object fully irradiates at least one detector, while in image a) the object does not fully irradiate any single detector. The maximum temperature measured from a long distance (image c)) is lower than that measured from a short distance. In the upper images: 1 is the object, 2 are the array detectors

In addition, each real optic distorts an image. These distortions result from chromatic and/or spherical aberration and many other imperfections of the optics.

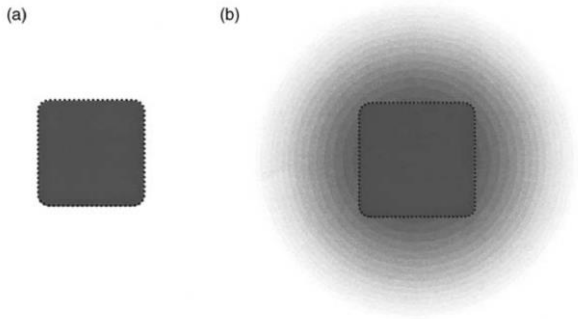


Figure 4-7 Determination of measurement area size: (a) ideal optics – irradiation of area of 2x2 detectors required for correct measurement; (b) real optics, image blurring – irradiation of area of4x 4 detectors required for correct measurement

5 THERMAL IMAGE PROCESSING

It is clear that the temperature at the surface of a specimen will be the result of a complex process but, neglecting the temperature gradients in the thickness, the problem can be drastically simplified to a two-dimensional diffusion problem. With this assumption, measuring the temperature field at the surface of a plane specimen can so be significant for the entire specimen [17], [18].

$$Q(x,y,t) = \rho C \frac{\partial T(x,y,t)}{\partial t} - k\left(\frac{\partial^2 T(x,y,t)}{\partial x^2} + \frac{\partial^2 T(x,y,t)}{\partial y^2}\right)$$

The scope of image processing is to obtain a denoised T(x,y,t) function to then calculate the heat source function Q(x,y,t).

Instead of performing the mean of the temperature for every frame, as Plekhov [19] and Maquin et al. [20] have done (the latter where working with uniaxial test), we proceed with a local approach focused on the process zone.

5.1 IR NON-DESTRUCTIVE TESTING

Infrared thermography in non-destructive testing provides images (thermograms) in which zones of interest (defects) appear sometimes as subtle signatures due to all factors that degrade infrared images from self-emission of the IR camera to the nonuniform properties of the surface where data are collected. Moreover, with long wavelengths in IR thermal bands with respect to visible bands, signals in the thermal bands are intrinsically weak since liberated photonic energy due to the oscillatory nature of particles inside matter is inversely proportional to the wavelength [21].

In the zone of interest appear sometimes as subtle signatures due to all actors that degrade infrared images from self-emission of the IR camera to the nonuniform proprieties of the surface where data are collected. The raw data has so to be processed.

IR images are mainly degraded by the following effects:

- Vignetting due to limited aperture
- Fixed pattern noise (FPN)
- Presence of dead pixel in the FPA matrix
- Radial distortion

Figure 5-1 Vignetting, also known as "light fall-off" (sometimes spelled "light falloff") is common in optics and photography, which in simple terms means darkening of image corners when compared to the centre. The image on the left suffers from vignetting

FPN could be cancelled out by subtracting an image of a uniform scene from the image of interest and if a map of dead pixel is known from the manufacturer.

Thermal units from the IR camera need to be converted into temperature. A calibration is needed. The procedure consists in position the IR camera in front of a reference temperature source (such as a blackbody or a thick Copper plate) brought to various known temperatures. As the reference temperature source is varied, the IR images are recorded.

5.2 Infrared Image Processing

The most common steps in infrared image processing can be listed as follow:

- Noise reduction: to augment the signal to noise ratio;
- Contrast balance: to highlight some features not evident in the original image;
- Edge detection: to define the discontinuities in the frame considered.

One of the most common procedures for noise reduction is the subtraction technique. Simple operations such as subtracting two images acquired at the same moment from two different experiments (**spatial reference technique**) or from images recorded closely (**temporal reference technique**) allow to remove unwanted effects present in both experiments such as non-uniform heating.

Moreover, the application of spatial linear or nonlinear filters is often exploited to reduce the high-frequency components characterizing thermal noise [22].

5.3 De noising

Many scientific datasets are contaminated with noise, either because of the data acquisition process, or because of naturally occurring phenomena (such as material inhomogeneity, electronic noises in the array). A first pre-processing step in analysing such datasets is denoising, that is, estimating the

unknown signal of interest from the available noisy data. There are several different approaches to denoise images.

Spatial filters have long been used as the traditional means of removing noise from images and signals. These filters usually smooth the data to reduce the noise, but, in the process, also blur the data.

A different class of methods exploits the decomposition of the data into the wavelet basis and shrinks the wavelet coefficients in order to denoise the data [23].

5.4 ALGORITHMS FOR IMAGE PROCESSING AND MATLAB IMPLEMENTATION

Image acquired are matrix that contains the temperature distribution. Two dimensions are required for the pixel position, one is containing the temperature data and the fourth specifies the number of frame (time).

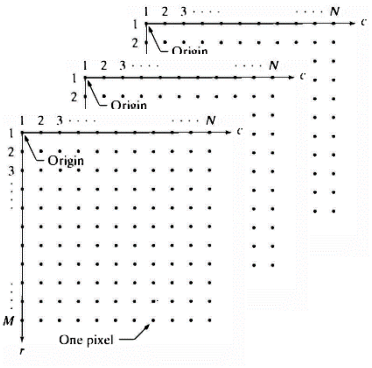


Figura 5-2 Image's matrix rapresentation

In MATLAB different classes of images exist: uint8, uint16, double, single, double in the range [0,1], logical.

We are using the class double in which numbers are double-precision floating point number because the input from the camera is the temperature, not a representative colour.

Intensity Trasformation and Spacial Filtering operate in the **spatial domain**. The spatial domain refers to the image plane itself, so all these methods operate directly on image's pixels. Two important categories of spatial domain processing are considered:

- Intensity (grey level) transformation
- Spatial filtering (also referred as neighbourhood processing or spatial convolution).

5.4.1 Spatial Filtering (or neighbourhood processing)

Neighbourhood processing consists of:

- Selecting a centre point, (x, y)
- Performing an operation that involves only the pixels in a predefined neighbourhood about (x, y)
- Letting the result of that operation be the "response" of the process at that point
- Repeating the process for every point in the image.

Use of the term linear spatial filtering differentiates this type of process from frequency domain filtering.

The linear operations of interest in this chapter consist of multiplying each pixel in the neighbourhood by a corresponding coefficient and summing the results to obtain the response at each point (x, y). If the neighbourhood is of size m X n, mn coefficients are required. The coefficients are arranged as a matrix (called a filter, mask, filter mask, kernel, template, or window) with the first three terms being the most prevalent. The terms convolution filter, convolution mask, or convolution kernel, also are used.

5.4.1.1 *Mean of a Neighborhood*

Mean filter reduce the amount of intensity variation between one pixel and the next to reduce noise. Replace each pixel value in an image with the average value of its neighbours, including itself. This has the effect of eliminating pixel values which are unrepresentative of their surroundings. Often a 3×3 square kernel is used although larger kernels (*e.g.* 5×5 squares) can be used for more severe smoothing.

$$H(i,j) = \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$I'(u,v) = \sum_{i=-1}^{1}\sum_{j=-1}^{1} I(u+i, v+j)\cdot H(i,j)$$

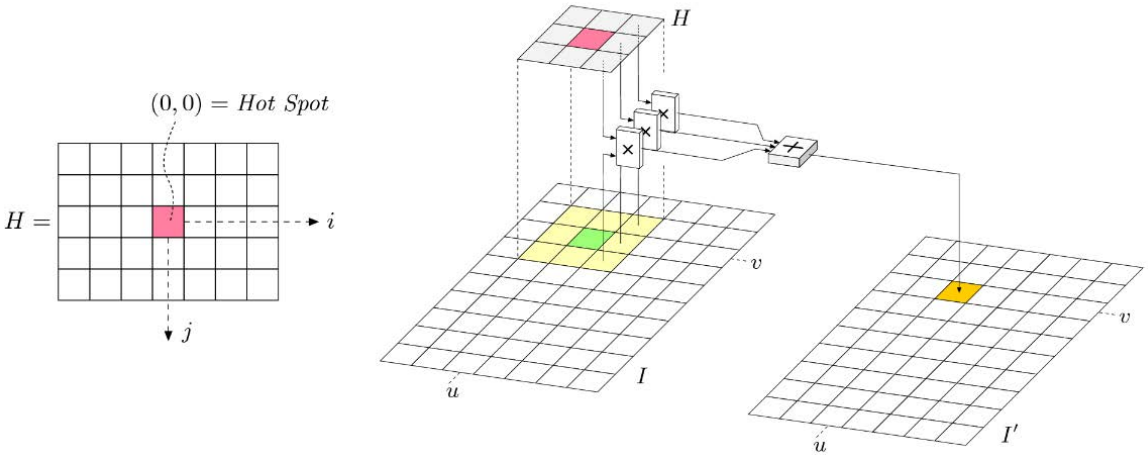


Figure 5-1 'I' is the original image, H is the kernel of the filter and I' the final image

Some problems emerge when handling pixels close to boundaries. One solution is to pad the borders with zeros

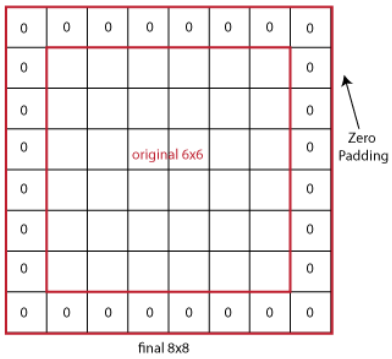


Figure 5-2 Image padding

One other is to crop the region of the image to be filtered to avoid the problem.

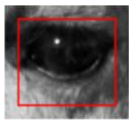


Figure 5-3 Image cropping

One other solution is the extend the borders.

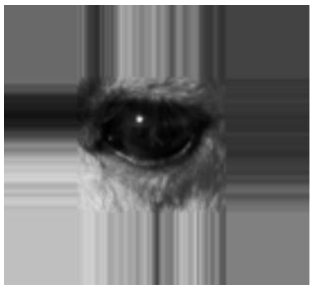


Figure 5-4 Image extending

And the last solution is to wrap the image.

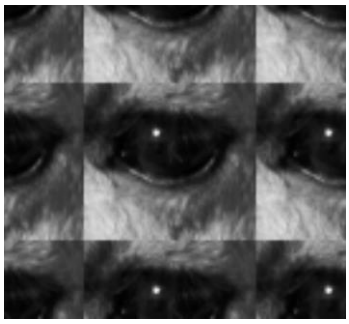


Figure 5-5 Image wrapping

5.4.1.2 *Wiener Filter*

In signal processing, the Wiener filter is a filter used to produce an estimate of a desired or target random process by linear time-invariant (LTI) filtering of an observed noisy process, assuming known stationary signal and noise spectra, and additive noise. The Wiener filter minimizes the mean square error between the estimated random process and the desired process.

The filter uses a pixelwise adaptive Wiener method based on statistics estimated from a local neighbourhood of each pixel, using neighbourhoods of size `m`-by-`n` to estimate the local image mean and standard deviation.

This approach often produces better results than linear filtering. The adaptive filter is more selective than a comparable linear filter, preserving edges and other high-frequency parts of an image. In addition, there are no design tasks, however, does require more computation time than linear filtering.

Wiener Filter works best when the noise is constant-power ("white") additive noise, such as Gaussian noise.

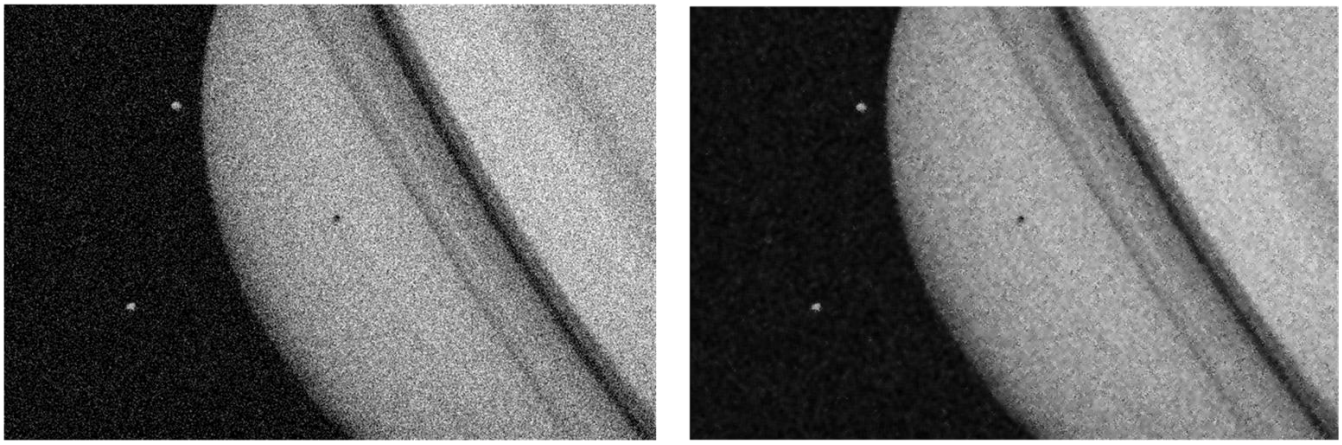


Figure 5-6 On the left the original image, on the right with the Wiener filter applied

It estimates the local mean and variance around each pixel.

$$\mu = \frac{1}{NM} \sum_{n_1,n_2 \in \eta} a(n_1, n_2) \quad \sigma^2 = \frac{1}{NM} \sum_{n_1,n_2 \in \eta} a^2(n_1, n_2) - \mu^2$$

where η is the N-by-M local neighbourhood of each pixel in the image. `wiener2` then creates a pixelwise Wiener filter using these estimates,

$$b(n_1, n_2) = \mu + \frac{\sigma^2 - \nu^2}{\sigma^2} (a(n_1, n_2) - \mu)$$

where ν^2 is the noise variance. If the noise variance is not given, `wiener2` uses the average of all the local estimated variances.

This kind of filters (even the nonlinear) are too simple for ours application. It's better to operate in the **frequency domain**.

5.4.1.3 *Filtering in the Frequency domain*

With this type of filtering we can perform lowpass filtering for image smoothing, highpass filtering (including high-frequency emphasis filtering) for image sharpening, and selective filtering for the removal of periodic interference.

Let f(x,y) for x=0, 1, 2, …, M-1 and y=0, 1, 2, …, N-1 denote a digital image of size M x N pixels. The 2D discrete Fourier transform (DFT) of f(x,y) denoted by F(u, v) is:

$$\mathrm{F}(\mathrm{u},\mathrm{v}) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-j2\pi(\frac{\mathrm{ux}}{\mathrm{M}}+\frac{\mathrm{vy}}{\mathrm{N}})}$$

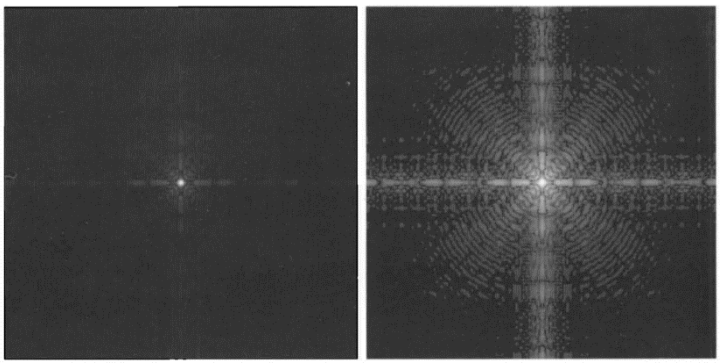


Figura 5-3 On the left a Fourier spectrum; on the right the spectrum after a log transformation to visualize it better

Defined for u=0, 1, 2, …, M-1 and v=0, 1, 2, …, N-1. The exponential could be expanded into sine and cosine wave where u and v determine the functions frequencies.

F(0,0), the value at the origin of the frequency domain, is equal to MN times the average value of f(x,y)

The frequency domain is the coordinate system spanned by F(u, v) with u and v as (frequency) variables. The M x N rectangular region defined by u=0, 1, 2, …, M-1 and v=0, 1, 2, …, N-1 is often referred to as the frequency rectangle. Clearly, the frequency rectangle is of the same size as the input image.

Even if f(x, y) is a real function, its transform is complex in general. The principal method for analysing a transform visually is to compute its spectrum (the magnitude of F(u, v), which is a real function) and display it as an image. Letting R(u, v) and I(u, v) represent the real and imaginary components of F(u, v), the Fourier spectrum is defined as

$$|\mathrm{F}(\mathrm{u},\mathrm{v})| = \sqrt{R^2(\mathrm{u},\mathrm{v}) + I^2(\mathrm{u},\mathrm{v})}$$

The phase angle of the transform is defined as

$$\varphi(\mathrm{u},\mathrm{v}) = \tan^{-1}\frac{\mathrm{I}(\mathrm{u},\mathrm{v})}{\mathrm{R}(\mathrm{u},\mathrm{v})}$$

These two functions can be used to express the complex function F(u, v) in its polar form:

$$\mathrm{F}(\mathrm{u},\nu) = |\mathrm{F}(\mathrm{u},\nu)|e^{j\varphi(\mathrm{u},\nu)}$$

The Power Spectrum is defined as the square of the magnitude:

$$\mathrm{P}(\mathrm{u},\nu) = |\mathrm{F}(\mathrm{u},\nu)| = R^2(\mathrm{u},\nu) + I^2(\mathrm{u},\nu)$$

If f(x, y) is real, its Fourier transform is conjugate symmetric about the origin; that is,

$$F(u, \nu) = F'(-u, -\nu)$$

This implies that the Fourier spectrum is symmetric about the origin also:

$$|F(u, \nu)| = |F'(-u, -\nu)|$$

It can be shown that the DFT is infinitely periodic in both the u and v directions, with the periodicity determined by M and N. Periodicity is a property of the inverse DFT also. An image obtained by taking the inverse Fourier transform is also infinitely periodic.

Considering the spectrum of a one-dimensional transform F(u), the periodicity propriety indicates that F(u) has a period of length M, and the symmetry propriety indicates that |F(u)| is centred on the origin. The values of |F(u)| from M/2 to M-1 are repetitions of the values in the half period to the left of the origin. It follows that computing the transform yields two back to back periods in the interval [0,M-1].

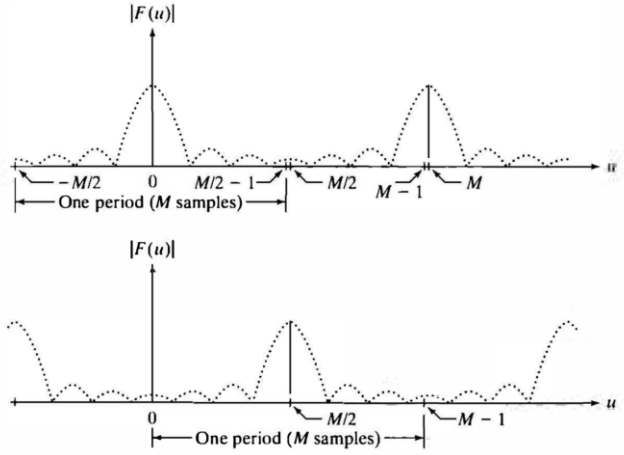


Figure 5-7 On the upper side the original spectrum. On the bottom the shifted spectrum

The desired period can be obtained by multiplying f(x) by $(-1)^x$ prior to computing the transform, moving the origin of the transform.

A similar process can be executed for the two dimensional DFT.

Visual analysis of the spectrum is simplified by moving the values at the origin of the transform to the centre of the frequency rectangle. The value of the spectrum at coordinates (M/2, N/2) of the original transformation is the same as its value at (0,0) on the shifted transform.

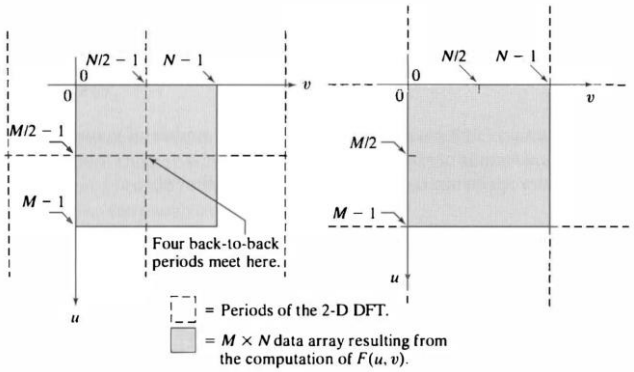


Figure 5-4 On the left: (MxN) Fourier spectrum (shaded), showing four back-to-back quarter periods. On the right: spectrum after multiplying f(x,y) by (-1)^x+y prior to computing the Fourier transform. The shaded period is the data that would be obtained by using the DFT

An example of application of the DFT is made to a simple rectangle.

Figure 5-8 Image of a rectangle for which the spectrum will be calculated

The corresponding DFT is the following. More energy is seen, particularly at the higher frequencies, along the vertical axis because the object's vertical cross sections appear as a narrow pulse. The border horizontal cross sections produce frequency characteristics that fall off rapidly at higher frequencies.

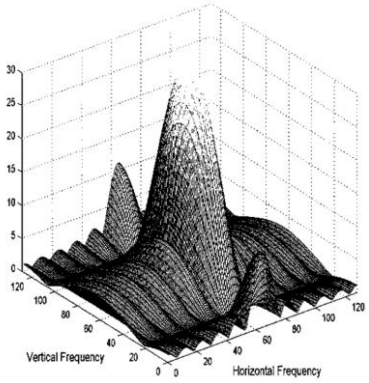


Figure 5-9 Spectrum of the horizontal rectangle. Low frequencies components for the horizontal direction are evident

For real images the analysis is less intuitive.

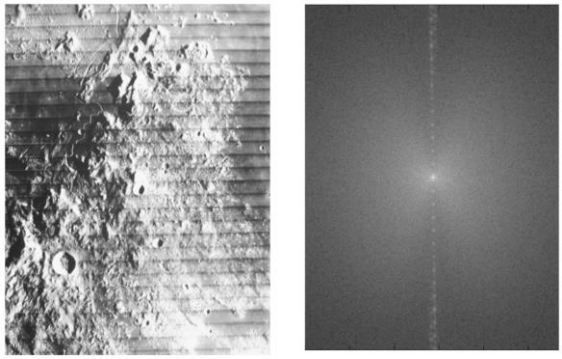


Figure 5-10 Moon surface with a periodic horizontal noise and its corrispective spectrum

Especially the ones in which there isn't repetition.

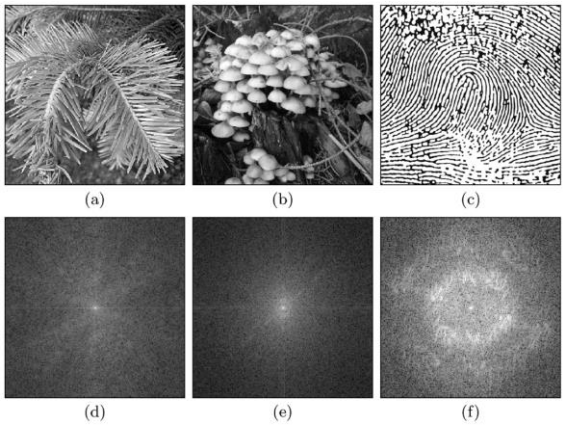


Figure 5-11 Natural images and their spectrum. Even if some repetition is present, it isn't evident in the spectra

Even if there is some repetition it won't stand out in spectra.

5.4.2 Lowpass Frequency Domain Filters

A low pass filter keeps the frequencies below a specified value. An ideal low pass filter works by multiplying the Fourier Transform by the value 1 if (u,v) is closer the centre than the specified value, and zero otherwise.

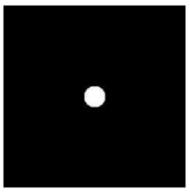


Figure 5-12 Sharp low-pass filter

But applying this sharp filter it to an image causes some problems as can be see:

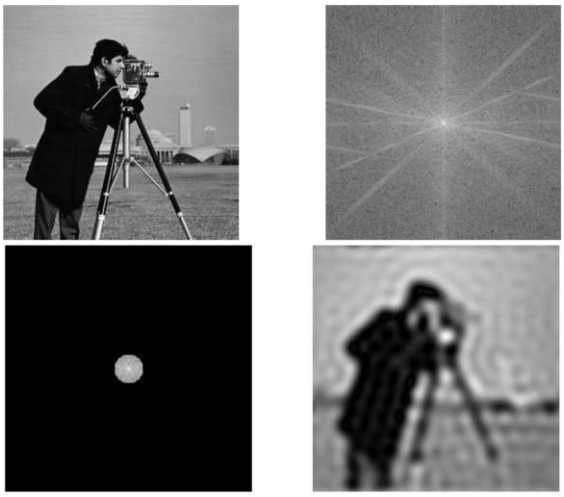


Figure 5-13 On the first row the original image and its DFT, on the second row the spectrum with the sharp filter applied and then the inverse transformation (filtered image)

5.4.3 Gaussian Filtering

Mathematically, applying a Gaussian blur to an image is the same as convolving the image with a Gaussian function [24].

Since the Fourier transform of a Gaussian is another Gaussian, applying a Gaussian blur has the effect of reducing the image's high-frequency components; a Gaussian blur is thus a low pass filter. The function is simply the product of two 1D Gaussian functions (one for each direction) and, the transformation to apply to each pixel, is given by:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

The Gaussian filter works by using the 2D distribution as a point-spread function and it's assumed to have mean of zero. This is achieved by convolving the 2D Gaussian distribution function with the image. A discrete approximation to the Gaussian function is needed, theoretically it would require an infinitely large convolution kernel as the distribution is non-zero everywhere. 99% of the distribution falls within 3 standard deviations.

Figure 5-14 Examples of different values of sigma applied

This means we can normally limit the kernel size to contain only values within three standard deviations of the mean.

5.4.3.1.1 3-Dimensional Gaussian Filtering

Three-dimensional Gaussian smoothing in the frequency domain can be performed. Anisotropic smoothing is achieved by the following formula [25]:

$$G(\mathrm{x},\mathrm{y},\mathrm{z}) = \frac{1}{(2\pi)^{\frac{3}{2}}\sigma_x\,\sigma_y\,\sigma_z}\mathrm{e}^{-\frac{2x^2}{\sigma_x^2}-\frac{2y^2}{\sigma_y^2}-\frac{2z^2}{\sigma_z^2}}$$

This permits to smooth also in the time dimension with a different sigma (avoiding using a general 3D gaussian filter). The results of this process can be read on chapter 5.

5.4.4 Other Types of Filters

5.4.4.1 *Kalman Filter*

It's an algorithm that uses a series of measurements observed over time, containing statistical noise and other inaccuracies, and produces estimates of unknown variables that tend to be more precise than those based on a single measurement alone.

They are especially used in object motion tracking but more recently are being used also for video denoising (our images are like a high frame rate video) [26].

The math behind the filter is quite complex and it transcend the scope of this work but to try to understand the main concept, an example on object taking is made.

For non-linear systems, an extended Kalman filter can be used.

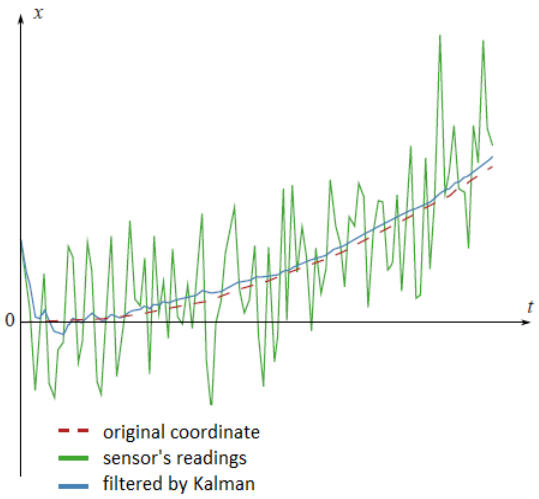


Figure 5-15 Kalman filter applied on one dimensional time varying signal

5.4.4.1.1 Kalman Filter: a more common example

In this example, a simple state of the system contains only position and velocity.

We don't know what the actual position and velocity are; there are a whole range of possible combinations of position and velocity that might be true, but some of them are more likely than others. The Kalman filter assumes that variables are random and Gaussian distributed. Each variable has a mean value μ, which is the center of the random distribution (and its most likely state), and a variance σ2, which is the uncertainty:

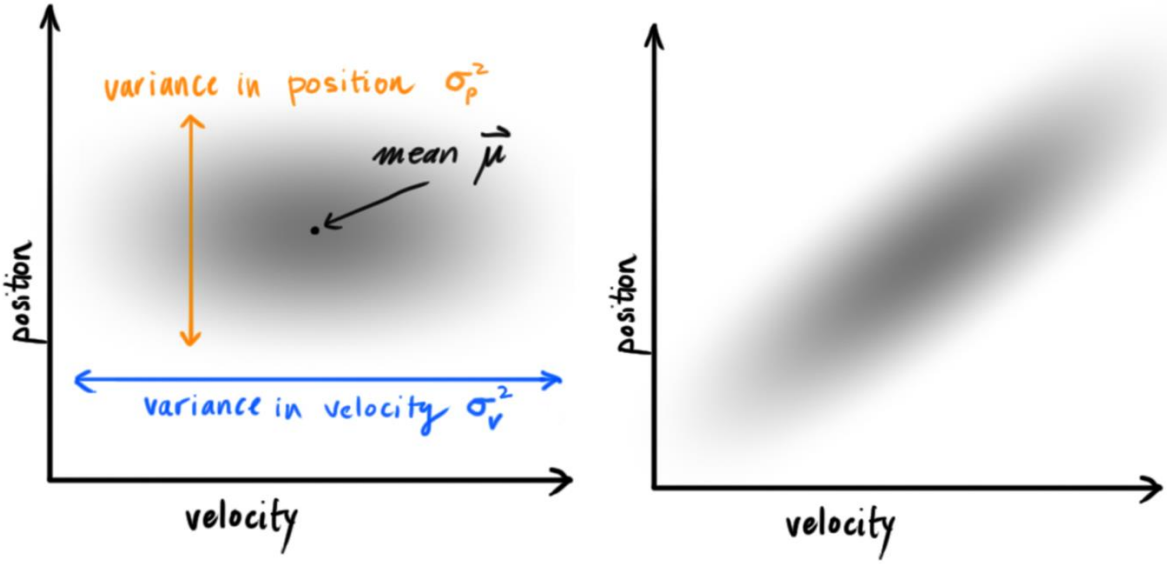


Figure 5-16 On the left plot of two uncorrelated variables. On the right plot of correlated variables [27]

Considering only a state of the system, position and velocity are **uncorrelated** (the state of one variable tells nothing about the other). If we start estimating a new position based on an old one, the two variables start to be correlated. If the velocity was high, the object probably moved farther, so the final position will be more distant. If it's moving slowly, it didn't get as far. The goal of the Kalman filter is to obtain more information by comparing the different states of the system during time.

This correlation is captured by a covariance matrix. Each element of the matrix Σij is the degree of correlation between the *ith* state variable and the *jth* state variable. (the covariance matrix is symmetric, which means that it doesn't matter if you swap *i* and *j*).

Every state in our original estimate could have moved to a range of states. Each state point (consisting of position and velcoity) is moved to somewhere inside a Gaussian blob with a certain covariance (the untracked influences are treated as noise with the same covariance).

The new best estimate is a prediction made from previous best estimate. And the new uncertainty is predicted from the old uncertainty, with some additional uncertainty from the environment.

5.4.4.2 *Kalman-like Filter for video*

This version, created by Christopher Philip Mauer, implements a recursive prediction/correction algorithm which is based on the Kalman Filter to remove high gain noise from time lapse image streams. These filters remove camera/detector noise while recovering faint image detail.

The user can specify:

- Initial noise variance estimate
- Filter gain

Wrong guesses for the initial variance will not prevent noise estimation, but merely delay the fitting process. High filter gain renders the output less sensitive to momentary fluctuations.

The filter performs the following steps:

1) Initialization
    - User input specifies the filter gain: G
    - User input specifies the noise variance estimate: V
    - Use the first image as the prediction seed: $I_{-k} = I_k$
    - Use the variance estimate as the error seed: $E_{-k} = V$
2) Correction
    - Compute the Kalman Gain: $K_k = E_{-k}/(E_{-k} + V)$
    - Update the image prediction with the (M_k) measurement: $I_k = G*I_{-k} + (1.0-G)*M_k + K_k(M_k-I_{-k})$
    - Update the variance estimate: $E_k=E_{-k}(1-K_k)$
3) Prediction
    - Predict the next image: $I_{-k+1} = I_k$
    - Predict the variance: $E_{-k+1} = E_k$
4) Update values
    - $E_{-k}=E_{-k+1}$
    - $I_{-k}=I_{-k+1}$
5) Repeat 2,3,4

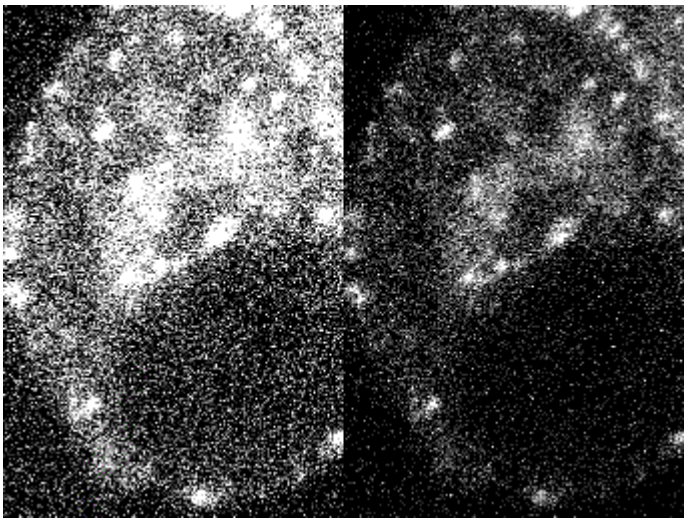


Figure 5-17 On the left original frame from a sequence, oh the right the filtered one with less noise [28]

5.5 Wavelets (Discrete Wavelet Transform)

The main aim of an image denoising algorithm is to achieve both noise reduction and feature preservation. In this context, wavelet-based methods are of particular interest [23].

Wavelet schemes are especially suitable for applications where scalability and tolerable degradation are the important considerations. Wavelet transform decomposes a signal into a set of basis functions. Wavelets are derived from a single prototype wavelet ψ called mother wavelet by scaling and translation.

Unlike Fourier bases, Wavelet transform provides excellent time and frequency representation simultaneously. With the sub sampling property, the performance of the Wavelet transform can be realized using iterative filter bank structures. Every time the filter bank is iterated, the number of samples for the next stage is halved so that only one sample is left at the end.

Considering an image f(x,y) of size M x N, a forward general discrete transform F(u, v) can be expressed as:

$$\mathrm{F}(\mathrm{u},\mathrm{v}) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\; g_{\mathrm{u},\mathrm{v}}\,(x,y)$$

Where x and y are spatial variables, u and v transform domain variables and $g_{\mathrm{u},\mathrm{v}}$ is the Forward Transformation Kernel. To reobtain f(x,y) given $\mathrm{F}(\mathrm{u},\mathrm{v})$:

$$f(x,y) = \sum_{\mathrm{u}}\sum_{\mathrm{v}} \mathrm{F}(\mathrm{u},\mathrm{v})\; h_{\mathrm{u},\mathrm{v}}\,(x,y)$$

Where $h_{\mathrm{u},\mathrm{v}}$ is the Inverse is the Inverse Transformation Kernels. The $\mathrm{F}(\mathrm{u},\mathrm{v})$ can be viewed as the expansion coefficients of a series expansion of a series expansion of f(x,y) with respect to $h_{\mathrm{u},\mathrm{v}}$.

The discrete Fourier transform expansion series would be

$$h_{\mathrm{u},\mathrm{v}}(x,y) = g^{*}_{\mathrm{u},\mathrm{v}}(x,y) = \frac{1}{\sqrt{MN}}\, e^{2j\pi(\frac{ux}{M}+\frac{vy}{N})}$$

Where j is the imaginary unit, * is the complex conjugate operator, u and v represent horizontal and vertical frequency respectively. In the DWT instead there are 3 variables (instead of special frequencies): scale j, horizontal and vertical translation, n and m.

The Discrete Wavelet Transform refers to a class of transformation that differ not only in the transformation kernels employed (and thus the expansion functions used) but also in the fundamental nature of those functions (like orthonormal or biorthogonal basis).

Different kernels pair can be chosen, their expansion functions are small waves of finite duration and varying frequency.

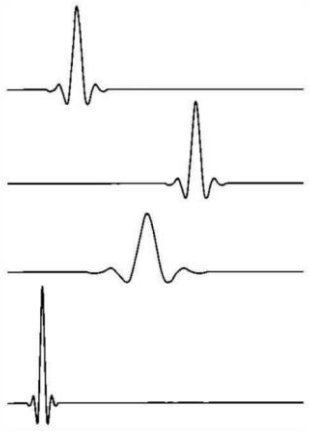


Figure 5-18 DWT expansion function are small waves of finite duration and varying frequency

The kernels can be represented as three separable 2D wavelets.

$$\psi^H(x,y) = \psi(x)\varphi(y)$$

$$\psi^V(x,y) = \varphi(x)\psi(y)$$

$$\psi^D(x,y) = \psi(x)\psi(y)$$

These 2D wavelets are respectively the Horizontal, Vertical and Diagonal wavelets. φ is the Scaling Function and is separable: $\varphi(x,y) = \varphi(x)\varphi(y)$.

Each of these 2-D functions is the product of two 1-D real, square-integrable scaling and wavelet functions:

$$\varphi_{j,k}(x) = 2^{\frac{j}{2}}\,\varphi(2^j x - k)$$

$$\psi_{j,k}(x) = 2^{\frac{j}{2}}\,\psi(2^j x - k)$$

Translation k determines the position of these 1-D functions along the x-axis, scale j determines their width (how broad or narrow they are along x) and $2^{\frac{j}{2}}$ controls their height or amplitude. Note that the associated expansion functions are binary scaling and integer translates of mother wavelet $\psi(x) = \psi_{0,0}(x)$ and scaling function $\varphi(x) = \varphi_{0,0}(x)$.

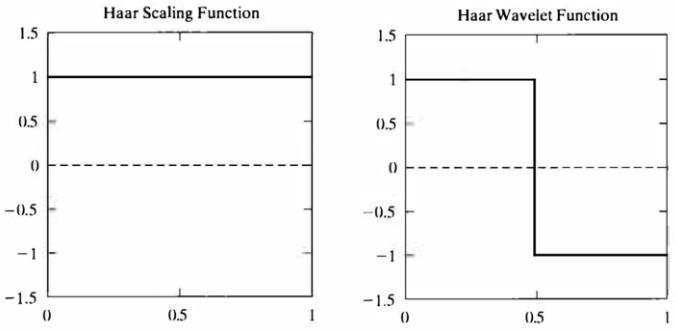


Figure 5-19 Wavelet (right) and Scaling function for Haar

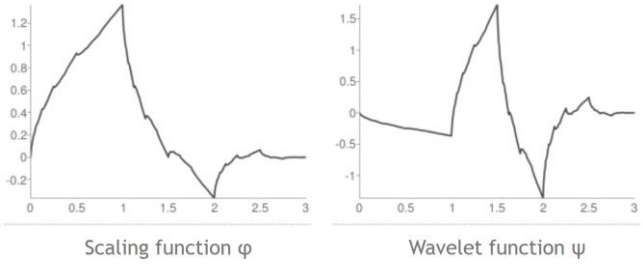


Figura 5-5 Wavelet and Scaling functions for Symlets 2 (sym2)

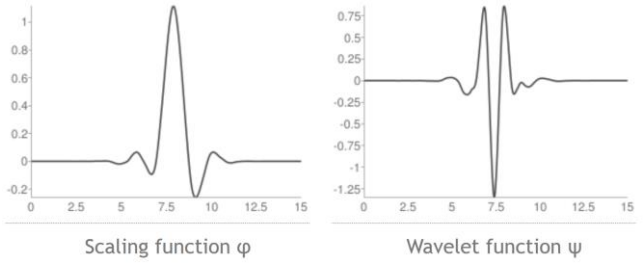


Figura 5-6 Wavelet and Scaling functions for Symlets 8 (sym8)

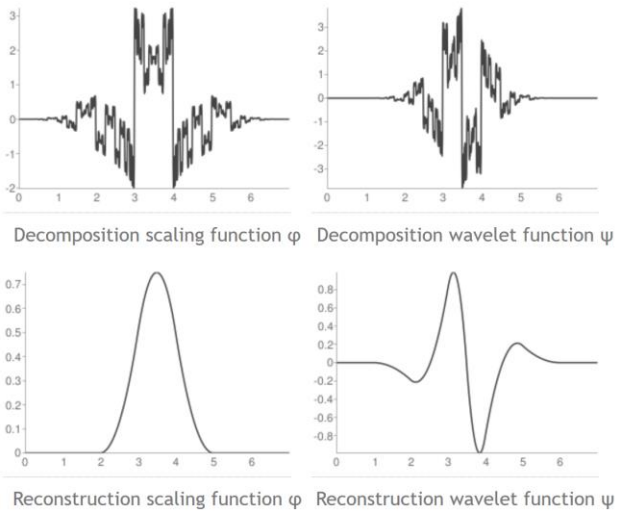


Figura 5-7 Wavelet and Scaling functions for Biorthogonal 3.3 (bior3.3)

Both the scaling function $\varphi(x)$ and the mother wavelet $\psi(x)$ can be expressed as linear combinations of double-resolution copies of themselves.

$$\varphi(x) = \sum_n h_\varphi(n)\sqrt{2}\,\varphi(2x-n)$$

$$\psi(x) = \sum_n h_\psi(n)\sqrt{2}\,\varphi(2x-n)$$

Where the expansion coefficients h_φ and h_ψ are called scaling and wavelet vectors. They are the filter coefficients of the fast wavelet transform (FWT), an iterative computational approach to the DWT.

There are four types of output for every scale j: $W_\varphi(j,m,n)$, $W_\varphi^H(j,m,n)$, , $W_\varphi^V(j,m,n)$, $W_\varphi^D(j,m,n)$.

Approximation coefficient are created decomposing the signal simultaneously using a high-pass filter $h_\psi(-m)$ and low-pass filter $h_\varphi(-n)$. The outputs giving the detail coefficients (from the high-pass filter) and approximation coefficients (from the low-pass). It is important that the two filters are related to each other and they are known as a quadrature mirror filter. After every filter a Downsampling is performed extracting every other point from a sequence of points. In first iteration $W_\varphi(j+1,m,n) = f(x,y)$, so the image serves as the input.

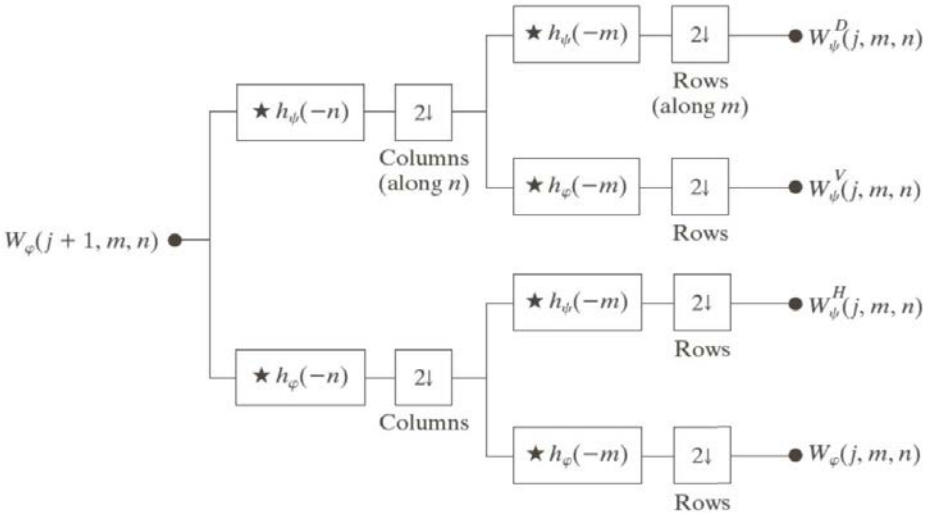


Figure 5-20 The 2-D fast wavelet transform (FWT) filter bank. Each pass generates one DWT scale. In the first iteration $W_\varphi(j+1,m,n) = f(x,y)$

The $W_\varphi(j,m,n)$ coefficients are created via two lowpass filter and are called approximation coefficients. $W_\varphi^H(j,m,n)$, , $W_\varphi^V(j,m,n), W_\varphi^D(j,m,n)$ are the horizontal, vertical and diagonal detail coefficients.

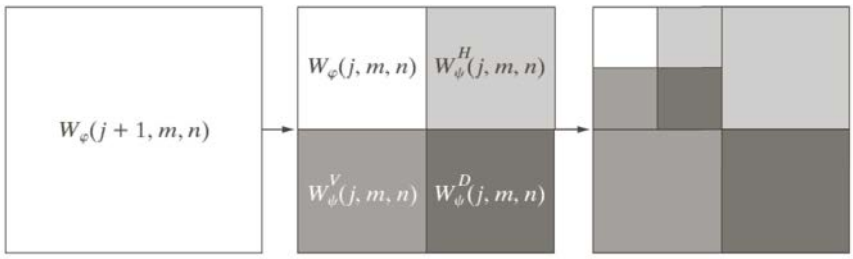


Figure 5-21 Original image, first level of decomposition, first and second level of decomposition

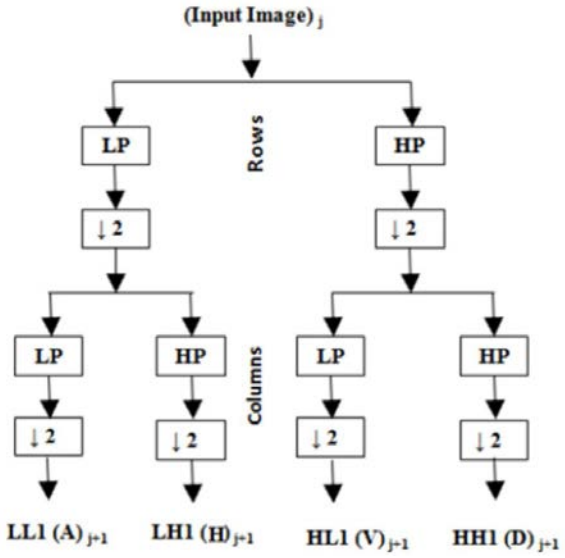


Figure 5-22 Wavelet decomposition (LP: Low pass filter, HP: High pass filter, A: Approximation coefficients, H, V, D: Horizontal, Vertical and Diagonal detail coefficients respectively, 'j' is the scale)

The output $W_\varphi(j, m, n)$ can be used as a subsequent input $W_\varphi(j+1, m, n)$ to the block diagram for creating even lower resolution components.

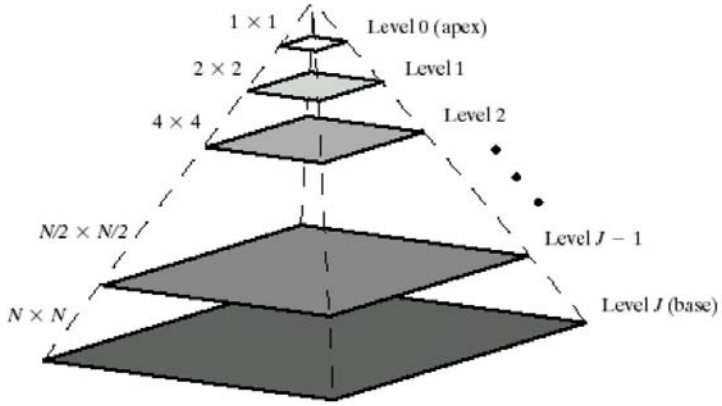


Figure 5-23 Pyramidal structured image deriving from wavelet decomposition

At each level, we just store the differences (residuals) between the image at that level and the predicted image from the next level. The image can be reconstructed by just adding up all the residuals.

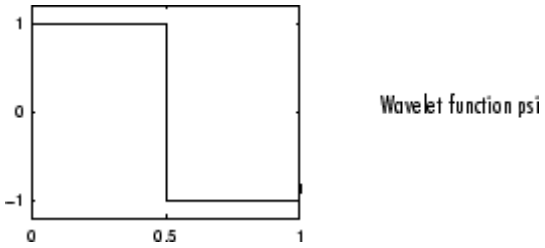
The wavelet decomposition of an image is carried out as follows: in the first level of decomposition, the image is split into 4 subbands, namely the HH, HL, LH and LL subbands. The HH subband gives the diagonal details of the image; the HL and LH subbands give the horizontal and vertical features respectively. The LL subband is the low resolution residual consists of low frequency components and its subbands are further split at higher levels of decomposition.

The coefficients for the various details and of the desired decomposition level could be zeroed or tresholded by a thresholding vector.

5.5.1 Common Wavelet Families implemented in MATLAB

5.5.1.1 *Haar*

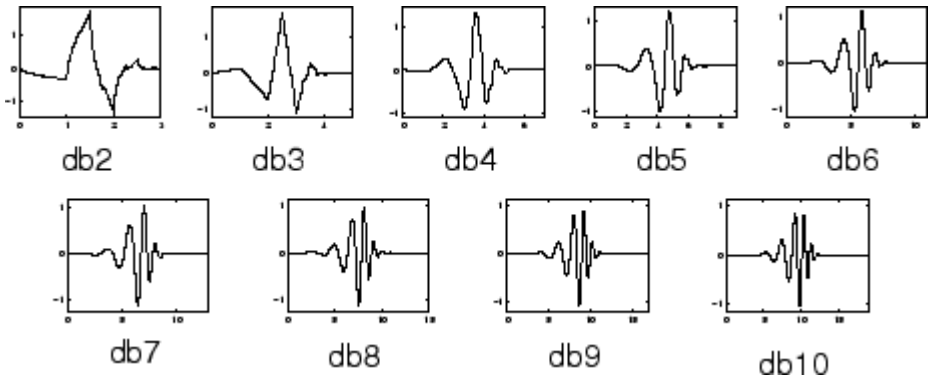
Any discussion of wavelets begins with Haar wavelet, the first and simplest. The Haar wavelet is discontinuous, and resembles a step function. It represents the same wavelet as Daubechies `db1`.



5.5.1.2 *Daubechies*

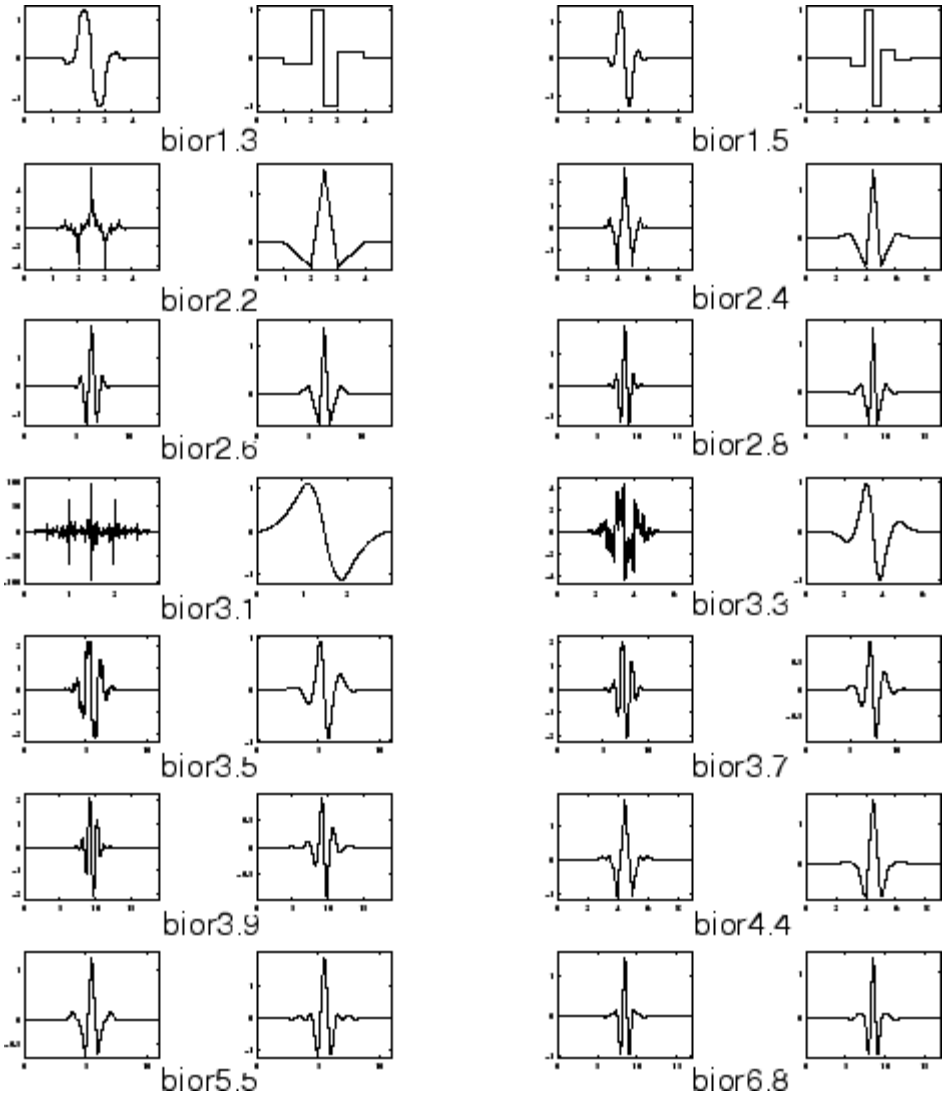
Daubechies invented what are called compactly supported orthonormal wavelets, making discrete wavelet analysis practicable.

The names of the Daubechies family wavelets are written `dbN`, where `N` is the order, and `db` the "surname" of the wavelet. The `db1` wavelet, as mentioned above, is the same as `Haar` wavelet. Here are the wavelet functions psi of the next nine members of the family:



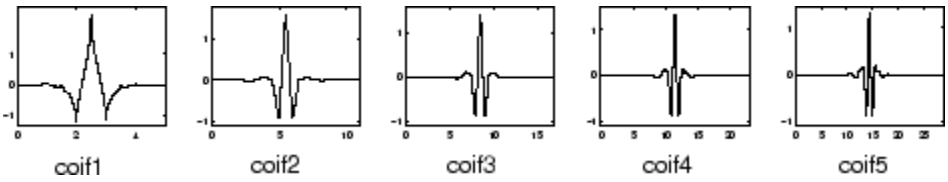
5.5.1.3 *Biorthogonal*

This family of wavelets exhibits the property of linear phase, which is needed for signal and image reconstruction. By using two wavelets, one for decomposition (on the left side) and the other for reconstruction (on the right side) instead of the same single one, interesting properties are derived.



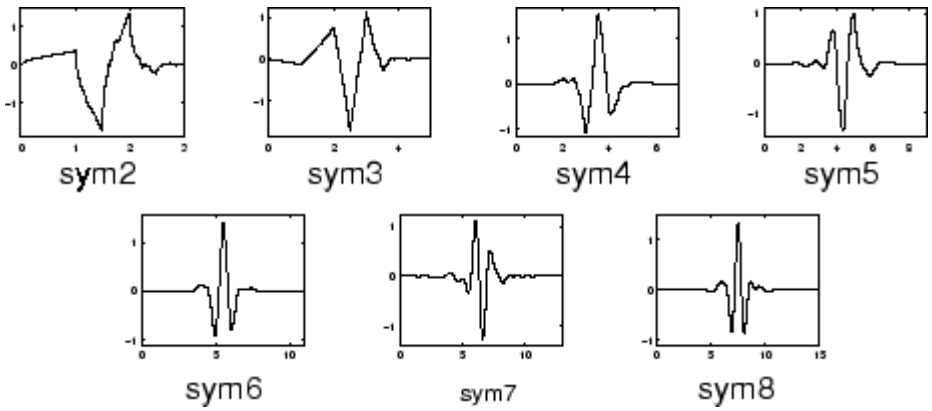
5.5.1.4 *Coiflets*

Built by Daubechies at the request of R. Coifman. The wavelet function has 2*N* moments equal to 0 and the scaling function has 2*N*-1 moments equal to 0. The two functions have a support of length 6*N*-1.



5.5.1.5 *Symlets*

The symlets are nearly symmetrical wavelets proposed by Daubechies as modifications to the db family. The properties of the two wavelet families are similar. Here are the wavelet functions psi.



5.5.2 Image Denoising

Under Besov norms, the magnitudes of wavelet coefficients are directly proportional to the irregularity of a given image. When noises are involved, such irregularity grows in the wavelet coefficients [29].

Noise commonly manifests itself as fine-grained structure in the image and DWT provides a scale based decomposition. Thus, most of the noise tends to be represented by wavelet coefficients at the finer scales. Discarding these coefficients would result in a natural filtering of the noise on the basis of scale. Because the coefficients at such scales also tend to be the primary carriers of edge information, this method threshold the DWT coefficients to zero if their values are below a threshold. These coefficients are mostly those corresponding to noise.

Wavelet based denoising techniques follow the similar steps irrespective of the shrinkage function.

1) Read the noisy image as input
2) Perform 2D Discrete Wavelet Transform and obtain Wavelet Coefficients (Subbands)
3) Estimate noise variance from the noisy image.
4) Calculate the threshold using suitable nonlinear shrinkage function.
5) Apply soft thresholding.
6) Perform inverse 2D Discrete Wavelet Transform on the thresholded wavelet coefficients.
7) Obtain the denoised image
8) Evaluate the quality of the denoised image.

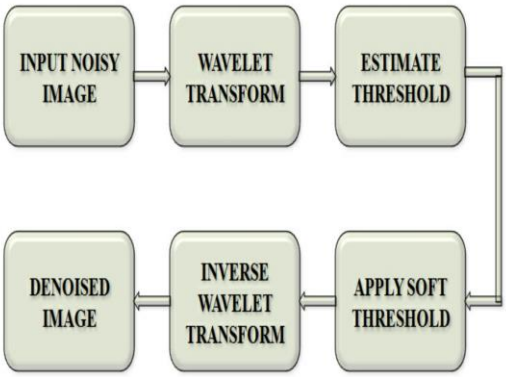


Figure 5-24 Scheme for wavelet filtering

The performance of the denoising algorithm relies on the optimal value of threshold. Fixing an optimal threshold is not an easy task but different methods will be provided.

5.5.3 Tresholding (shrink)

In the wavelet domain, the noise is uniformly spread throughout coefficients while most of the image information is concentrated in a few large ones. Noise gives many small coefficients, spread more or less everywhere in the parameter space. Therefore, the first wavelet-based denoising methods were based on thresholding of detail subands coefficients: putting to zero all the coefficients that lie below a fixed treshold. However, most of the wavelet thresholding methods suffer from the drawback that the chosen threshold may not match the specific distribution of signal and noise components at different scales and orientations [30].

If the threshold is too large, noisy components may not be eliminated. On the other hand if the threshold is too small, it may remove the image details also resulting in overly smoothed images. The inefficient threshold may affect the edge details; this may degrade the visual quality [31].

One has to choose between two versions:

- **Hard Thresholding**: here the small coefficients, are replaced by 0 and the rest remains untouched. As a consequence, artificial discontinuities are created.
- **Soft Thresholding** or **Wavelet Shrinkage**: in order to remove these discontinuities, all the remaining coefficients are shifted by ± T, so as to make them continuous.

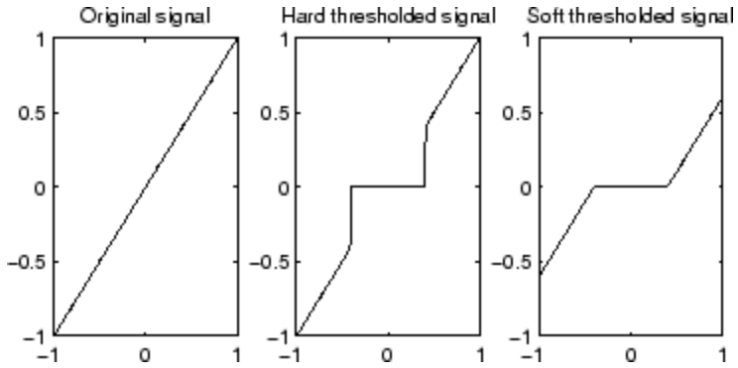


Figure 5-25 Soft and Hard threshold

The hard thresholding is defined as:

$$F = \begin{cases} Y \; if \; |Y| > T \\ 0 \; otherwise \end{cases}$$

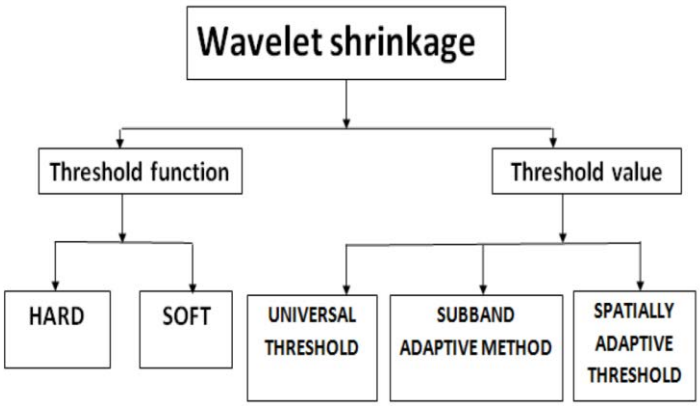
The soft thresholding is instead defined as:

$$F = sgn(Y) \cdot \max(0, |Y| \geq T)$$

Where T is the threshold value (T=0,5 in the image), Y is the input and F the output. The soft thresholding rule is chosen over hard thresholding, which is found to introduce artefacts in the images.

Three categories of threshold selection exist:

- Universal threshold: a threshold value is uniquely chosen for all wavelet coefficients
- Sub band adaptive thresholding: threshold value is selected differently for each detail sub band
- Spatially adaptive threshold: each detail wavelet coefficient has its own threshold value



5.5.3.1 *Universal Threshold*

The universal threshold can be defined as

$$T = \sigma\sqrt{2\log N}$$

N being the image size, σ being the noise variance. The universal threshold can give a better estimate of the image with the soft threshold. However, the estimated threshold value depends on the image size. With a particular 'σ', universal threshold yields larger threshold for big images and comparatively small threshold for smaller images, also it requires the prior knowledge about the noise distribution.

5.5.3.2 *Visu Shrink*

It follows the hard threshold rule. An estimate of the noise variance 'σ' is defined based on the median absolute deviation which is a robust estimator.

$$\hat{\sigma}^2 = \left[\frac{median(|X_{ij}|)}{0.675}\right]^2, X_{ij} \in HH1$$

$$T = \hat{\sigma}\sqrt{2\log N}$$

Visu Shrink does not deal with minimizing the mean squared error. Another disadvantage is that it cannot remove speckle noise. Yet, with additive gaussian noise assumption Visu Shrink exhibits better denoising performance than the universal threshold.

5.5.3.3 *SURE Shrink*

A threshold chooser based on Stein's Unbiased Risk Estimator (SURE) was proposed by Donoho and Johnston and is called as Sure Shrink. It is a combination of the universal threshold and the SURE threshold. It has the distinct advantage of offering an analytic unbiased estimator. The goal of Sure Shrink is to minimize the mean squared error of the estimate.

Sure Shrink suppresses noise by thresholding the empirical wavelet coefficients. Sureshrink is smoothness adaptive, which means that if the unknown function contains abrupt changes or boundaries in the image, the reconstructed image also has the same. The risk for a particular threshold value 't' can be estimated. The optimal threshold can be selected by minimizing the risks in 't'.

Threshold value for t_j for each resolution level j in the wavelet transform is used, which is referred to level dependent thresholding.

The Sure Shrink threshold t* is defined as follows:

$$t^* = \min(t, \hat{\sigma}\sqrt{2\log N})$$

Where, t denotes the value that minimizes Stein's Unbiased Risk Estimator, $\hat{\sigma}$ is the median absolute deviation and N the image size. Sure Shrink minimizes the mean squared error and also it is smoothness adaptive, which means that if any unknown function includes abrupt changes or boundaries in the image, the reconstructed image also has the same.

5.5.3.4 *Bayes Shrink*

Unlike universal threshold, Visu Shrink and Sureshrink, Bayes Shrink sets different thresholds for every subband. As noise is additive in nature so noisy image is additive sum of original image and noise and in terms of variance is:

$$\sigma_y^2 = \sigma_x^2 + \sigma_v^2$$

Where σ_y is the variance of noisy image, σ_x the variance of original image and σ_v the variance of noise.

Since, the diagonal coefficients of first level wavelet decomposition (HH1) contains significant amount of information about the noise components, the noise variance 'σ_v' is calculated using the robust estimator:

$$\sigma_v^2 = \left[\frac{median(|X_{ij}|)}{0.675}\right]^2, X_{ij} \in HH1$$

Variance of the corrupted image is estimated as:

$$\hat{\sigma}_y^2 = \frac{1}{J}\sum_{j=1}^{J} W_j^2$$

Where Wj are the wavelet coefficients in each scale 'j' and 'J' is the total number of wavelet coefficients. The threshold value using Bayes shrink is given by

$$T = \begin{cases} \dfrac{\hat{\sigma}_v^2}{\hat{\sigma}_x^2} \; if \; \hat{\sigma}_v^2 \leq \hat{\sigma}_x^2 \\ max\{|W_j|\} \; otherwise \end{cases}$$

Where:

$$\hat{\sigma}_x^2 = \max(\hat{\sigma}_y^2 \leq \hat{\sigma}_v^2, 0)$$

The estimation in equation holds good for images corrupted by Gaussian noise. Nevertheless, it is less sensitive to the noise around edges, but completely denoises the flat regions of the image. Modified Bayes shrink overcomes this issue. The threshold is given by:

$$T' = \beta \frac{\hat{\sigma}_y^2}{\hat{\sigma}_x}$$

Where

$$\beta = \sqrt{\frac{\log J}{2j}}$$

'J' is the total of coefficients of wavelet. 'j' is the wavelet decomposition level present in the sub-band coefficients under consideration. The modified Bayes shrink yields the best results for denoising and preserves edges better than the original Bayes shrink.

5.5.3.5 *Normalshrink*

Normal shrink an adaptive threshold estimation method based on the generalized Gaussian distribution (GGD) modeling of subband coefficients. The threshold is computed by

$$T' = \beta \frac{\hat{\sigma}_v^2}{\hat{\sigma}_y}$$

where σ_v and σ_y are the standard deviation of the noise and the subband data of noisy image respectively. β is the scale parameter, computed as

$$\beta = \sqrt{\frac{L_j}{M}}$$

L_j is the length of the subband at j-th level, M is the total number of decompositions, σ_v^2 is the estimated noise variance of HH1 subband and σ_y is the standard deviation of the image subband. This method is computationally more efficient and adaptive because the parameters required for estimating the threshold depend on subband data. Performance of normal shrink is similar to Bayes shrink. But Normal shrink preserves as well as removes noise better than Bayes shrink.

5.5.3.6 *Minmax Threshold*

The minimax principle was initially used in statistics to design estimators. Since the denoised signal can be assimilate to the estimator of the unknown regression function, the minimax estimator is the option that realizes the minimum, over a given set of functions, of the maximum mean square error[26]. The optimal threshold is derived from minimising the constant term in an upper bound of the risk involved in the estimation. Two oracles namely diagonal linear projection (DLP) and the diagonal linear shrinker (DLS) are used. DLP tells when to "keep" or "kill" each wavelet coefficient, whereas DLS states how much shrinking is applied to each wavelet coefficient.

Minimax threshold does not give good visual quality, but it has the advantage of giving predictive performance.

5.5.3.7 *Neigh Shrink Sure*

Neigh Shrink Sure is an improvement over Neigh Shrink, which has disadvantage of using a non optimal universal threshold value and the same neighbour window size in all wavelet sub bands

In simple Neigh Shrink, for each noisy wavelet coefficient to be shrinked, a square neighboring window centred at it. In sub band thresholding, the threshold and neighboring window size keep unchanged in all sub bands.

Neigh Shrink Sure, instead, can determine an optimal threshold and neighboring window size for every sub band by the SURE rule. They combine the unknown noiseless coefficients from sub bands into the corresponding 1-D vector. As using Stein's approach for almost any fixed estimator based on the data, the risk can be estimated in an unbiased way.

5.5.3.8 *Examples of Application*

For smooth images like 'Peppers', Visu Shrink, Sure shrink, Bayes shrink and wavelet based minimax threshold are visually appealing. On the other hand, for images with more details (Barbara), Visu Shrink, Bayes shrink and minimax threshold are not able to preserve edges. Sure shrink exhibits visually good results for images with more details.

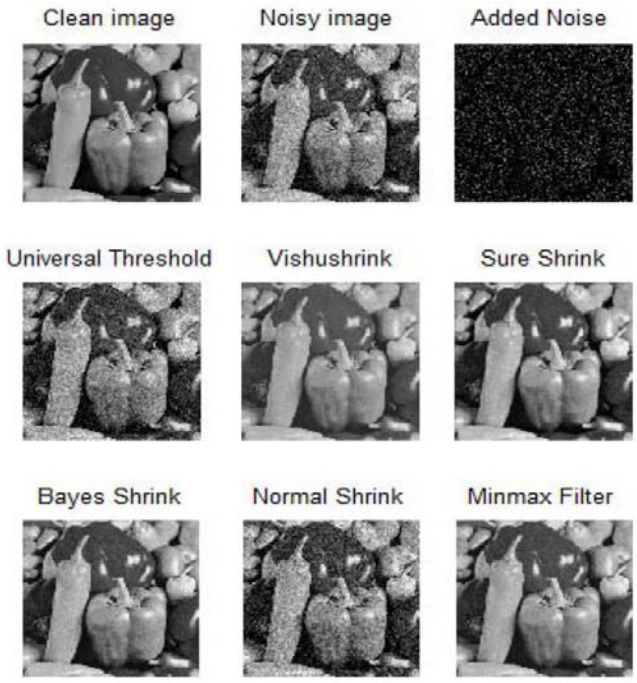


Figure 5-26 Example of filters application on a Smooth Image (Peppers)

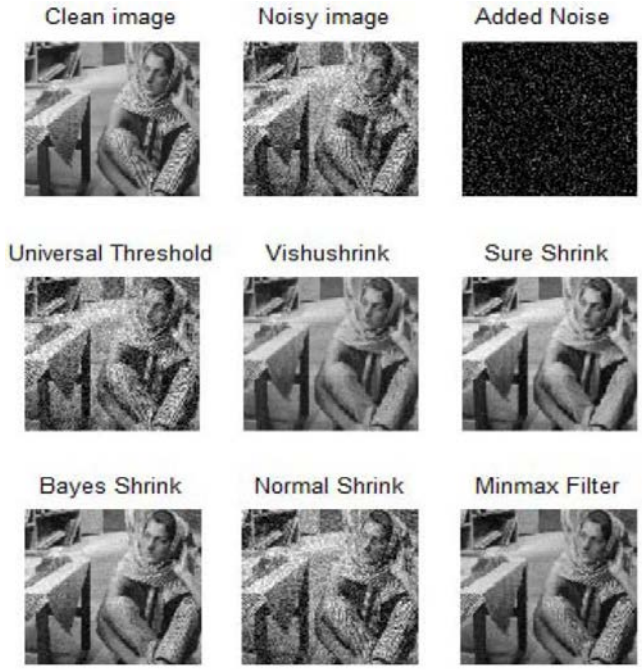


Figure 5-27 Example of application of filter on a more Detailed Image (Barbara)

Sure shrink performs well in terms of improving visual quality for both smooth and detailed images among the shrinkage methods.

Bayes shrink performs considerably better in improving visual quality. However, the images denoised by wavelet based denoising are prone to checkerboard artifacts due to the limited directional selectivity of wavelets. This effect is resolved with the use of highly directional representations to improve denoising performance.

5.5.4 Inverse DWT

After computing the 2D wavelet transform of an image and having altered the transform coefficients, the inverse transform has to be computed.

5.6 OTHERS USEFUL ALGORITHMS TO IMPROVE OR ANALYSE THE IMAGE

5.6.1 Removing Defective Pixels

Abnormal or defective pixels can be defined as being those significantly deviating from the average behaviour of their neighbours, whether due to manufacturing defects (in the most common case) (Dargaud, 2009 and Rodríguez, 2009), periodic incidence of radiation during sensor operation, as in astronomical imaging (Cresitello-Dittmar, 2001), or due to interference by some type of parasitic radiation in charge-coupled (CCD) device applications involving high-frequency energy (Li et al., 2006) depending on what it is intended to capture. Two basic types of anomalous pixels can be found in the

literature (Dargaud, 2009 and Brändström, 2009), according to their greyscale: dead pixels (always having a very low average level or having a very little sensitivity) and hot pixels (retaining a very high level instead, being common in astrophysical images where they can be confused with astronomical bright bodies).

Based on the work of Giron and Correa [32] a typical value of the pixels and a threshold has to be defined. Instead of using the median of all the pixels of the frame, we found that the mean produced more accurate results. Moreover, instead of calculating the mean on the whole image, it's better to divide the frame in square areas 50 pixel wide and perform the calculation on each one. A smaller box size increases only the computational time without obtaining much better results.

For every zone, a threshold is defined as three times the standard deviation of the temperature values. Then every defective pixel is marked and for each one a small area 8 pixels wide is considered. Then the mean of the new area is calculated (defective pixels excluded) and the value is substituted to the corrupted pixel.

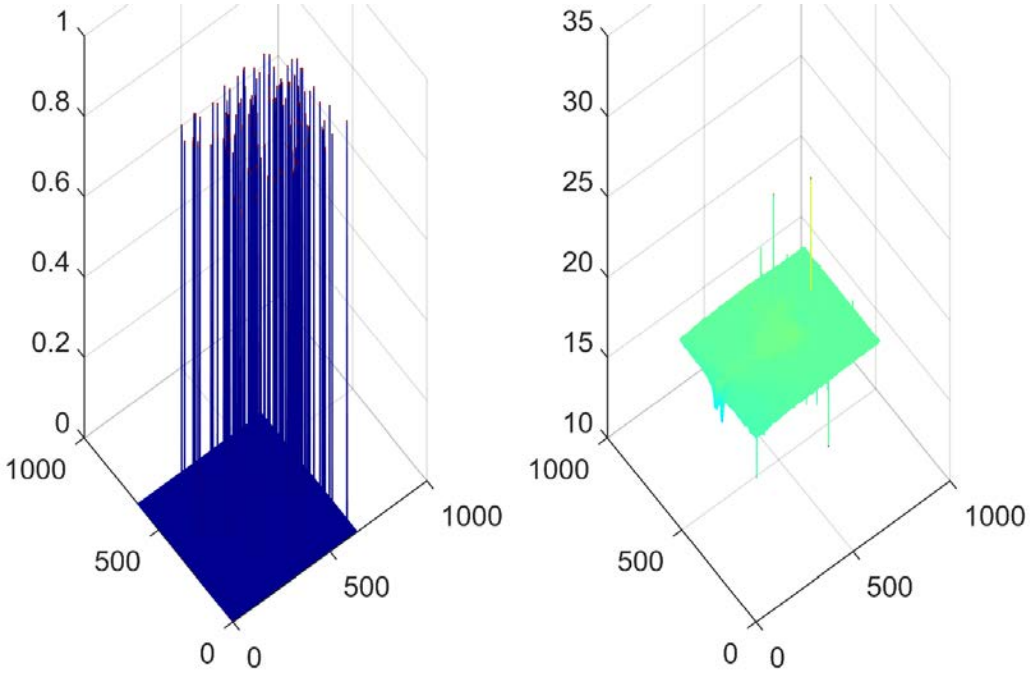


Figure 5-28 Defective pixel for a frame. On the left are highlighted the defective pixel and on the right the original image. The highest peaks can be easily seen. These peaks cause to have wrong limits for the image, which doesn't correspond to the physical temperature

The major use of this algorithm is to improve the visualization of the images. For some wavelet decomposition parameters it increase the quality of the results, avoiding always existing non physical peaks even after shrinkage.

5.6.2 Crack Tip Detection Algorithm

Manual crack tip detection is very difficult and time consuming. For the heat conductivity of the material a precise temperature peak is not so clear. An automatic method has to be introduced.

The main idea is to follow the maximum temperature in every frame (for each data of 100 frames the crack is considered stationary) and the perform the mean of the positions.

That is not sufficient to obtain an accurate position of the crack tip. Phenomena like motion of the specimen and friction on the crack surface create errors.

If we chose instead of the maximum position, the mean of the locations of the 99,9% percentile of the temperatures still we don't reach the desired accuracy. It's better to keep only the maximum and exclude some outliers.

The outliers are defined in the temperature, spatial and loading domain.

Temperature and size of the picture are normalized to one, then only peak in the 0.2 to 0.8 temperature range are keep. In spatial domain values outside the 0.01-0.99 for the horizontal dimension and 0.2-0.8 for the vertical dimension are ignored, this has to be done because on borders a peak could be find and even one pixel on the border can change much the crack tip position moving the mean position.

The most important points' filter for the points is to exclude the points when the temperature is decreasing (loading phase because the thermo-elastic effect causes cooling during extension).

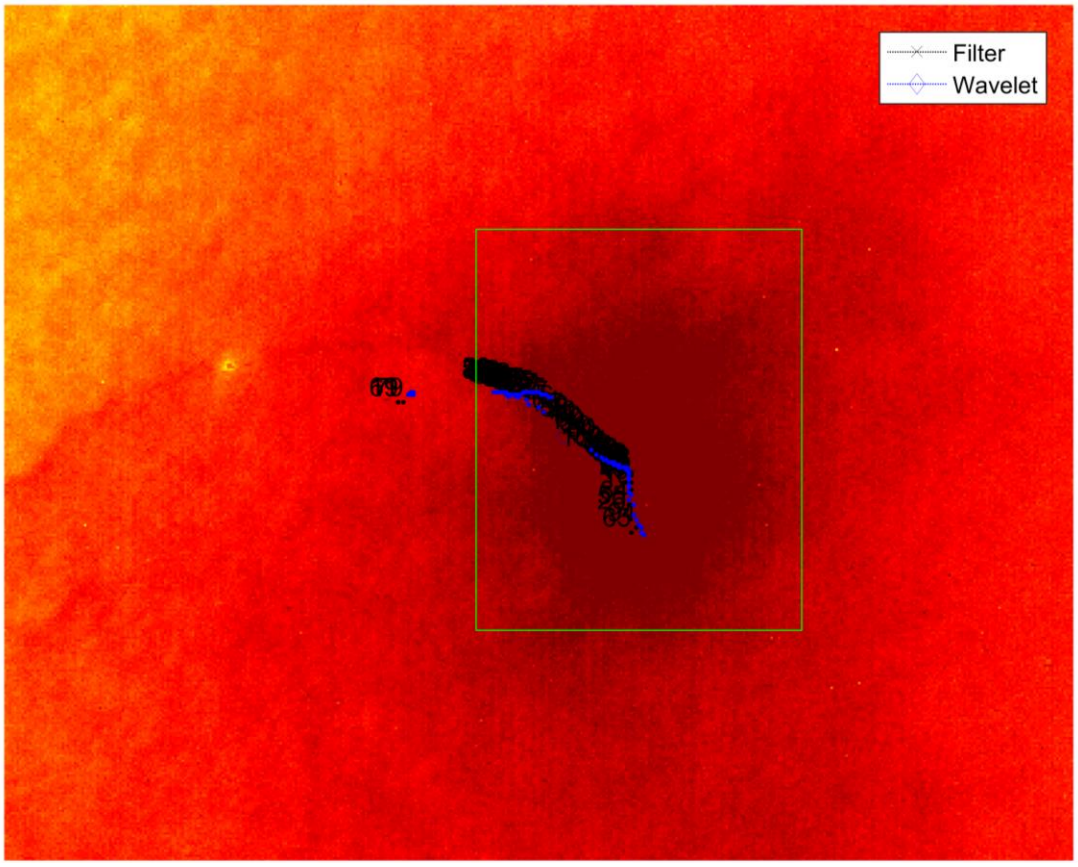


Figure 5-29 Crack area without excluding outliers. The blue small points represent the maximum point of every frame (the number of frame is written near the point)

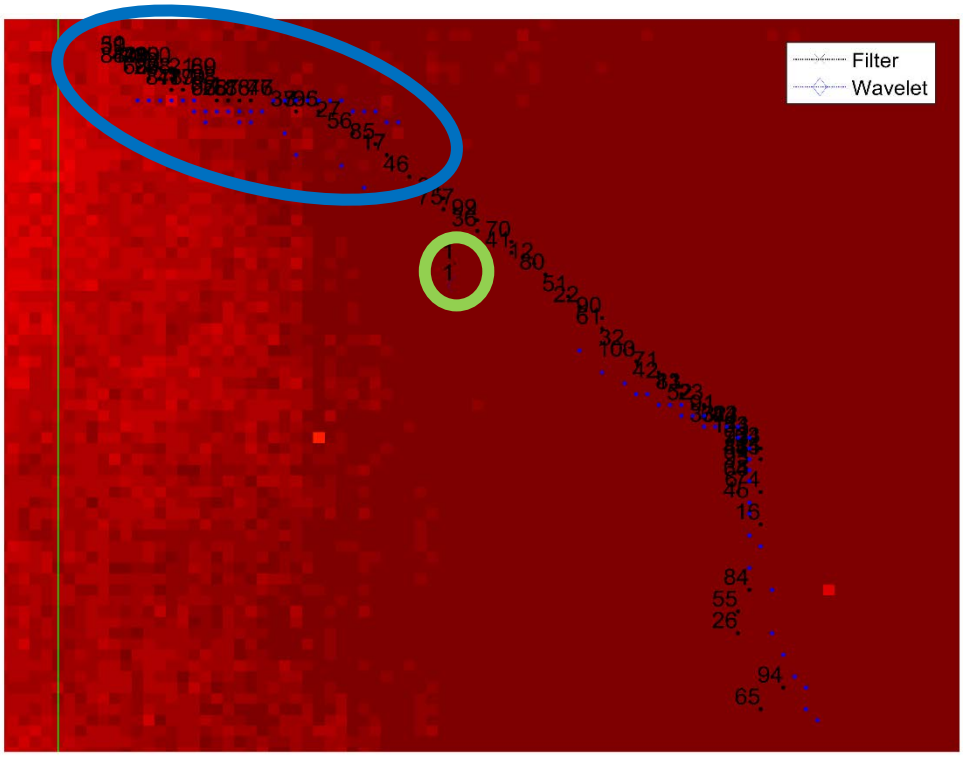


Figure 5-30 A zoom of the crack tip area can be seen in the lower figure. The green circle are the crack tip position (coincident both with gaussian and wavelet filtering). The blue circle contains the values in the unloading phase, this cause the mean of the position

The plot of maximum temperature versus frame number (time) is self-explanatory of the process adopted.

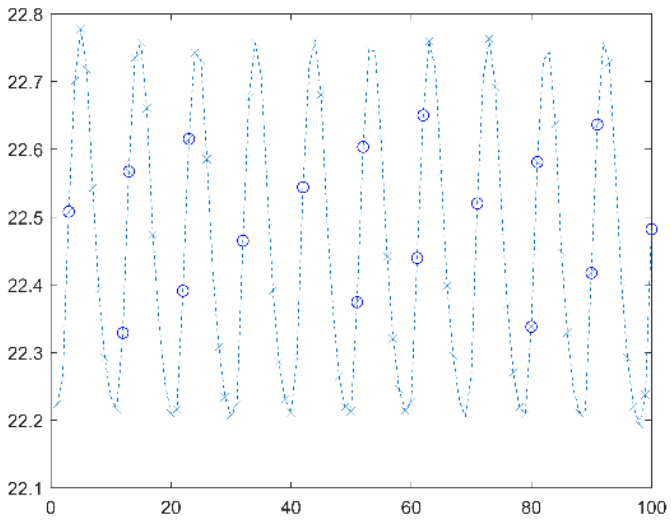


Figure 5-31 On the Y axis, the maximum temperature of the specimen for the frame. On the X axis the frame number for the data set. The point circled are the keeped ones

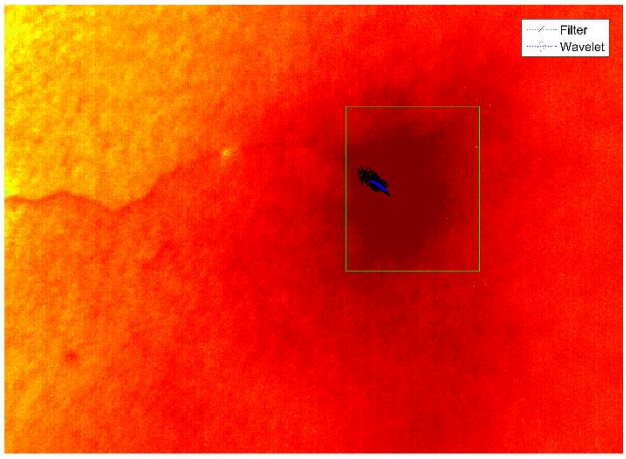


Figure 5-32 Crack area after having excluded outliers. The blue small points represents the maximum point of evry frame (the number of frame is written near the point)

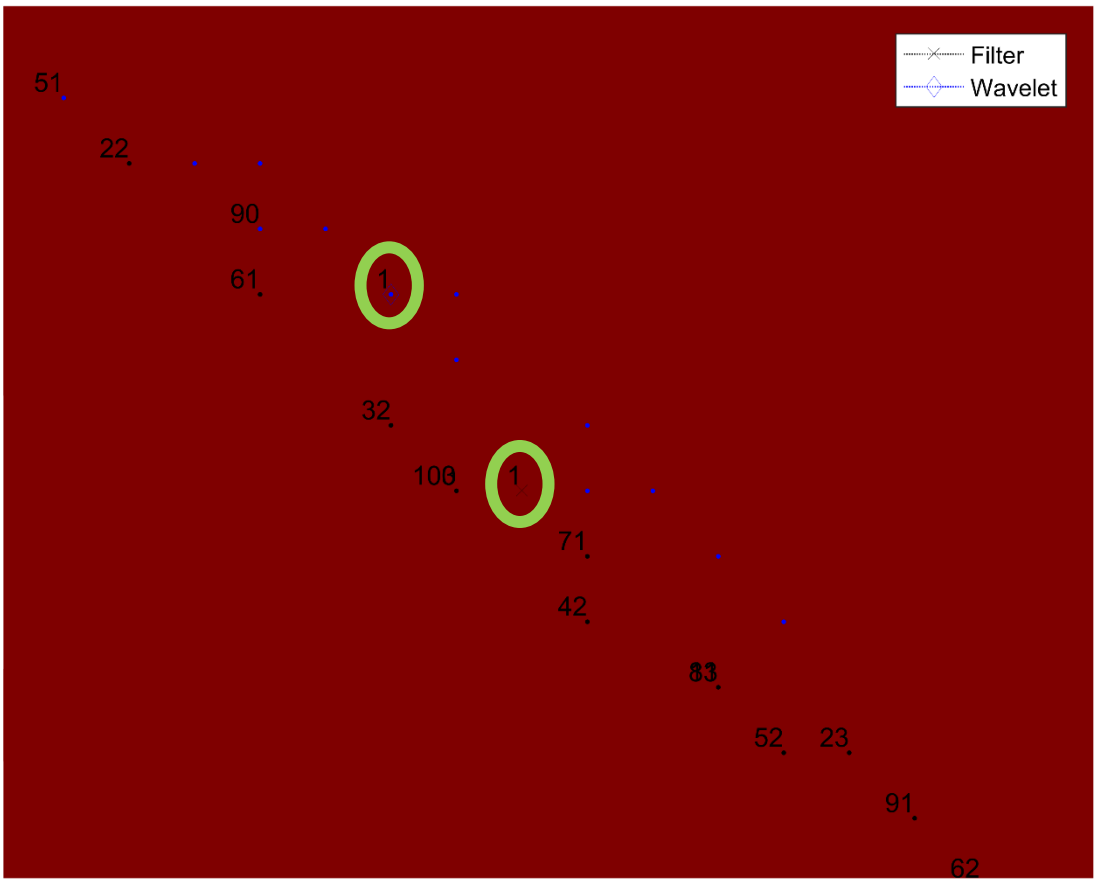


Figure 5-33 Crack tip and magnification of the iterested zone. The points with the two methods are coincident (3 pixel far). There is still a small scattering of the crack tip for evry point for the motion of the specimen and the noise that cannot be removed without losing information of the image.

The results are quite impressing and the crack can be followed almost perfectly. Only when the crack hasn't started yet it's difficult to determine the initiation point. Knowing the pixel size of the image then the crack length can be calculated. The most useful scope of this procedure is to perform the successive calculation only on a Region Of Interest that contains the plastic zone.

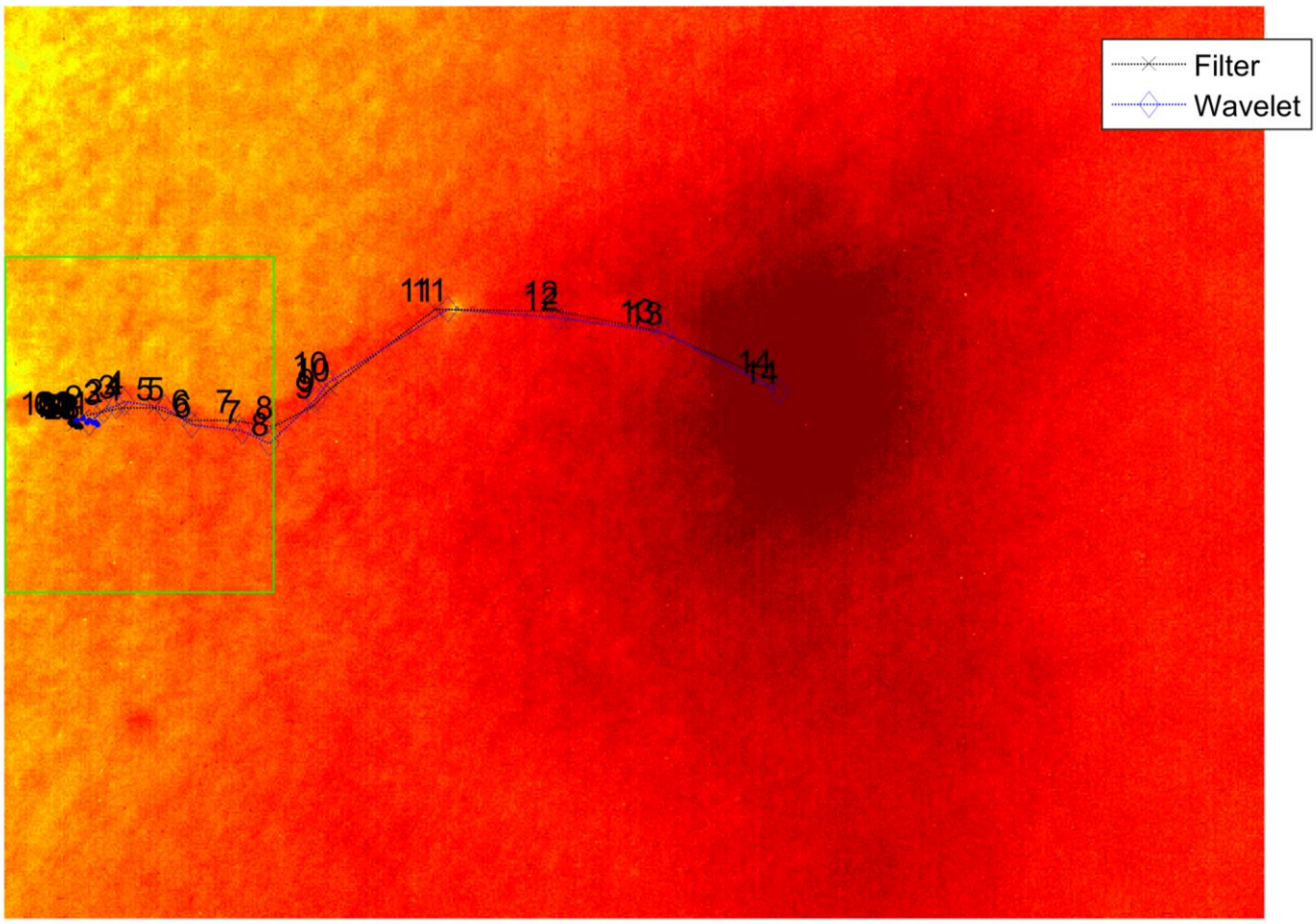


Figure 5-34 There results are equivalent for both approaches

5.6.3 Motion Compensation by cross correlation

An attempt has been made to compensate the motion of the specimen. Phase correlation is an approach to estimate the offset between two similar images. It is commonly used in image registration (the process of transforming different sets of data into one coordinate system) and relies on a frequency-domain representation of the data.

Image registration or image alignment algorithms can be classified into intensity-based and feature-based. One of the images is referred to as reference and the others are referred to as target images.

Applying the phase correlation method to a pair of images produces a third image which contains a single peak. The location of this peak corresponds to the relative translation between the images.

Unlike many spatial-domain algorithms, the phase correlation method is resilient to noise, occlusions, and other defects typical of medical or satellite images, all more complex images than temperature profile of cracks. Additionally, the phase correlation uses the fast Fourier transform to compute the cross-correlation between the two images, generally resulting in large performance gains.

The method has been applied only for translation, however it could also be extended to determine rotation and scaling differences between two images by first converting the images to log-polar coordinates. Due to properties of the Fourier transform, the rotation and scaling parameters can be determined in a manner invariant to translation.

Given two input images: f(x,y) as a reference and w(x,y) as a template and then calculating the DFT of both images.

$$F(u,v); \quad W(u,v)$$

Then the cross-power spectrum is calculated by taking the complex conjugate of the second result, multiplying the Fourier transforms together elementwise, and normalizing this product elementwise.

$$R = \frac{F \circ W^*}{|F \circ W^*|}$$

Where ∘ is the Hadamard product (takes two matrices of the same dimensions, and produces another matrix where each element ij is the product of elements ij of the original two matrices) and this impose that the images has to be of the same size or one has to be padded (adding new pixels around the edges of the image).

The normalized cross-correlation r is obtained by applying the inverse Fourier transform.

$$r = IFT(R)$$

The offset between the images is determined by finding the position of the peak in r.

Following the work of Lewis [33], although it is well known that cross correlation can be efficiently implemented in the transform domain, the normalized form of cross correlation preferred for feature matching applications does not have a simple frequency domain expression. For this reason normalized cross correlation has been computed in the spatial domain.

The window containing the feature w is positioned at s, t. the correlation coefficient is defined as:

$$\gamma(s,t) = \frac{\sum_{x,y}\left[f(x,y) - \bar{f}_{s,t}\right]\left[w(x-s,y-t) - \overline{w}\ \right]}{\left\{\sum_{x,y}\left[f(x,y) - \bar{f}_{s,t}\right]^2 \sum_{x,y}\left[w(x-s,y-t) - \overline{w}\ \right]^2\right\}^{0.5}}$$

Where

$f(x,y)$ is the image

$\overline{w}$ is the mean of the template

$\bar{f}_{s,t}$ is the mean of $f(x,y)$ in the region under the template

The value $\gamma(s,t)$ ranges from -1 to 1. A high value for $|\gamma(s,t)|$ indicates a good match between the template and the image, when the template is centred at coordinates (s,t).

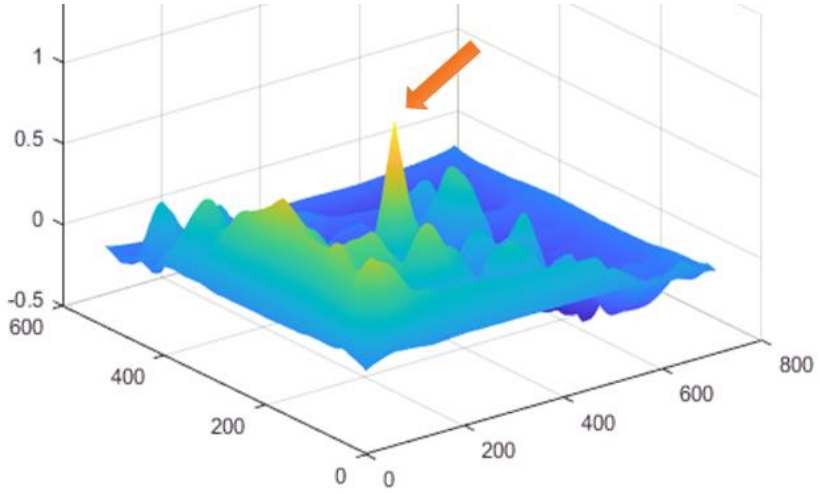


Figure 5-35 Example of plot of the Correlation Coefficient γ(s,t). The high peak corresponds to the coordinates (s,t) offset of the two images.

Normalized cross-correlation (NCC) is not invariant with respect to imaging scale, rotation, and perspective distortions but it's as a default choice in many applications where a fast algorithm is therefore of interest.

The results are not satisfactory yet. We have to use a Sobel operator to highlight the edges of the notch and the remove all the edges found that aren't in contact with the image borders to eliminate noise in the image.

- Frame 1: is the template and must be smaller than the reference image
- Frame 2: is the reference image (only a small part of the original image)

After the offset is calculated then the Frame 2 is moved to be coincident with the Frame 1 so after a loop cycle, all the recording should be still with respect of the first frame.

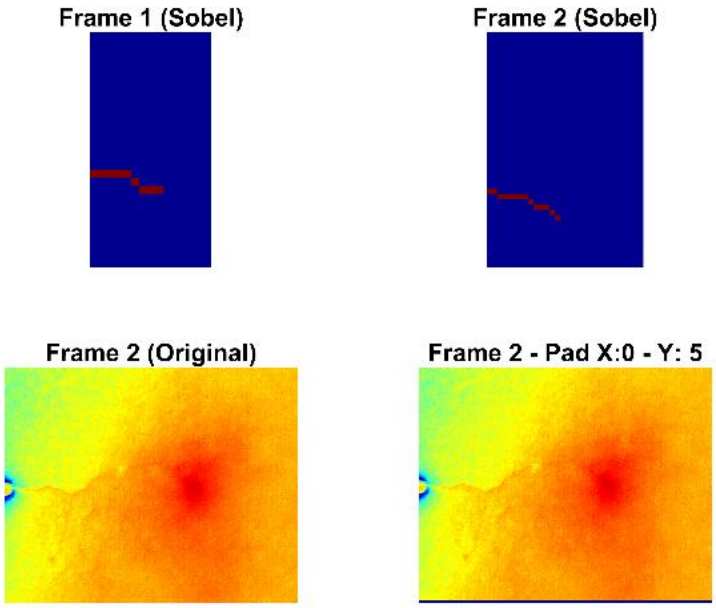


Figure 5-36 On the upper side the template (left) and cropped reference image (right) are present. Below the original 2nd Frame and on the right the offset applied let it be coincident with the previous frame

Even with the edge detection algorithm the results aren't yet precise but the improvement is huge. A filter different than Sobel could tried to be used. Also a round approximation problem could be the cause that bias the y offset always up.

The expected behaviour would be a sinusoidal wave with mean zero and not always shifted up.

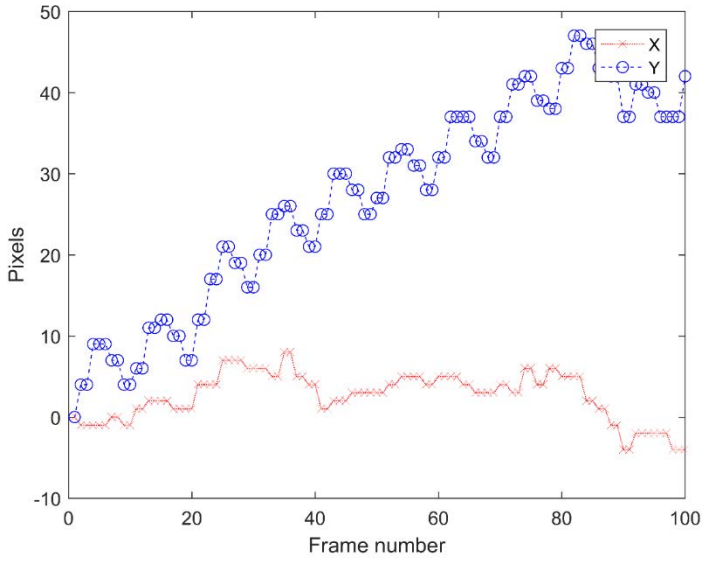


Figure 5-37 Offset for the Y (blue) and X (red) directions on a set of data of 100 frames

6 COMPARISON BETWEEN CASES

Two main denoising approaches has been followed: one using filtering in the frequency domain and the other using wavelets and then a shrinkage to eliminate noise. The last method produces better results in terms of smoothness of the temperature and the heat generation distribution. However even if the second method produce results more graphically easy to understand, the final results are comparable in both methods as it can be seen in the small difference between crosses and circles in the graph below.

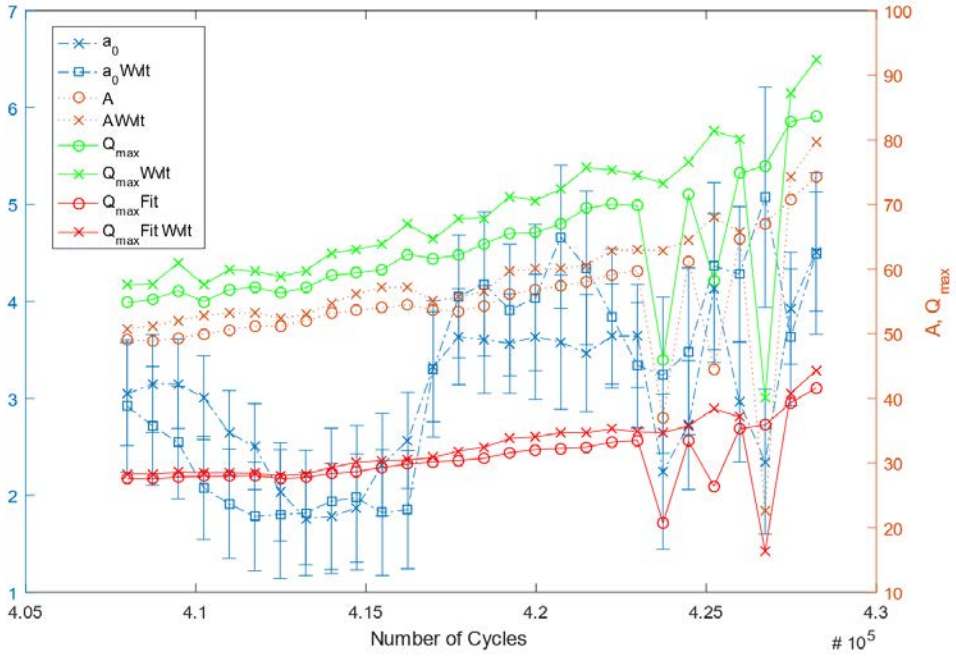


Figure 6-1 Comparison for different parameters between conventional filtering and wavelet filtering

The main objective is the calculation of the Heat generated in the process, the frames contains intrinsic noise that causes some numerical instabilities when the 2D heat equation is solved. This problem is accentuated when the pixel size and the time size decreases.

6.1 CONVENTIONAL FILTERING

The removal of the defective pixels doesn't remove the noise. First a Wiener Filter in combination with Median Filter is applied to reduce some of the spikes then a Gaussian filter is applied. Values of sigma near 8 had to be chosen otherwise the heat generation function contains too much noise.

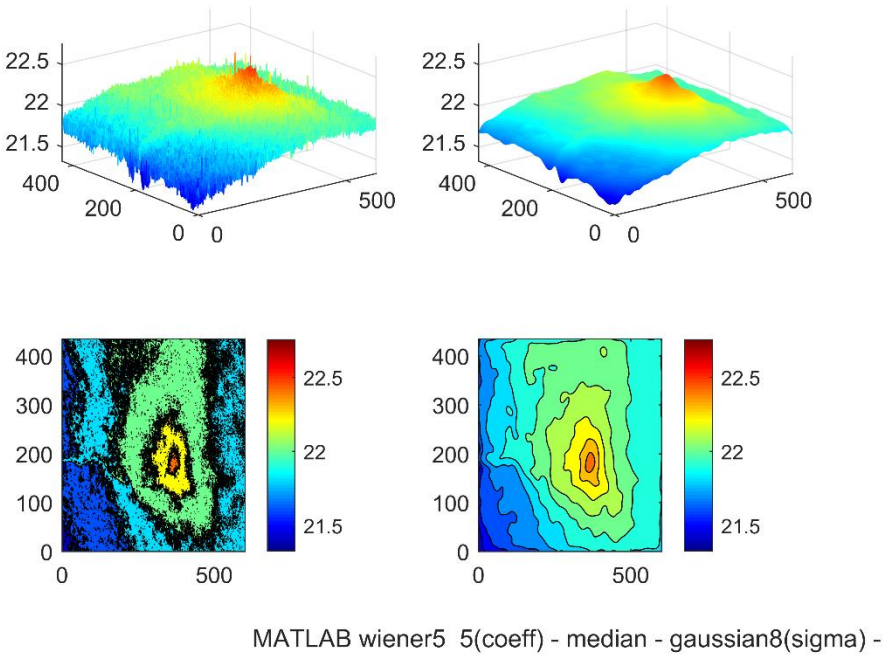


Figure 6-2 On the left the original image (after removing defective pixel) and on the right the Filtered Image with its contour plot on bottom

If we zoom on the image, it can be seen that isn't completely smooth but with a higher sigma the information on the crack tip are lost.

Also, an anisotropic Gaussian filter in three dimension has been tested, for x and y dimensions sigma has been set to 8 as before but in the third dimension a much smaller value has been used otherwise the periodicity of the temperature will be lost. The heat source function appear smother but some effects at the border of the image could appear. Calibrating all the parameters to avoid this border effect of the heat without losing the periodicity of the temperature is difficult. Even if this problem should not interfere with the process zone it has been preferred to use the Gaussian filter only on x and y dimensions.

Also from the FFT logarithmic spectrum the improvements from the original image can be seen easily, all the high frequency values have been removed.

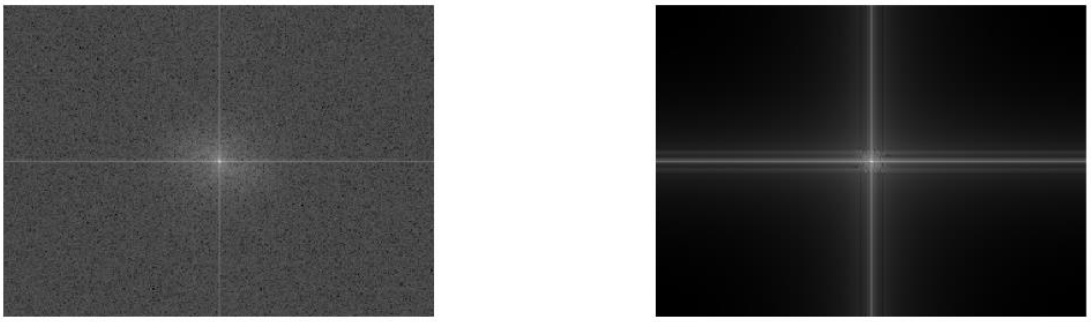


Figure 6-3 On the left spetrsum of the original image. On the right spectrum of the filtered frame, high-frequency components have been removed. As the previous image it has been applied a Median and Wiener filter and finally a Gaussian with sigma equal to 8

The final results for the temperature represents a clear improvement.

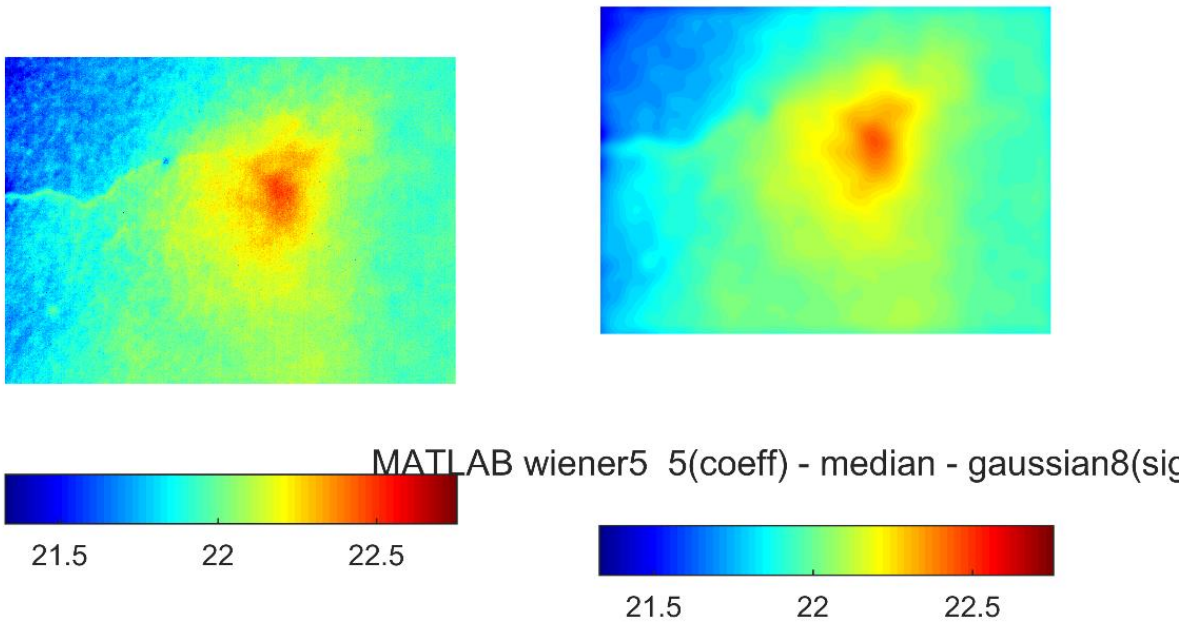


Figure 6-4 Corespective images of the previous spectra

The Heat generation function, as said before, still contains some noise but the overall results are very good. The peak correspondent to the plastic zone can be identified and also the crack surface. Some of the noise can also derive from an inhomogeneous material and it's difficult to eliminate that, we have to recur to wavelets.

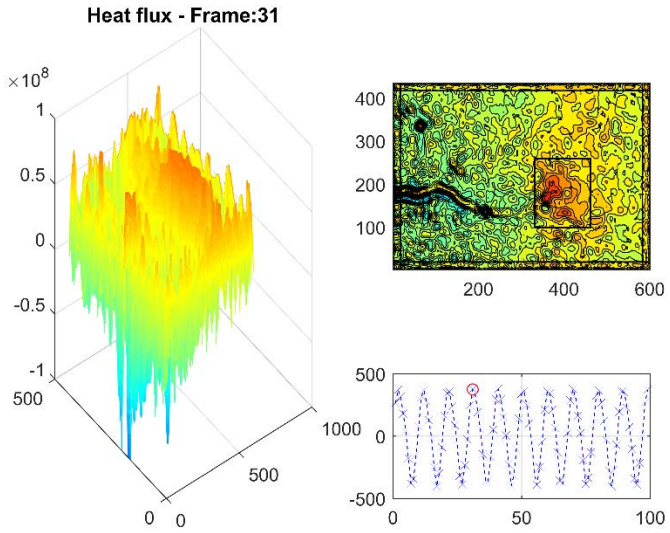


Figure 6-5 Heat source function calculated for an image frame. Median and Wiener filter has been applied to the image, finally a Gaussian filter is applied on the frame. On bottom right the integration of the function over the region of interest versus number of frame is plotted

When using wavelets it's recommended to crop the image and exclude the notch area. Otherwise it will cause some problems in the heat generation function. The integration on the ROI (region of interest) is still the same but the graphical interpretation is more difficult. At the start of the crack growth the results values of the integrated heat function are biased if the cropping isn't used.

The Universal Treshold has been applied to all the frames. Also the other methods have been tried but they only increase the computational cost without improving the quality more than the Universal Treshold. The other shrinkage methods seems to be more useful for real complex images.

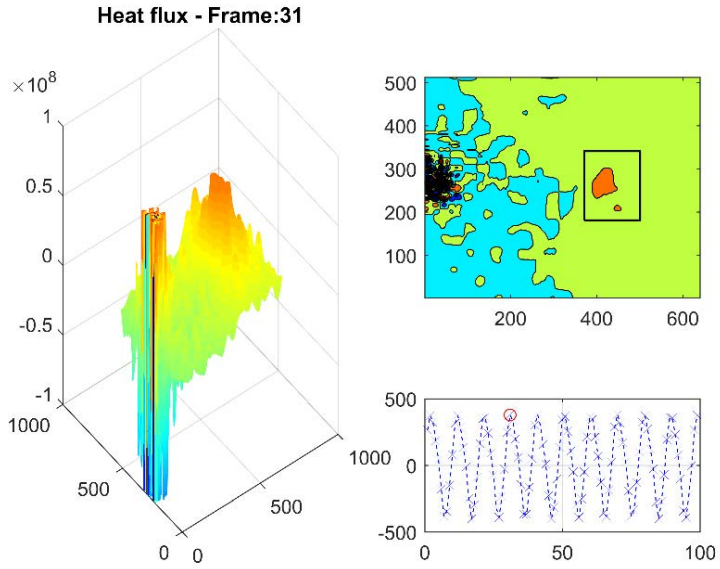


Figure 6-6 No-cropped frame, some problems in the notch area can be seen using wavleets. Level of decomposition is 5 and the Wavelet mother function is db10

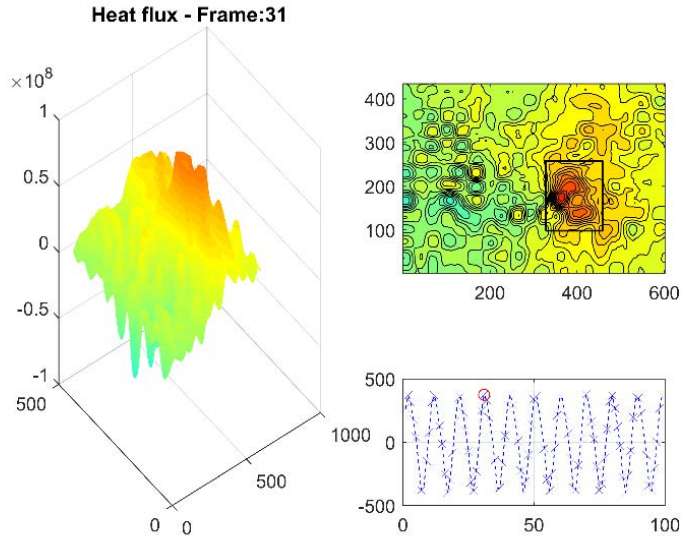


Figure 6-7 Cropped frame. Level of decomposition is 5 and the Wavelet mother function is db10

A level decomposition of 4 (less smoothing) or 5 is recommended, two Wavelet familes has proven to be the best for this type of application: Symlets and Daubechies. Symlets (which derive from Daubechies) offer in general a more smooth result.

Vandone in [22] experimented the use of sym8 mother function with six levels of decomposition. For thresholding he used the MATLAB's command wdencmp with this parameters:

$$threshold = [0.5\ 1\ 1\ 1\ 1\ 0]$$

For each detail level only the wavelet coefficients higher than a value equal to the maximum coefficient multiplied for the threshold value are considered. In this particular case, the coefficients from the second to the forth level are neglected, all the coefficients of the fifth level are considered, and just the highest coefficients of the first detail level are kept.

Because of the difficulties of fine tuning a threshold for very level, the universal Threshold has been applied. The results are very similar. A comparison of different wavelets of the Symlet family but with different surname will follow.

6.1.1 Symlet

For a sym6 the corresponding FFT spectrum is the following. Compared to Coventional Filtering is even more “clean” from high frequency components.

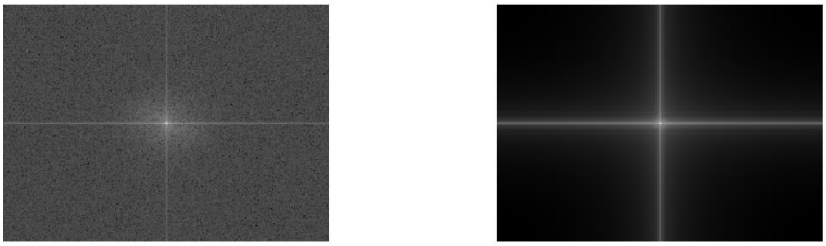


Figure 6-8 On the left FFT Spectrum of the original frame. On the right the spectrum of the same frame after a wavelet decomposition of level 5 with sym6 and the Universal Threshold applied

Comparing the images has no reason because they are very similar to human eye. It has more sense to compare the heat generation function directly. From the following graph it emerges that the functions are clearly smoother. Still some unnecessary peaks are present but this doesn’t prejudice the results of the integration.

Urealistic sharp peaks can be seen using sym6 with level of decomposition of 5.

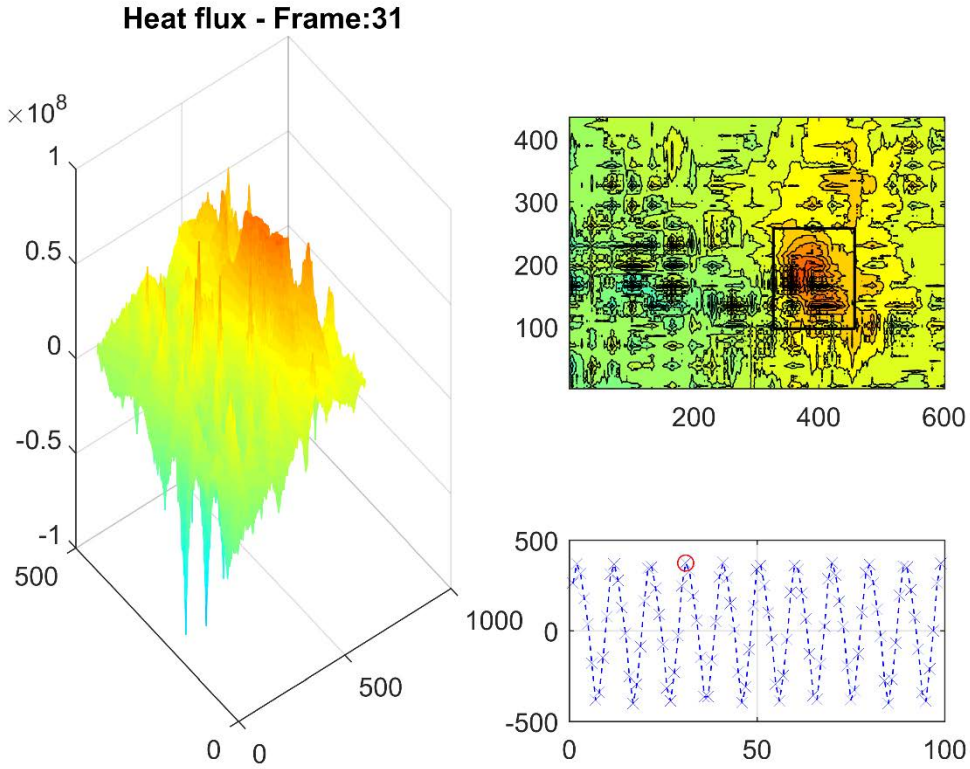


Figure 6-9 Heat Generation Function for a specific frame of the data set. Some unphysical peaks can still be observed far from the crack tip area (level 5 – sym6)

Keeping always the same level of decomposition of 5, it's preferred to use the sym14 to have a more clear view of the heat generation process. The FFT spectrum of the frame remain almost the same, but a small improvement in filtering permits to have much better results in the energy domain.

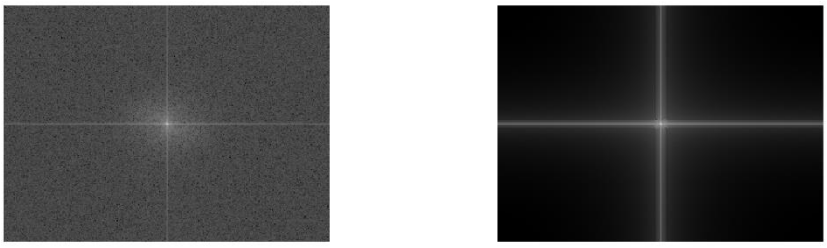


Figure 6-10 On the left FFT Spectrum of the original frame. On the right the spectrum of the same frame after a wavelet decomposition of level 5 with sym14 and the Universal Threshold applied

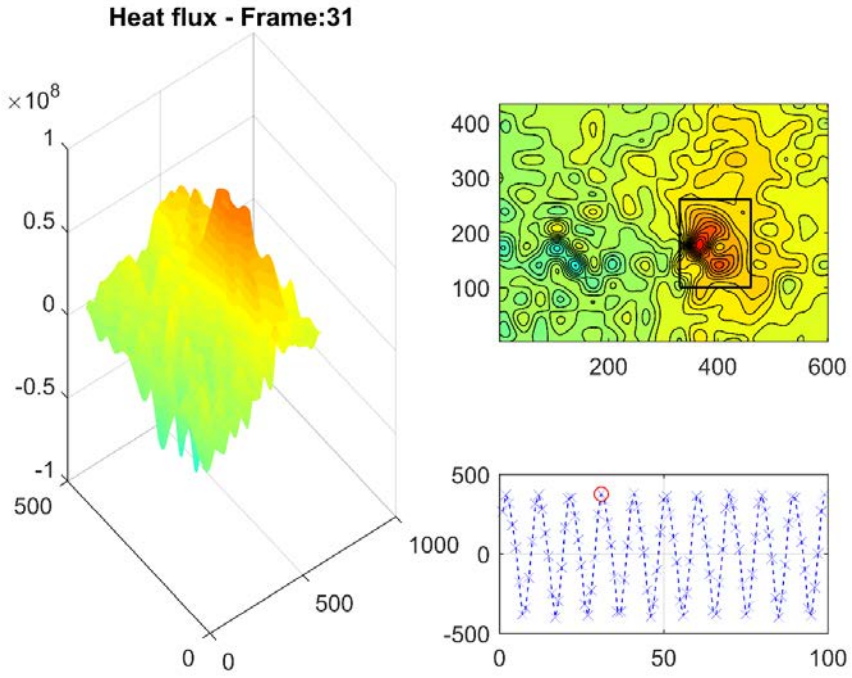


Figure 6-11 Heat source function calculated from a temperature field filtered through wavelet decomposition level of 5 and sym14

If the level of decomposition is lowered to 4, too many details start to appear.

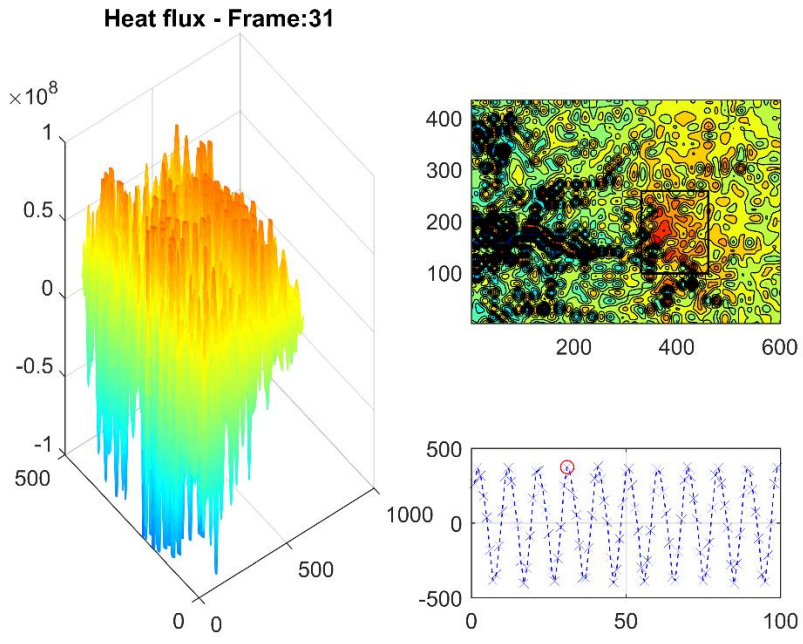


Figure 6-12 Results for the heat source function using a wavelet decomposition level of 4 and sym14

6.1.2 Daubechies

This wavelet family permit to keep more discontinuity compared to Symlets but this causes also the noise to increase. The preferred level of decomposition is still 5.

A good compromise between filtering and information preservation is db10.

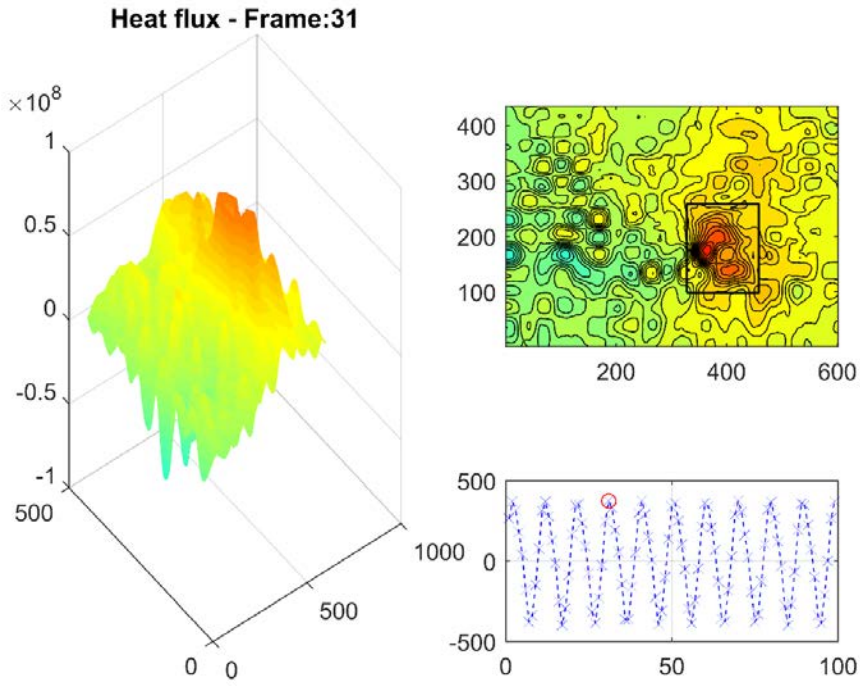


Figure 6-13 Results for the heat source function using a wavelet decomposition level of 5 and db10

Similar shape results are obtained with db8 but the values from integration differ, even if not by so much.

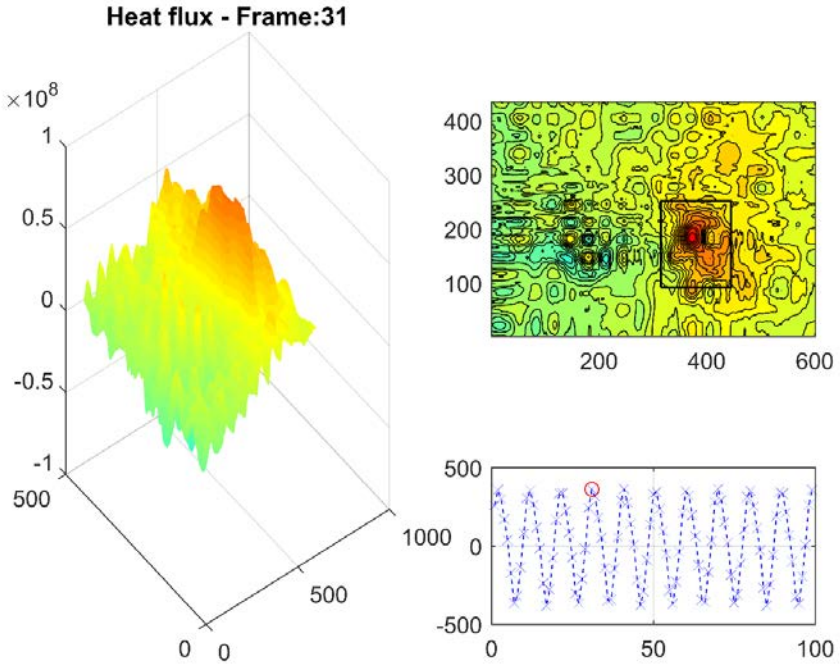


Figure 6-14 Results for the heat source function using a wavelet decomposition level of 5 and db8

Instead the wavelet surname db6 more unphysical peaks compared to the previous. The heat generation function integration produce still the same results as db8.

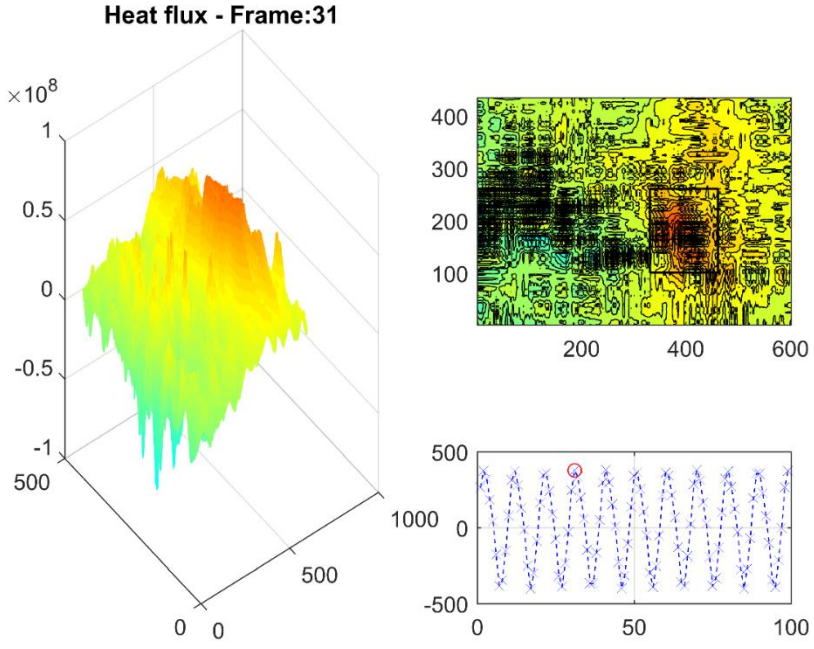


Figure 6-15 Results for the heat source function using a wavelet decomposition level of 5 and db6

7 NUMERICAL MODELS

A two-dimensional model of the specimen has been created for the calculation of the first intensity factor and strain energy density valued at different crack lengths. The numerical analysis has been performed assuming a linear elastic material behaviour even for the final phase of the crack growth and using crack tip singular elements. For the small thickness of the plates, plane stress conditions have been assumed.

A second model to characterize the heating of the process zone has been created to be then compared with the experimental results.

7.1 STRESS INTENSITY FACTOR CALCULATION

A static analysis with a vertical displacement applied on the pin hole (displacement control) has been performed. The model is symmetrical in correspondence of the crack surface and the material is linear elastic. On the crack tip, singular elements have been used.

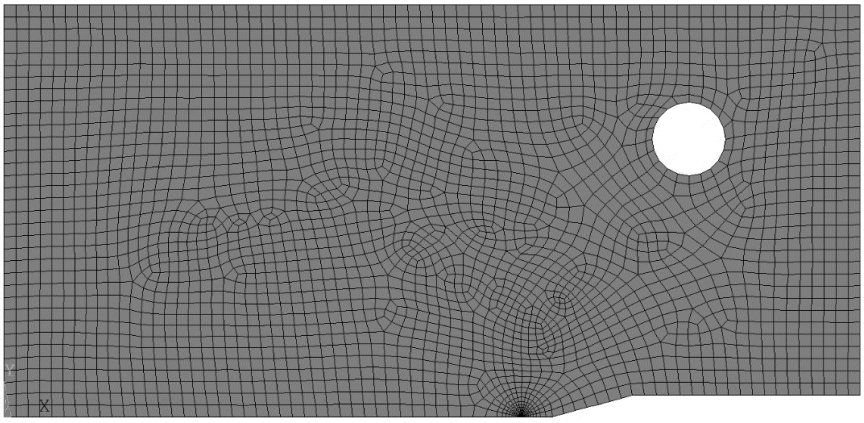


Figure 7-1 Mesh for the structural analysis

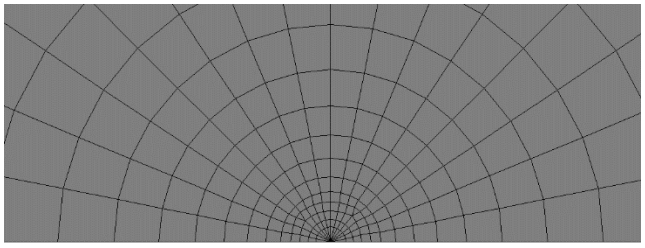


Figure 7-2 Details of the crack tip

The same procedure has been used for the SEN(T) specimens. Plane stress conditions have been applied.

7.1.1 Crack Tip Singular Elements

More accurate results, whiteout having to improve too much the mesh, could be obtained by capturing the crack tip singular stress field [34].

In elastic materials, the crack tip stresses are singular as $1/\sqrt{r}$ and this singularity can be built into the finite element calculation from the start. A number of methods to produce singular crack tip stresses exist, some of which require special elements and some of which can be used with standard elements. We will focus on 2D quarter-point elements that can be implemented using standard elements.

In the finite element method the displacement field and the coordinates are interpolated using shape functions, $N_i(\xi,\eta)$ and nodal coordinates (x_i,y_i).

Just moving the coordinates of the mid-side nodes to be L/4, is necessary to create a $1/\sqrt{r}$ singularity.

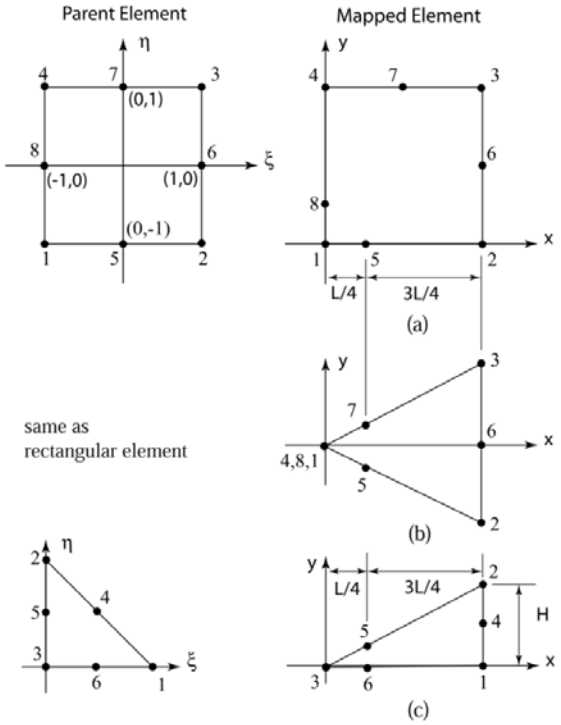


Figure 7-3 Parent elements in (ξ,η) plane and their mapped quarter-point elements. (a) 8 node rectangular element, (b) Triangular element formed by collapsing nodes 4,8,1 to a single point (c) Natural triangular element. (Recommended element for linear problems)

7.1.1.1 *8 node rectangular and degenerated triangular element*

The rectangular element has the drawback that it does not allow the crack to be surrounded by singular elements and hence to accurately capture the θ variation of stress, which might be needed for crack direction calculations. In addition, except along the element edges, the 1/ √r singularity exists only in a small region near the crack tip . The collapsed element has the drawback that the meshing will be somewhat more difficult to implement and that unless the element edges are straight the 1/√r does

not exist on straight lines coming from the crack tip and the accuracy in computing Stress Intensity Factors will be degraded.

7.1.1.2 *6 node triangular element*

A better choice might be the natural triangle quarter point element. Not only can the crack tip be easily surrounded by elements, the meshing is simple and the element gives the 1/√r on all lines coming from the crack tip.

The 6 triangular element is realized trough the Ansys's command KSCON. The KCTIP options has been set to 1 to skew midside nodes of the first row of elements to the 1/4 point for crack tip singularity and the ratio of the second row of elements compared to element size of the first row has been set to 0.5. 7 elements in the circumferential direction has been used.

7.1.1.3 *Stress Intensity Factor calculation*

The stress intensity factors at a crack for a linear elastic fracture mechanics analysis may be computed (using the KCALC command). The analysis uses a fit of the nodal displacements in the vicinity of the crack. The actual displacements at and near a crack for linear elastic materials are:

$$u = \frac{K_I}{4G}\sqrt{\frac{r}{2\pi}}\left((2k-1)\,cos\left(\frac{\theta}{2}\right) - cos\left(\frac{3\theta}{2}\right)\right) - \frac{K_{II}}{4G}\sqrt{\frac{r}{2\pi}}\left((2k+3)\,sin\left(\frac{\theta}{2}\right) - sin\left(\frac{3\theta}{2}\right)\right) + o(r)$$

$$v = \frac{K_I}{4G}\sqrt{\frac{r}{2\pi}}\left((2k-1)\,sin\left(\frac{\theta}{2}\right) - sin\left(\frac{3\theta}{2}\right)\right) - \frac{K_{II}}{4G}\sqrt{\frac{r}{2\pi}}\left((2k+3)\,cos\left(\frac{\theta}{2}\right) - cos\left(\frac{3\theta}{2}\right)\right) + o(r)$$

$$w = \frac{2K_{III}}{G}\sqrt{\frac{r}{2\pi}}sin\left(\frac{\theta}{2}\right) + o(r)$$

Where u,v,w are the displacement in the local cartesian coordinate system centred on the crack tip and r, θ in a cylindrical system. If the plane strain condition is valid k=3-4v, otherwise if the plane stress condition is valid $k = \frac{3-\nu}{1+\nu}$. o(r) represents terms of order r or higher.

If $\theta = 180°$, the higher order terms are neglected and the model is symmetric about the crack plane then we have:

$$K_1 = \sqrt{2\pi}\frac{2G}{1+k}\frac{|v|}{\sqrt{r}}$$

The final factor $\frac{|v|}{\sqrt{r}}$ is evauated based on the nodal displacements and locations.

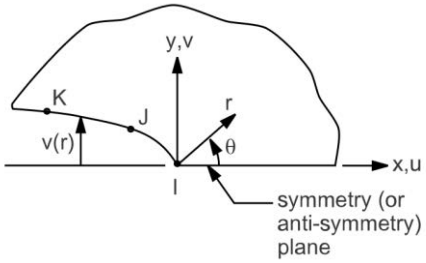


Figura 7-1 Nodes used for approximate displacement

7.2 NUMERICAL SIMULATION OF HEAT GENERATION

Since conduction plays a role in the heat transfer problem, the temperature field is affected by the applied loading rate. Applying the same loads, if the dissipated energy calculated numerically equals the real one, the temperature change should be the same. Moreover, in the future, a local stored energy value could be calculated as in [3]. A FEA model has been implemented both in ANSYS and ABAQUS software [35].

In ANSYS the material model has first been defined as a multilinear kinematic hardening to account for plastic work dissipation and large deformation has been turned on to be possible to solve the model (otherwise displacement would go to infinite).

In Abaqus the material model chosen is the bilinear kinematic, to implement a multilinear model, as in Ansys, a UMAT subroutine has to be written and then compiled by a Fortran complier.

Each of them has got some limitations, in fact ANSYS can't provide an instantaneous value for the Plastic Dissipated Work and so the generated heat is cumulated. Only the first cycles can approximate the thermal dissipation of the specimen. Also Abaqus present some problems because only a multilinear isotropic model can be defined instead of a kinematic one (that is the indicated for cyclic loads). The correspondent heat flux generated isn't the true value.

A comparison of the two software is made to create the two model as a future reference to improve.

The mesh doesn't need to be fine for having good results in the thermal behaviour. A coarse mesh is also needed to not increase excessively the computational time.

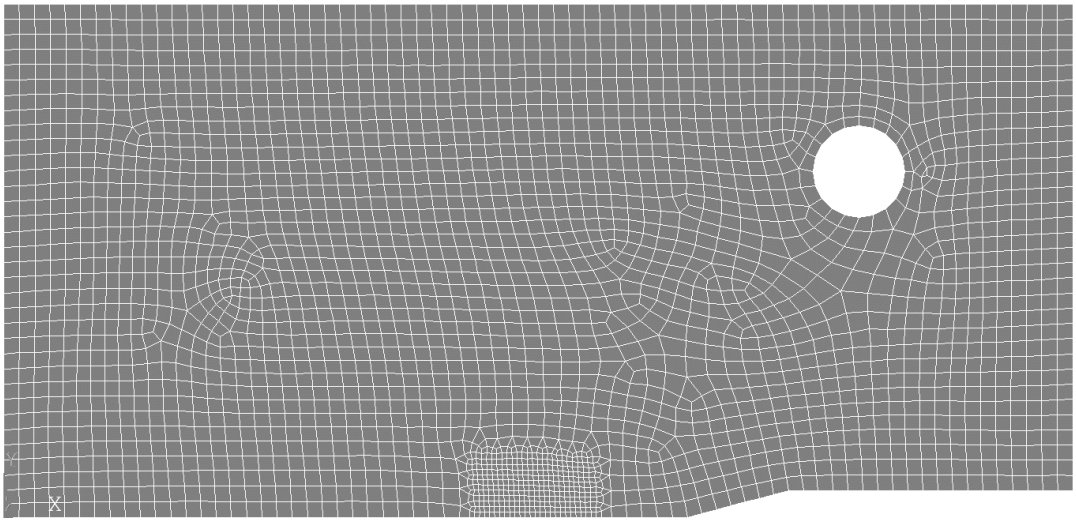


Figure 7-4 Mesh for transient thermic analysis

For the fact that plastic dissipation cannot be included without problems, a simulation with only the thermos-elastic effect present has been executed to see how much of the heat generated can be attributed to this phenomenon.

7.2.1 Taylor Quinney model for plastic heat generation

The thermal and mechanical fields are coupled through the plastic work heat fraction β.

Over many years the Taylor e Quinney coefficient was generally assumed as a constant independent of plastic deformation and strain rate and its value was accepted as between 0.85-0.95 for most metals.

In Ansys is defined trough the material propriety QRATE, in Abaqus from the GUI the correspondent name is Inelastic Heat Generation. In Abaqus, through the use of the UMAT routine, the inelastic heat generation can be written as a function of true deformation.

7.2.2 Element Type

Three ANSYS elements support the thermoplastic effect, "which manifests itself as an increase in temperature during plastic deformation due to the conversion of some of the plastic work into heat" (Theory Reference, 11. Coupling, 11.3. Thermoplasticity) in the following elements:

- PLANE223 - 2-D 8-Node Coupled-Field Solid
- SOLID226 - 3-D 20-Node Coupled-Field Solid
- SOLID227 - 3-D 10-Node Coupled-Field Solid

Setting KEYOPT(1) to 11 activates displacement and temperature degrees of freedom for these elements. During direct coupled analysis, structural-thermal coupling can have Strong (matrix) coupling which produces an unsymmetric matrix, or Weak (load vector) coupling which produces a symmetric matrix and requires at least two iterations per substep. Coupling choice is set with KEYOPT(2). As quoted below, Weak coupling is recommended in static and transient analysis. The Coupled-Field

Analysis Guide, 2. Direct Coupled-Field Analysis mentions Strong coupling if contact elements are used. Note the following comments:

> *"Because of the possible extreme nonlinear behaviour of weakly coupled field elements, you may need to use the predictor and line search options to achieve convergence."*

"To speed up convergence in a coupled-field transient analysis, you can disable the time integration effects for any DOFs that are not a concern. For example, if structural inertial and damping effects can be ignored in a thermal-structural transient analysis, you can issue TIMINT,OFF,STRUC to turn off the time integration effects for the structural degrees of freedom."

Turning on structural time integration doesn't permit to reach convergence even in the first iterations with a sinusoidal load.

In the Coupled-Field Analysis Guide, 2. Direct Coupled-Field Analysis, 2.6. Structural-Thermal Analysis it states:

> *"In a structural-thermal analysis with structural nonlinearities using elements PLANE223, SOLID226, and SOLID227, it is recommended that you use weak (load vector) coupling between the structural and thermal degrees of freedom (KEYOPT(2) = 1) and suppress the thermo-elastic damping in a transient analysis (KEYOPT(9) = 1). When using the SOLID226 element, you should also select the uniform reduced integration option (KEYOPT(6) = 1). These options will be automatically set if ETCONTROL is active.*

Thermo-elastic damping is a source of intrinsic material damping due to thermo-elasticity present in almost all materials. As the name thermo-elastic suggests, it describes the coupling between the elastic field in the structure caused by deformation and the temperature field.

> *"PLANE223, SOLID226, and SOLID227 also support a thermoplastic effect calculation in static or transient analyses. For more information, see Thermoplasticity in the Mechanical APDL Theory Reference."*

Because the analysis is conducted in 2D, the PLANE223 is used. The element has eight nodes with up to four degrees of freedom per node. Structural capabilities include elasticity, plasticity, hyperelasticity, viscoelasticity, viscoplasticity, creep, large strain, large deflection, stress stiffening effects, and prestress effects. It also has mixed formulation capability for simulating deformations of nearly incompressible elastoplastic materials, and fully incompressible hyperelastic materials. Thermoelectric capabilities include Seebeck, Peltier, and Thomson effects, as well as Joule heating.

In addition to thermal expansion, structural-thermal capabilities include the piezo-caloric effect in dynamic analyses. The thermoplastic effect is available for analyses with structural and thermal degrees of freedom.

PLANE223 assumes a unit thickness.

In Abaqus is selected a Coupled Element with Thermo-Structural coupling. Transient structural effects and also thermos elasticity can't be neglected to speed up the simulation.

7.2.3 Thermo-plasticity

ANSYS theory Guide provides the equation at the basis of the FEM model. In a thermoplastic analysis, the stress equation of motion and heat flow conservation equation are coupled by the plastic heat density rate defined as:

$$\dot{Q}^p = \beta \dot{W}^p$$

Where:

β = fraction of plastic work

$\dot{W}^p$= plastic work rate = $\{\sigma\}^T\{\dot{\varepsilon}^p\}$

Deriving as said before from:

$$\frac{\{F^a\}}{vol} = \rho \frac{\partial^2}{\partial t^2}\{w\} \quad \text{(Newton's second law)}$$

$$\rho c\left(\frac{\partial T}{\partial t} + \{v\}^T\{L\}T\right) + \{L\}^T\{q\} = \ddot{q} \quad \text{(first law of thermodynamics)}$$

Where:

{w} = vector of displacements of a general point

$\{L\} = \begin{Bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{Bmatrix}$ = vector operator

{q} = heat flux vector

$\ddot{q}$ = heat generation rate per unit volume

The coupled-field finite element matrix equation for the thermoplastic analysis is:

$$\begin{bmatrix} [M] & [0] \\ [0] & [0] \end{bmatrix} \begin{Bmatrix} \{\ddot{u}\} \\ \{\ddot{T}\} \end{Bmatrix} + \begin{bmatrix} [C] & [0] \\ [0] & [C^t] \end{bmatrix} \begin{Bmatrix} \{\dot{u}\} \\ \{\dot{T}\} \end{Bmatrix} + \begin{bmatrix} [K] & [0] \\ [0] & [K^t] \end{bmatrix} \begin{Bmatrix} \{u\} \\ \{T\} \end{Bmatrix} = \begin{Bmatrix} \{F\} \\ \{Q\} + \{Q^p\} \end{Bmatrix}$$

$\{Q^p\}$ = element plastic heat generation rate load = $\int_{vol} \dot{Q}_n^p \{N\} d(vol)$

$[K^t]$= element diffusion conductivity matrix

{N} = element shape functions

n = number of substep

7.2.4 Thermo-elasticity (thermos-elastic damping)

The first recognition of the thermo-elastic effect is traditionally attributed to Weber in 1830. He observed that, when a vibrating wire receives a sudden increase in tension, it experiences a delay in the change of its fundamental frequency. The thermos-elastic effect was given a theoretical foundation in 1853 by William Thomson, later known as Lord Kelvin. Thomson derived a linear relationship between the temperature change of a solid and the change in the sum of the principal stresses for isotropic materials [36].

Thermo-elasticity is neglected in the analysis by selecting KEYOPT(9) = 1. However the coupled thermo-elastic constitutive equations are reported for completeness.

$$\{\varepsilon\}=[D]^{-1}\{\sigma\}+\{\alpha\}\Delta T$$

$$S=\{\alpha\}^{T}\{\sigma\}+\frac{\rho C_p}{T_0}\Delta T$$

Where:

S = entropy density

[D] = elastic stiffness matrix

{α} = vector of coefficients of thermal expansion

C_p = specific heat at constant stress or pressure

Using {ε} and ΔT as independent variables, and replacing the entropy density S by heat density Q using the second law of thermodynamics for a reversible change we obtain:

$$\{\sigma\}=[D]\{\varepsilon\}-\{\beta\}\Delta T$$

$$Q=T_0\{\beta\}^{T}\{\varepsilon\}+\rho C_v\Delta T$$

Where:

{β} = vector of thermoelastic coefficients = [D] {α}

C_v = specifc heat at constant strain or volume

The Finite Element Matrix equation is the following:

$$\begin{bmatrix}[M] & [0]\\ [0] & [0]\end{bmatrix}\begin{Bmatrix}\{\ddot{u}\}\\ \{\ddot{T}\}\end{Bmatrix}+\begin{bmatrix}[C] & [0]\\ [C^{tu}] & [C^{t}]\end{bmatrix}\begin{Bmatrix}\{\dot{u}\}\\ \{\dot{T}\}\end{Bmatrix}+\begin{bmatrix}[K] & [K^{ut}]\\ [0] & [K^{t}]\end{bmatrix}\begin{Bmatrix}\{u\}\\ \{T\}\end{Bmatrix}=\begin{Bmatrix}\{F\}\\ \{Q\}\end{Bmatrix}$$

Where:

$[K^{ut}]$ = element thermo-elastic stiffness matrix

$[C^{tu}]$ = element thermo-elastic damping matrix

It's easy to see that the second and the third matrix are no more symmetric, that increases the computational time when also the thermo-elastic effect is included.

7.3 ELASTOPLASTIC MATERIAL MODEL

Because the heat is also produced as a result of the plasticity at the tip of the crack, the linear elastic model has to be avoided. For simplicity and absence of material data in literature, the material is supposed to be rate-independent.

The von Mises yield criterion predicts that yielding will occur whenever the distortion energy in a unit volume equals the distortion energy in the same volume when uniaxially stressed to the yield strength.

$$\sigma_e = \sqrt{\frac{1}{2}[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2]}$$

When von Mises equivalent stress exceeds the uniaxial material yield strength, general yielding will occur.

If plotted in 3D principal stress space, the von Mises yield surface is a cylinder. The cylinder is aligned with the axis $\sigma_1=\sigma_2=\sigma_3$.

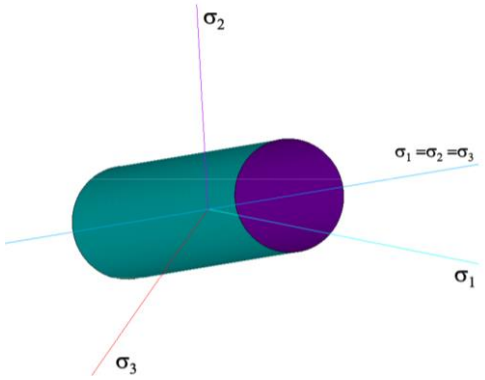


Figure 7-5 Von Mises yield criterium plotted in a three-dimensional stress space

Note that if the stress state is inside the cylinder, no yielding occurs. This means that if the material is under hydrostatic pressure ($\sigma_1=\sigma_2=\sigma_3$) no amount no amount of hydrostatic pressure will cause yielding. Another way to view this is that stresses which deviate from the axis ($\sigma_1=\sigma_2=\sigma_3$) contribute to the von Mises stress calculation.

At the edge of the cylinder (circle), yielding will occur. Hardening rules will describe how the cylinder changes with respect to yielding.

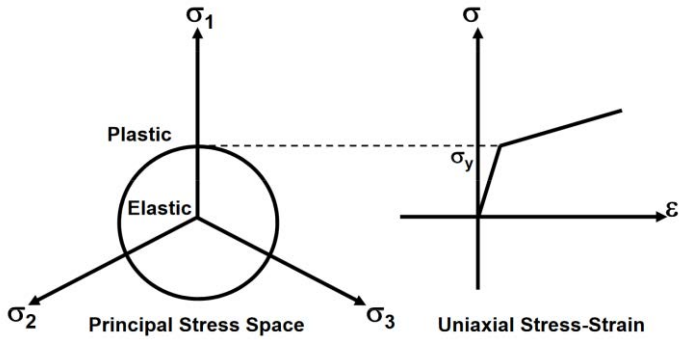


Figure 7-6 Yield criterium in a three dimensional stress space and uniaxial stress and strain relationship

The hardening rule describes how the yield surface changes (size, center, shape) as the result of plastic deformation and, more important for us, when the material will yield again if the loading is continued or reversed. This is in contrast to elastic-perfectly-plastic materials.

There are two basic hardening rules to prescribe the modification of the yield surface:

- Kinematic Hardening where the yield surface remains constant in size and translates in the direction of yielding
- Isotropic Hardening where the yeld surface expands uniformly in all directions with plastic flow

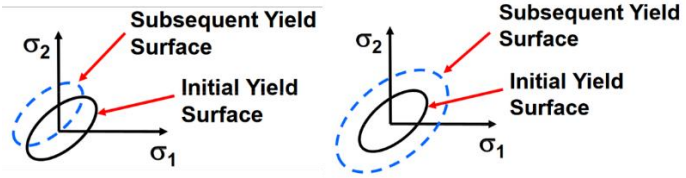


Figure 7-7 Differences in yielding surface change after plastic deformation. The left is for a kinematic hardening, the one on the right for isotropic hardening

7.3.1 Kinematic Hardening

The stress-strain behaviour for linear kinematic hardening is illustrated below:

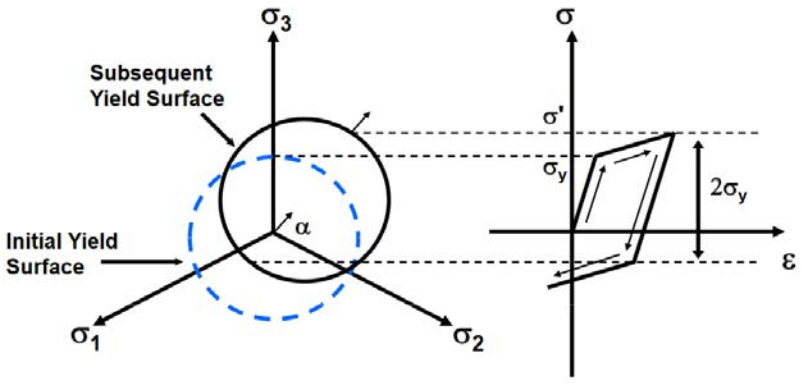


Figure 7-8 Yield criterium in a three-dimensional stress space and uniaxial stress and strain relationship for Kinematic Hardening

Subsequent yield in compression is decreased by the amount that the yield stress in tension increased, so that a 2σy difference between the yields is always maintained: this is known as the Bauschinger effect

An initially isotropic material is no longer isotropic after it yields and experiences kinematic hardening.

For very large strain simulations, the linear kinematic hardening model can become inappropriate because of the Bauschinger effect.

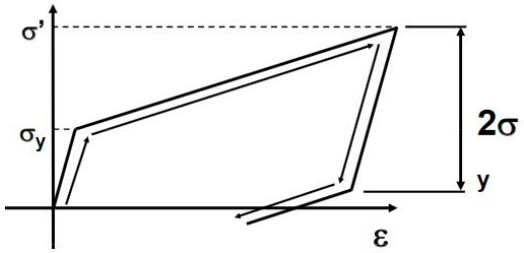


Figure 7-9 Stress and strain relationship for Kinematic Hardening

Kinematic hardening is generally used for small strain, cyclic loading applications.

7.3.2 Isotropic Hardening

Isotropic hardening states that the yield surface expands uniformly during plastic flow. The term 'isotropic' refers to the uniform dilatation of the yield surface.

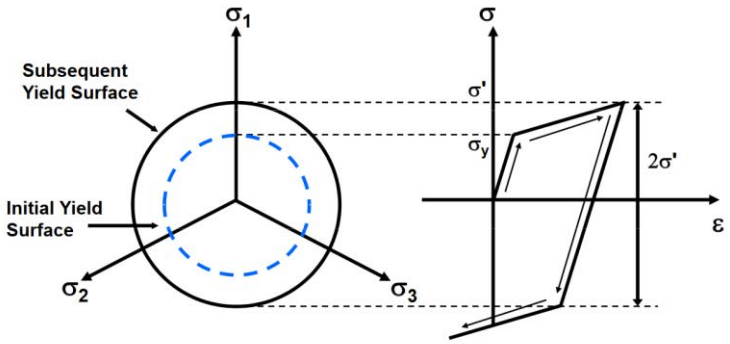


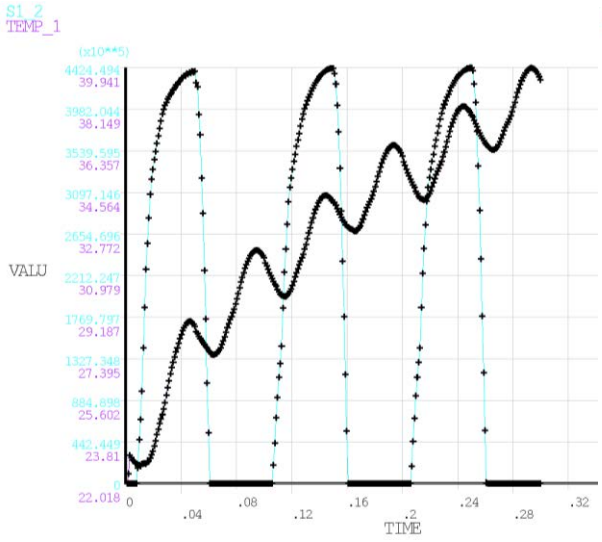
Figure 7-10 Yield criterium in a three dimensional stress space and uniaxial stress and strain relationship for Isotropic Hardening

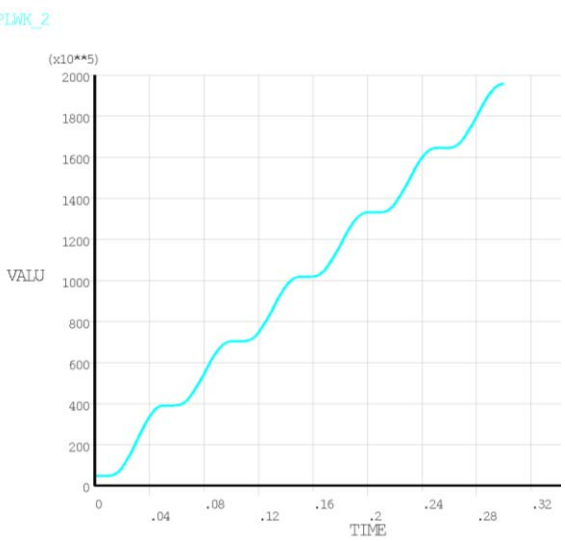
The subsequent yield in compression is equal to the highest stress attained during the tensile phase. Isotropic hardening is often used for large strain or proportional loading proportional loading simulations. It is usually not applicable for cyclic loading.

The stress-strain curve can be represented by bilinear model (after yielding it is only specified the tangent of the plastic slope) or a multilinear model (where more points are defined).

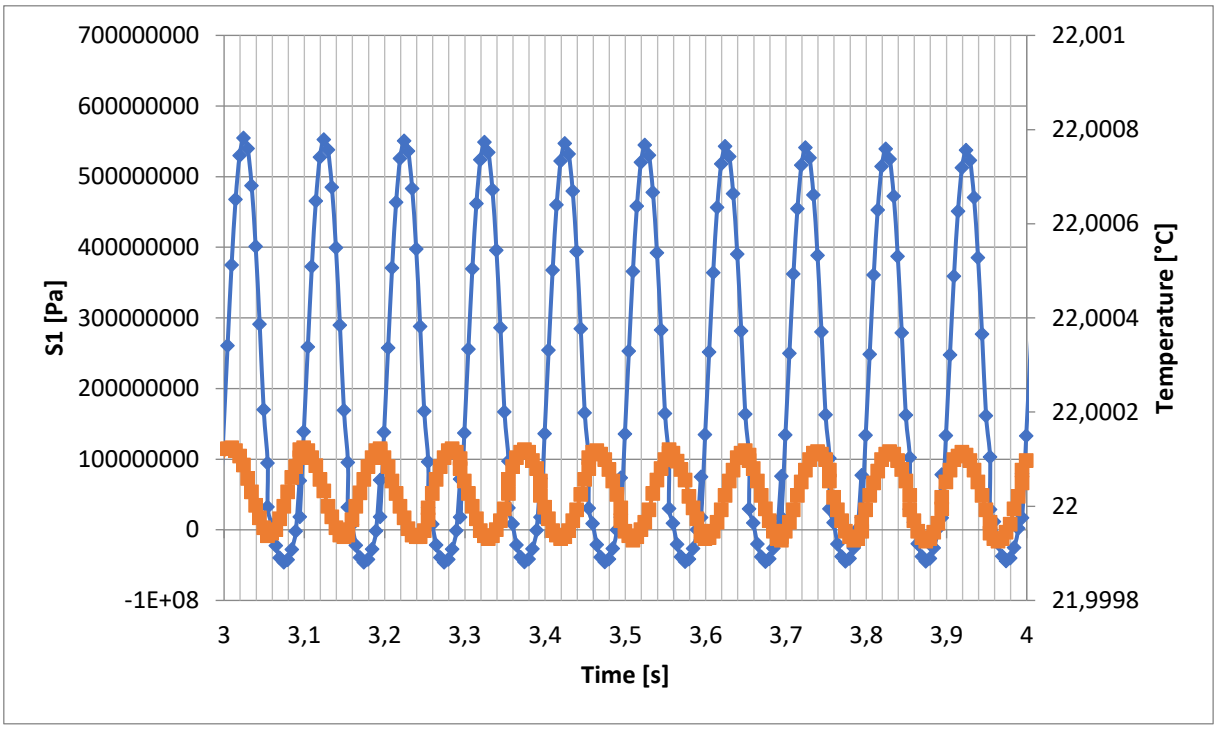
7.4 COMPARISON OF ELASTO-PLASTIC FEM'S RESULTS AND IMPROVEMENTS

The main problem in ANSYS can be see easily see by watching the Dissipated Plastic Work versus time. Instead of being an instantaneous value, it's a cumulated one. For this reason the temperature is increasing continuously, without any asymptote as expected. Temperature cycle aren't in phase with the load cycles, that could derive from thermal dissipation. Moreover also in the phase of compression the dissipated plastic work contribute to the temperature increase that could derive from the previously created plastic zone or from the fact that the contact of the two surfaces of the crack isn't modelled. Setting TIMEINT,STRUCT,ON the model would not converge with a sinusoidal load.





The results from Abaqus predict an expected shape for the temperature but the temperature increase is very small, in the order of 0.1 °mC, too low to be appreciated. The isotropic model isn't in fact the most appropriated model for cyclic loading. Only in the first cycle the temperature increment was in the order of 0.01°C that seems more reasonable. Moreover, the frequency of the two waves doesn't appear to be the same, the reason could be some numeric instability that can be highlighted more for analysis lasting for much more time.



In conclusion, to obtain a more realistic model, Abaqus is the numeric code to be used but a multilinear kinematic model need to be implemented by a user subroutine.

7.5 Results for thermo-elastic effect only

A coupled thermo-structural transient analysis as the previous paragraph has been conducted, the only difference is the absence of the thermoplastic phenomenon. The mesh is the same as the structural one. The results show to be periodic with the same frequency of the load as expected.

The crack is stationary in the simulation. This should not cause errors if the crack growths happens in a period of time few times higher than 10s (quasi-static hypotesis). It has been tried to release the Y degree of freedom for the node on the crack tip to simulate crack growth but it's no use. Like Skinner has done [37], before restarting the analysis, the reaction force should be gradually set to zero. This would require a lot of effort, especially to not lose the equilibrium, so it has not be implemented.

The simulation time has been set on 10 seconds but that's not enough to avoid the transient regime of temperature. However, the amplitude seems to remain constant during the transient phase.

For a crack length of 1 0mm the temperature amplitude is negligible (0.01 mK), the thermoelastic effect can be considered negligible. Instead for a crack length of 35 mm the temperature amplitude is of 0.0193 K. The mean temperature is always around 22 °C (21.994 °C from the simulation) because neglecting the heat dissipation the temperature cannot increase from the starting temperature, as opposite as a real case.

In fact, under cyclic loading, the stabilized temperature change due to thermoelastic coupling is equal to zero. Instead mechanical dissipation is always positive. As a consequence, the stabilized temperature change obtained during the experiments is expected to be due to the mechanical dissipation only.

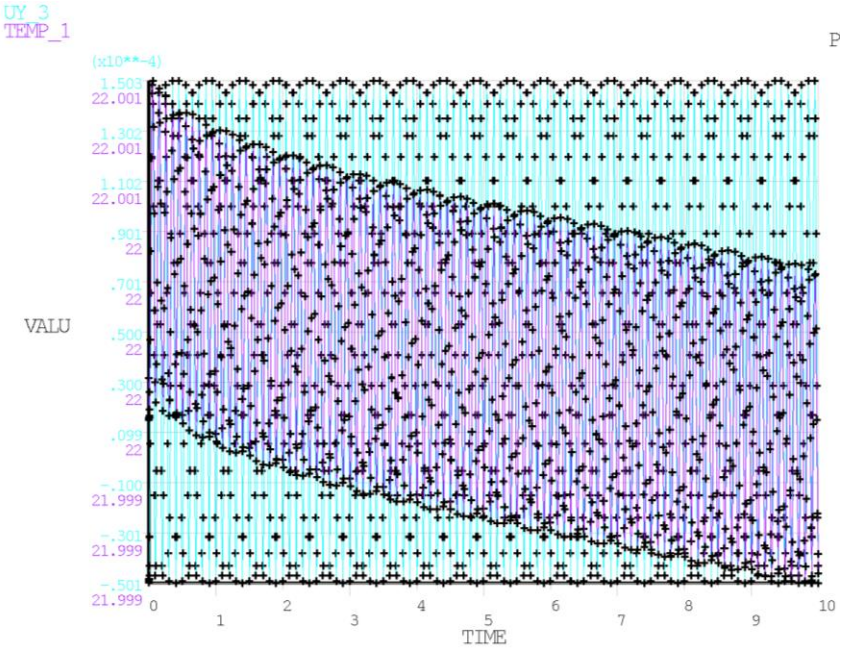


Figure 7-11 Results for displacement at the CTOD position [m] and Temperature on the crack tip [°C] versus time [s] for a crack length of 10mm. The scale doesn't give the iimpression that the temperature remains constant.

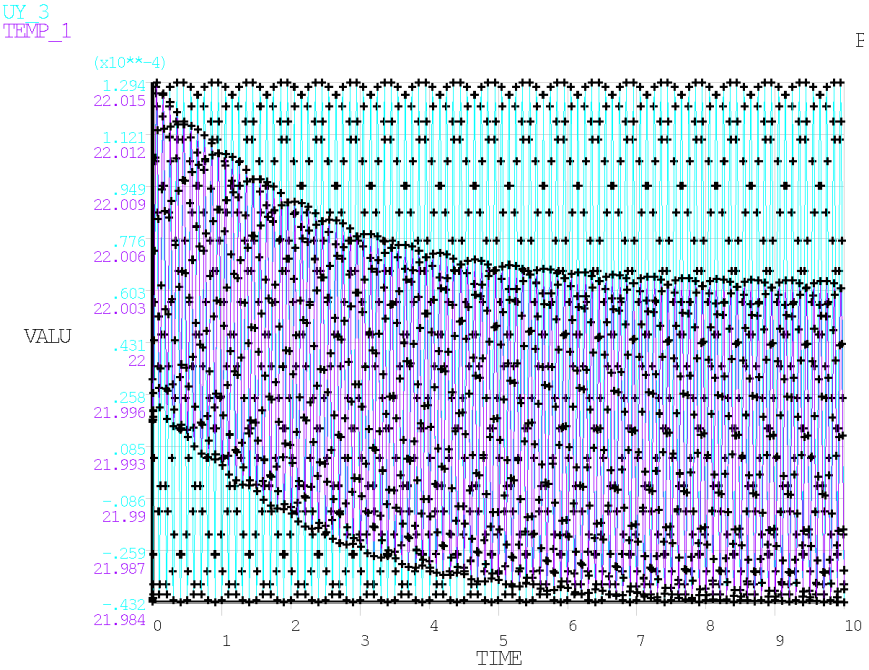


Figure 7-12 Results for displacement at the CTOD position [m] and Temperature on the crack tip [°C] versus time [s] for a crack length of 35mm

8 EXPERIMENT SETUP

The objective of the experiment is to record the temperature distribution on a C(t) specimen during a crack growth test (after the specimen has been changed to a SEN(t) specimen). The fatigue test is quite standard but careful attention had to be given to the thermographic equipment. Lots of externals infrared sources can interfere with the camera and alter the results.

Using thermocouple isn't sufficient to appreciate the temperature rise deriving from the heat dissipation, especially considering the low frequencies used for high cycles' fatigue testing and that the crack is moving.

A literature review has been done to avoid such problems.

An acoustic emission tool has been added to the set-up to measure the release of elastic energy as the crack growths.

8.1 PREVIOUS ATTEMPTS IN LITERATURE

The first study that was found on the idea of using heat release to study fatigue dates back to the early 1920s. Moore and Kommers (1921) had then already come up with the idea that the change of heat dissipation rate as a function of applied stress could give some indication about the fatigue limit.

Their experimental set up enabled them to study the influence of several parameters on fatigue life, such as surface roughness, overstressing, stress concentration, etc. Some seventy years later, this idea is still pursued (Luong, 1995, 1998; Rosa and Risitano, 2000, for instance) but very few successful industrial applications have been reported (Berard et al., 1997) [20], [38].

One of the main issues that remain open is the fact that these approaches do not rely on physical bases. Therefore, it is difficult to relate the heat release measurements to the micro-mechanisms finally leading to fatigue fracture.

8.1.1 Temperature as a variable

The use of temperature as observable state variable has been adopted in different approaches on the behaviour of metallic materials, such as large strain plasticity (Chrysochoos et al., 1989; Kamlah and Haupt, 1998; Zehnder et al., 1998; Macdougall, 2000), phase transformation in SMAs (Balandraud et al., 2001; Maquin et al., 2006; Shaw and Kyriakides, 1997; Gadaj et al., 2002; Schlosser et al., 2007), low cycle fatigue behaviour (Yang et al., 2005; Vincent, 2007), stress field measurement by thermo-elasticity (Harwood and Cummings, 1989; Wong et al., 1987) and its extension to residual stress determination (Offermann et al., 1998), or monitoring of crack propagation (Guduru et al., 2001; Charles et al., 1975), a partial review of this type of studies being available in (Pandey and Chand, 2003).

Temperature can strongly increase due to the latent heat production. In our cases mechanical dissipation due to plasticity and the thermos-elastic effect are the main contribution to the heating.

8.1.2 Moving from temperature to energy

Historically, the use of thermocouples (Moore and Kommers, 1921; Galtier, 1993) has limited the studies to the observation of the mean temperature changes due to the dissipative phenomena, considering the usual loading frequencies used in fatigue testing.

Moreover even with the advent of focal plane arrays, the thermal behaviour is not intrinsically linked to the behaviour of the material, so it is necessary to go back to the heat sources using the thermomechanical heat equation. Knowledge of the temperature field at the surface of a plane specimen can simplify to a two-dimensional diffusion problem while neglecting the temperature gradients in the thickness (Louche and Chrysochoos, 2001; Boulanger et al., 2004). It becomes then possible to reach the field of thermo-mechanical activated sources, due to thermo-elastic coupling and dissipative phenomena (Berthel et al., 2008).

$$Q(x,y,t) = \rho C \frac{\partial T(x,y,t)}{\partial t} - k\left(\frac{\partial^2 T(x,y,t)}{\partial x^2} + \frac{\partial^2 T(x,y,t)}{\partial y^2}\right)$$

8.2 Evaluate Fatigue Crack Growth rate

The procedure is based on ASTM E647-15, which scope is to determinate the fatigue crack growth rate from near-threshold to controlled instability (K_{max}).

Residual stresses can affect the subsequent fatigue crack, so noticeable crack mouth-opening displacement at zero applied force are searched as an indication of this phenomenon. Also local displacement measurements made before and after machining the crack starter notch are useful for detecting the potential magnitude of the effect (a simple gage can be used to measure distance between two hardness indentations at the mouth of the notch).

8.2.1 Specimen

The experiment requires the use of thin specimens in such a way that the surface temperature field is representative of the temperatures in the thickness. For this reason, the thickness isn't to high but still in the range of W/20<B<W/4. Moreover the small thickness eliminates the need of a through-thickness crack curvature correction.

The specimen is based on Compact Specimen C(T) of the ASTM E647 (Standard Test Method for Measurement of Fatigue Crack Growth Rates). Some dimension like the location of the hole and its diameter have been changed to fit the grip.

The material chosen is AISI 316L (stainless steel) for the reasons that will be explained in the chapter "Materials".

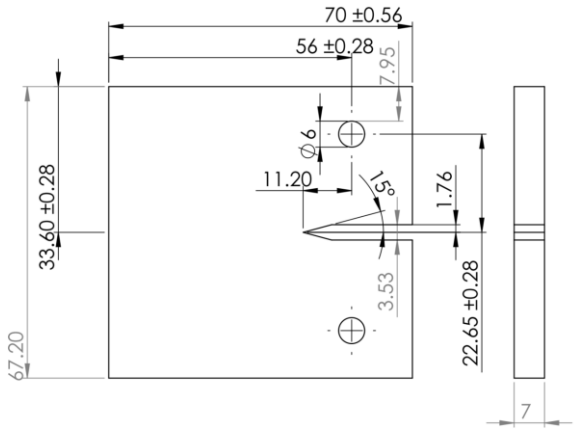


Figure 8-1 Dimension for the specimen in AISI 316L. The unit of measure is [mm]

Because of a defect on the CT grips, new specimens that can take advantage of flat grips has been designed. The width is of 25mm. The notch starts from 12 mm, in this way the pre-cracking can be performed on the machine that has a maximum load of 10 kN. The crack growth is around 10mm in this way the 1x magnification lens can see all the process without stopping the experiment.

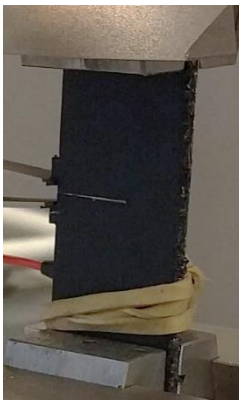


Figure 8-2SEN(t) specimen painted black to increase the emissivity

8.2.2 Crack length

A CTOD (crack tip opening device) is used to measure the crack length, this values will be compared with the values calculated from the images by the new algorithm.

8.3 REVIEW OF LITERATURE'S SET-UP AND DATA PROCESSING

Maquin and Pierron introduce in their experiments the use a dummy specimen to eliminate the variations of the temperature of the environment. They achieved this by positioning a rectangular coupon of the same material as the tested one next to the test specimen. Even if it's a good method to reduce the thermal noise it has the cons of decrease the resolution on the specimen. In fact, their test was a uniaxial cyclic loading where there wasn't any localized phenomenon. Moreover, the grips would need to be cooled down to have the same boundary conditions as the dummy specimen [20].

Because of the reason that the heat sources term is composed of two main contributions:

- the thermo-elastic coupling term that will be evidenced as a harmonic temperature variation at the excitation frequency
- the intrinsic dissipation term which will tend to heat up the specimen

They have filtered temperature variations caused by the thermo-elastic effect but this was only possible because they performed the mean of the temperature on the region of interest. This was possible because for monoaxial loading because it's reasonable to assume a uniform temperature. Note that the temperature changes caused by the thermo-elastic effect are much larger than the ones caused by the intrinsic dissipation. They set the frequency to 15 Hz such that it is high enough so that the dissipation leads to temperatures high enough to be detected and so that the thermo-elastic effect can easily be extracted.

Even if this approach is not directly applicable in our case, where we have a crack and we can't take the mean temperature on the specimen, some other method to highlight heat dissipation will be developed.

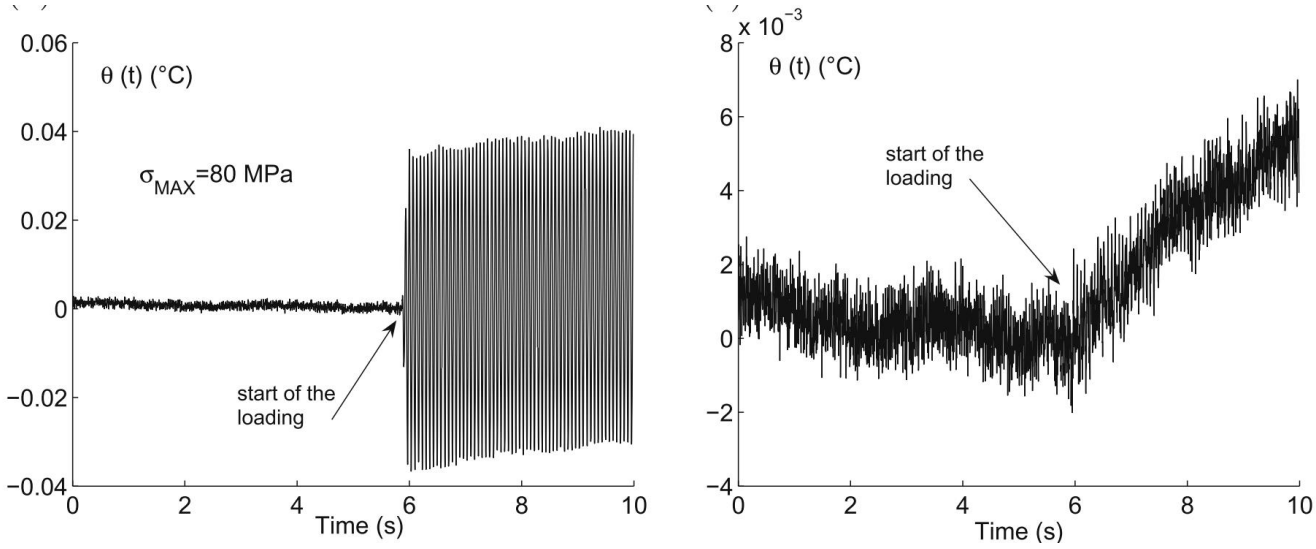


Figura 8-1 Raw and filtered thermal signals: (a) raw temperature signal and (b) temperature signal without the thermo-elastic term. Note that here is an order of magnitude of difference between original data and the filtered one

They provided careful protection of the environment around the test specimen to avoid external radiative reflections onto the specimen.

Bodelot, Sabatier et al. focus their attention to on the camera set up. First of all, the emissivity of the object has to be determined. This can be realized through a classically used method consisting in a comparison between the radiation of the object and the one of an extended blackbody, both being at the same temperature [39].

Then for a better accuracy, they turned on the camera more than 4 hours before beginning any measurements including calibration in order to get rid of thermal drift. For our Telops' camera 1 hour

is enough to reach stabilisation. The working place around the camera is protected from any reflections of the environment thanks to black tissue. Every element close to the camera is also covered by black tissue or painted with high emissivity black paint in order to dispose of any residual reflection.

In order to disturb as less as possible the thermal measurements, the working zone is protected from its environment as it has already been explained, light is turned off and the room temperature is kept steady thanks to air conditioned [39].

8.3.1 Spatial resolution

From the FOV chosen by Bodelot, Sabatier et al., their pixel size is of 6.5 μm (resolution of 1368x1024 pixels for an area of approximately 8.9x6.7 mm).

Fedoova, Bannikov et al. and also Iziumova and Plekhov used a spatial resolution of 0.1 mm (resolution of 320x256 pixels) [40], [41].

Meneghetti and Ricotta used a spatial resolution of 20 μm (resolution of 320x256 pixels) [42].

8.3.2 Sensitivity of the experimental configuration

The detection threshold is measured experimentally from the specimen at rest. After some seconds of recording the heat equation source is calculated. It should be expected to remain zero because no load is applied. Dividing the specific heat source ($[\frac{J}{m^3}]$) by ρC, the temperature increase required to produce the threshold amount of energy is evaluated. The maximum back ground noise is of 7 10^7 J/m^3 which corresponds to 1.4 mK of noise after filtering.

Luckily it would be lower than the noise equivalent temperature difference which for thermal cameras is usually of 20 mK.

The accuracy can vary from 0.1°C to 0.2°C depending on the exposure time (lower exposure time corresponds to higher electrical noise) but we are more interested in energy accuracy.

8.3.3 Adiabacity

Most metals can be safely considered adiabatic for frequencies above 2 Hz, but other factors such as the coating and system electronics might introduce concern below 10 Hz. Higher loading frequencies usually can improve the detail in the data since heat transfer is minimized [36].

The presence or absence of nonadiabatic response can be observed by checking the phase of the response signal with respect to the load signal. If heat conduction takes place within a component or coating, a detectable phase shift will occur between the reference and loading signals.

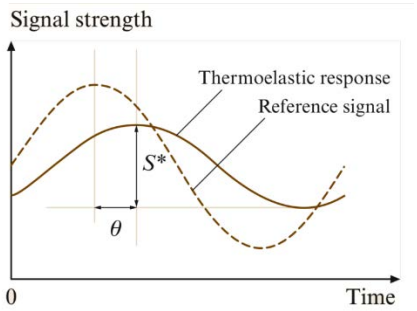


Figure 8-3 The form of the thermos-elastic data recorded during a cyclic analysis: the intensities of the reference signal and therm-oelastic signal S∗ versus time t, and the phase relationship ϑ between them

8.3.4 Specimen's surface

It's necessary to use some high emissivity paint on the surface to have a value close to 1. For this reason, the surface has to be smooth. The coating will introduce a phase between the temperature on the specimen and the measured from the camera but the layer is thin enough not to influence the results.

Meneghetti and Ricotta polished the surface by using progressively finer emery papers, starting from grade 100 up to grade 1000, and after that a diamond abrasive powder.

8.3.5 Data processing

Different expedient has been used to remove noise. One is to cancel out Fixed Pattern Noise by subtracting an image of a uniform scene from the image of interest.

Others used motion compensation to improve results. Common to everyone is noise smoothing using median filtering to remove spiky noise and then apply a Gaussian filter.

Vandone is the only to have tried to use wavelets for filtering but he doesn't provide a comparison with conventional filtering [22].

8.4 EXPERIMENT SET-UP

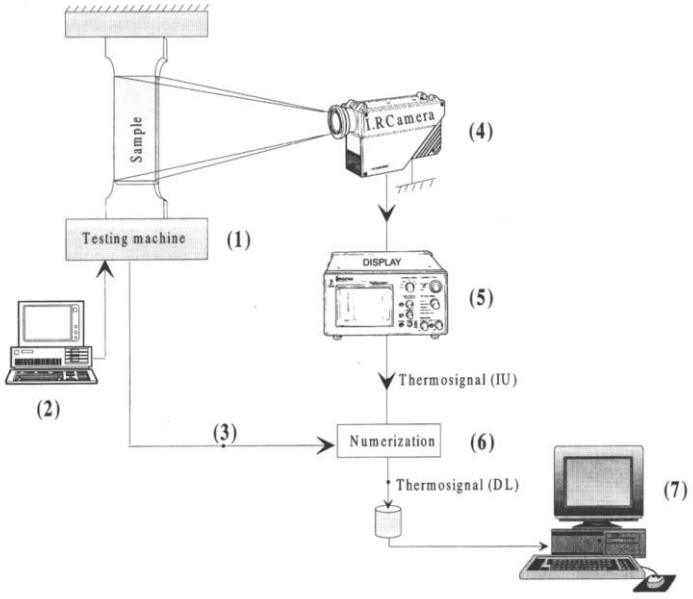


Figura 8-2 Experimental set-up: (1) testing machine (Instron E10000), (2) control computer, (3) recording of global load and displacement, (4) infrared camera (Telops High Speed Camera), (5) display unit, (6) numerizer (12 bits), (7) storage and processing of thermal image [43] *The AE is not included in the graph.*

8.4.1 Fatique Machine

As a testing machine, The ElectroPuls™ E10000, all-electric test instrument designed for dynamic and static testing on a wide range of materials and components, is used. It has high dynamic performance, capable of performing to more than 100 Hz and ±10 kN dynamic linear load capacity. For our application the frequency will be limited to 10Hz to take advantage of the high fps of the camera.

Figure 8-4 Instron ElectroPlus E10000 Fatique and Static Testing Machine

It can also be equipped with a hoven to perform high temperature testing but the window is not translucent to IR light.

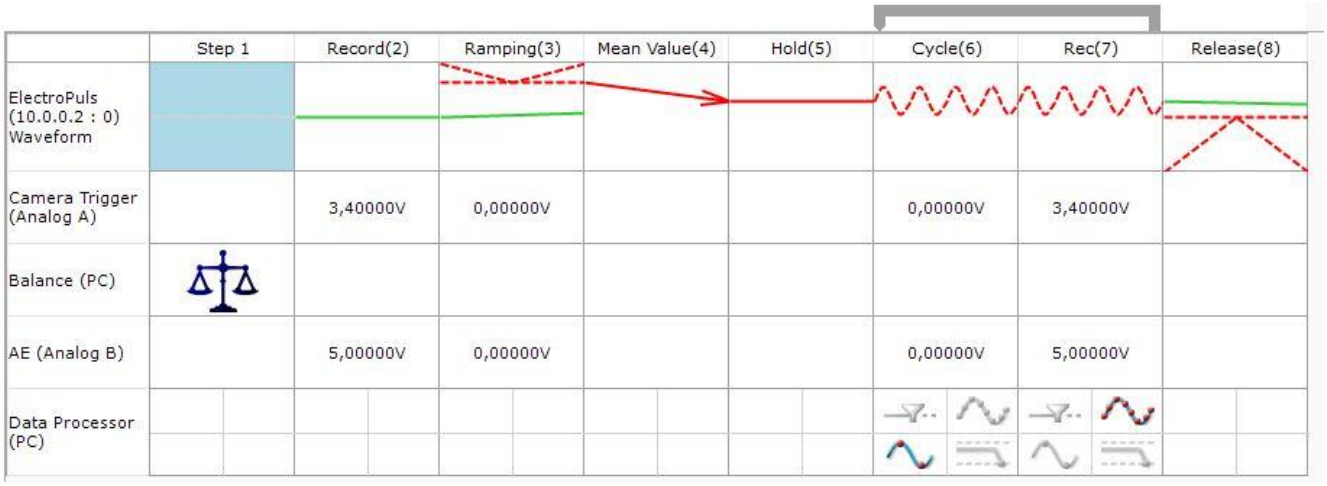
8.4.2 Acoustic Emission

Acoustic emission is the phenomenon of radiation of acoustic (elastic) waves in solids that occurs when a material undergoes irreversible changes in its internal structure. When crack growth or local yielding occurs they results in a sudden release of energy, part of which will be converted to elastic waves. These elastic waves are readily detected by piezoelectric transducers. The amplitude of the received signals can also be used to give an indication of the rate of growth of the defect.

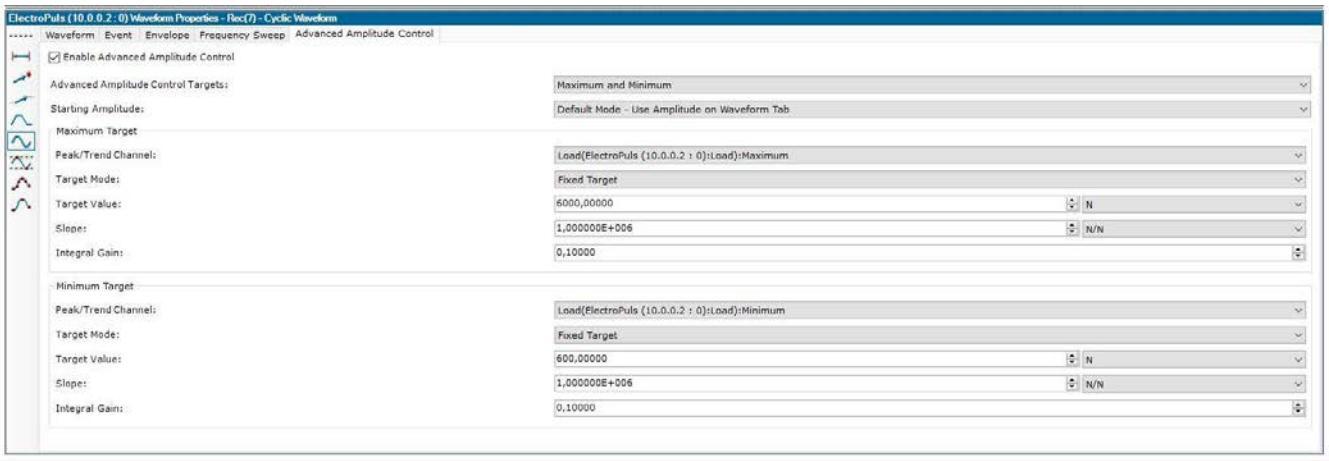
A PCI-2 AE system from Physical Acoustic Corporation has been used.

8.4.3 Test Sequence and Triggering (WaveMatrix Software)

The following testing method has been adopted. A loop of the cycling phase and recording phase has been implemented until the failure of the specimen. Complete data recording of test parameters is done only for the recording phase to save space.



An Advanced Amplitude Control has to be used to obtain always a sinusoidal wave. The controller needs the values for maximum and minimum to be respected. It's also necessary to increase the gain of the machine. Be careful that if the load is too high or the specimen too stiff, the system becomes unstable.



The first 0.2 s of the recording phase has to be discarded (for the camera is done from the Trigger Delay) because of a transient phase between the steps of the sequence. Using Advaced Amplitude Control this phase is very short. To completely avoid this discontinuity, a single step should be created and then control the analog outputs by programming the machine with C# in Advanced Calculations.

8.5 Common problems and Systematic errors

Temperature measurement with infrared cameras are classified in the following way:

- Errors of method;
- Calibration errors;
- Electronic path errors.

Under real conditions, errors in the method can follow from the reasons below or interactions occurring during measurement:

- Incorrect evaluation of object emissivity, ambient temperature, background temperature, camera-to object distance and relative humidity;
- Influence of the ambient radiation (direct or reflected from the object) arriving at the camera detector;
- Incorrect evaluation of atmospheric transmission and atmospheric radiance;
- Detector noise.

8.5.1 Thermal Reflections

The reflected temperature is the apparent temperature recorded on the sample surface due to the reflection of some radiation coming from the surrounding. An easy method to estimate this reflected temperature is written below:

- Crumble up a large piece of aluminium foil;
- Uncrumble it and put it in front of the sample facing the camera;
- Set the emissivity to 1 and measure the temperature.

- Set this value as reflected temperature in the camera

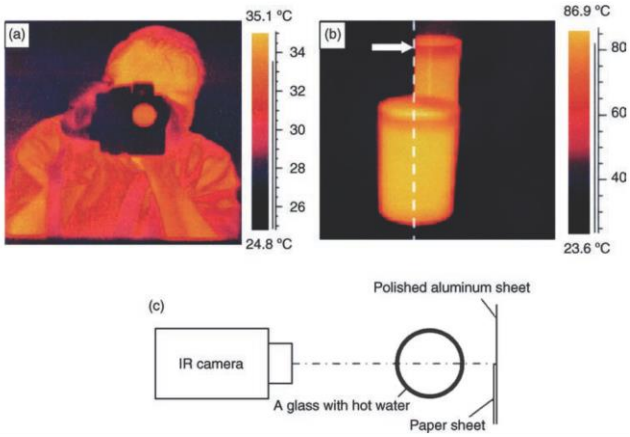


Figure 8-5 (a) A polished aluminum sheet, whose temperature is close to the ambient temperature creates a mirror reflection; (b) image of a glass of hot water located against the sheet and its background reflection (right-hand edge of paper sheet is marked with dashed line), (c) view

The data processing methods employed to improve the detection threshold include image subtraction, filtering and smoothing. The first image (before the start of the cyclic loading) can be subtracted from the subsequent images. The effect of thermal reflections are supressed through this operation. This feature has been implemented but no used because especially in the last phase, when high deformations are present, it has no use.

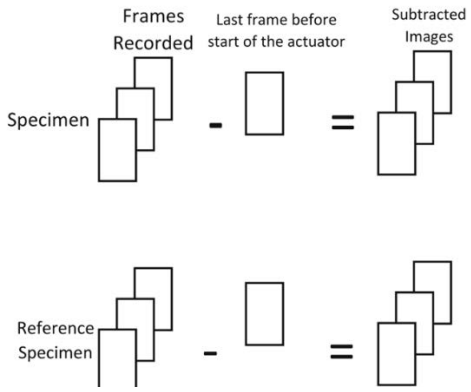


Figure 8-6 Image subtraction to remove fixed noise and reflections

8.5.2 Emissivity

Stick on the object's surface a piece of material (sticker) of high and accurately known emissivity (e.g. 0.95) and of good thermal conductivity, or paint part of it with a special paint of known and high emissivity.

- Heat up the object to a temperature at least 50°C higher than the ambient temperature.
- Set up the camera spot point (SP) on the part of the object with the sticker (or previously painted).

- Set on the camera the known emissivity of the sticker (or the paint) and measured earlier values of atmospheric temperature, the ambient temperature, the camera-to-object distance and atmospheric humidity.
- Read the spot point temperature of the area of known emissivity.
- Move the spot point outside the area of known emissivity.
- Change the parameter of the object emissivity in the camera and read the spot point temperature until it is the same as for the 'clean' area of known emissivity.

A variant of this method consists of determining the object's temperature by using a contact method. Then, the emissivity parameter in the camera should be tuned until the same reading of temperature is obtained. The value of the parameter set last represents the object's emissivity. In another method, a hole of depth at least six times larger than its diameter is drilled in the surface. Such a hole can be treated as a black body of emissivity 1. This method is an approximate method, since the hole distorts the temperature field of the object's surface.

For more accurate measurements, it is possible to set your own emissivity value.

- Attach a piece of electrical tape of known emissivity (usually around 0.97) onto the specimen;
- Set the emissivity mentioned above in the camera device, focus and write down the tape temperature;
- Assuming that the tape and the specimen are in condition of thermal equilibrium, point now to the sample surface and change the emissivity setting until the camera displays the same temperature as the one previously recorded.

8.5.3 Other important variables

8.5.3.1 *Humidity and Atmosphere temperature*

Atmospheric transmission depends strongly on the radiation wavelength so infrared devices operate in bands of maximum possible transmission (short-wave or long wave) [15].

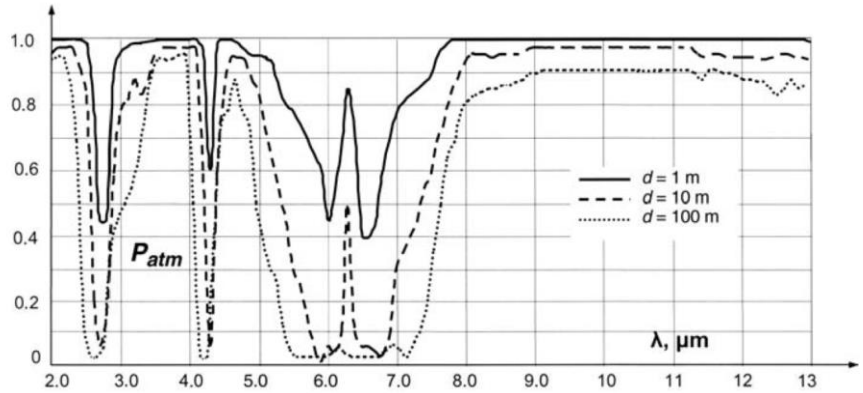


Figure 8-7 Changes in the atmospheric transmissivity for different values of camera-to-object distance d, in fact the characteristic of the atmospheric transmission is a function of camera to object distance. SW cameras operate between 2-5 μm, instead LW cameras operate between 8-14 μm. Also relative humidity influence the transmittance

The most important role in absorption of the infrared radiation for wavelength 4.3 µm is played by carbon dioxide, present in exhaled air. For example (Rudowski 1978), after 3 hours in a closed room of about 40 cubic meters in volume, the concentration of CO_2 exhaled by two people was such that at a distance of 0.8 m, 70% of radiation, for the carbon dioxide wavelength, was absorbed by the air.

Also the presence of water into the atmosphere has effect in the absorption of the radiation emitted by the specimen. Anyway, if no humidity instruments are available, it is possible to set the value at 50%. It's important to measure it also for the ASTM E647 fatigue test.

The temperature of the atmosphere normally has not a big influence, assuming of working in standard conditions (-20 to 400 °C).

8.5.3.2 *Measurements trough inspection window*

Moreover, if the specimen is put in the heating chamber to perform high temperature testing, a glass window is present between the camera and the specimen. The material for this window must be transparent within the camera's operational bandwidth (SW: 2–5 µm; LW: 8–14 µm). To make possible visual control of the measurement, the window should also be transparent within the visible radiation bandwidth.

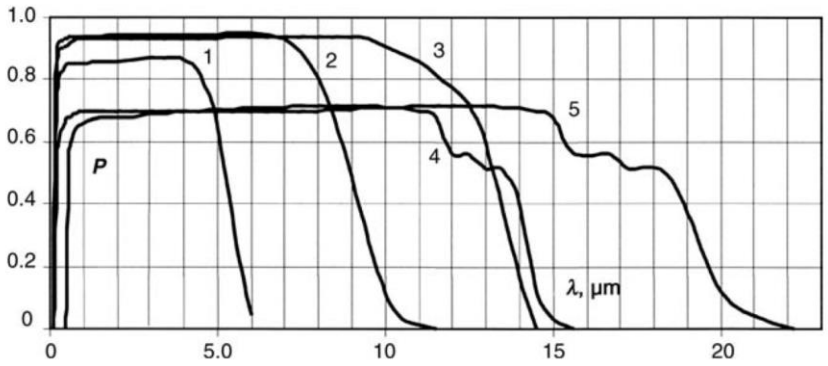
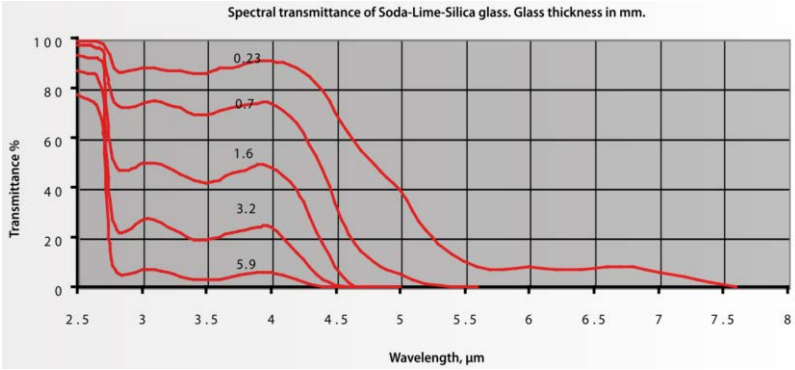


Figure 8-8 Transmission coefficient of several materials used for inspection windows: 1, Al_2O_3 ; 2,CaF_2 ; 3, BaF_2 ; 4, ZnS; 5, ZnSe. The bands can change depending on the thickness



When an IR camera needs to see through a material to measure the temperature of an object behind it, signal attenuation and ambient reflections can make accurate temperature readings difficult or

impossible. In some cases, an IR filter can be placed in the camera's optical path to overcome these problems.

Inserting a spectral filter into the camera's optical path is called spectral adaptation. An optical (IR) filter must be selected that blocks all wavelengths except the band where the object absorbs. This ensures that the object has high emissivity within that band.

8.5.3.3 *Distance Object-Camera*

The minimum focal length of the camera has to be respected.

8.5.3.4 *Ambient Drift*

Another important consideration in the calibration process is the radiation caused by the heating and cooling of the camera itself. The radiation that results directly from the camera is called parasitic radiation and can cause inaccuracies in camera measurement output, especially with thermographically calibrated cameras.

9 RESULTS AND CONCLUSIONS

The test has been conducted with a maximum load of 6500N and varying loading ratio. Three very different values has been selected: R=0.1, R=0.3 and R=0.5. For R=0.1 two specimens has been tested to evaluate the repeatability of the testing configuration. Also a series of test, in which the loading ratio has been kept constant to 0.1, with varying maximum load has been conduct.

Using the 1x magnification lens, a spatial resolution of 15 μm has been achieved. The resolution has been lowered to 512x400 pixels. In fact, decreasing the frame height a higher frame rate can be used. In the future a lower pixel height can be used (the pixel width doesn't change the maximum fps, only the space). The exposure time has been fixed to 2500 μs, this causes more noise but permits to lower the fps. If the frame height will be lowered then the exposure time could be increased to remove noise.

The area of integration is 50 pixels width and 70 pixels height, which corresponds to 0.75mm x 1,05mm. the area has been chosen by graphical individuation of the process zone and it's a little bigger to account for motion and small error in the crack tip position.

The small roughness of the painted surface is supposed to be the main source of noise, the low value for the exposure time causes instead noise in the time dimension. For conventional filtering a median and Wiener filter have been applied first and then a anisotropic Gaussian filter ($sigma_{x,y}$=8.8 $sigma_t$=0.1). For wavelet filtering db10 has been used. Because of similar results between the two methods, only result for conventional filtering are presented.

Because of an incompatibility of CameraLINK acquisition card with Windows 10, a lag is present between the input and the effective start of recording. Until new drivers for Windows 10 are released, a test should be performed to measure the precise time of this lag. After manual synchronisation between the Instron machine's data log and image frames the following graph, correlating load and the heat source function, has been produced.

The thermo-elastic effect in the unloading phase the generation of heat is positive, in the loading phase instead (traction) the heat is negative. For this reason, the mean is around zero because, excluding a small plastic dissipation, this effect is reversible.

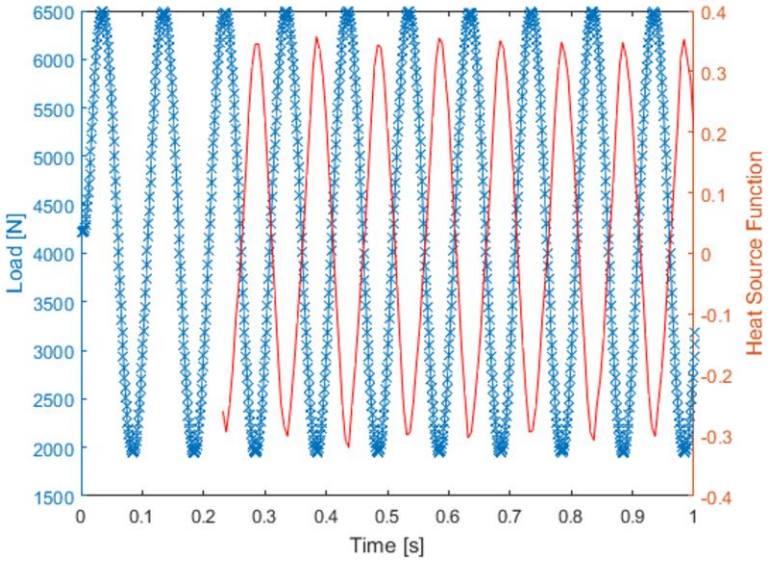


Figure 9-1 Correlation between load and heat function

For this specimen, the maximum load is of 6500 N and R=0.1. The name of the folders of camera and AE data is 'Specimen 2'. The specimen is marked as #1. The surface is ruined by some scratches but the last phase is not affected too much because the temperature increase is enough to cover the defects. The first growing phase is more affected by paint melting or oil residuals from the machining. This problem put out of focus the crack's process zone and caused problems also with the crack tip detection algorithm.

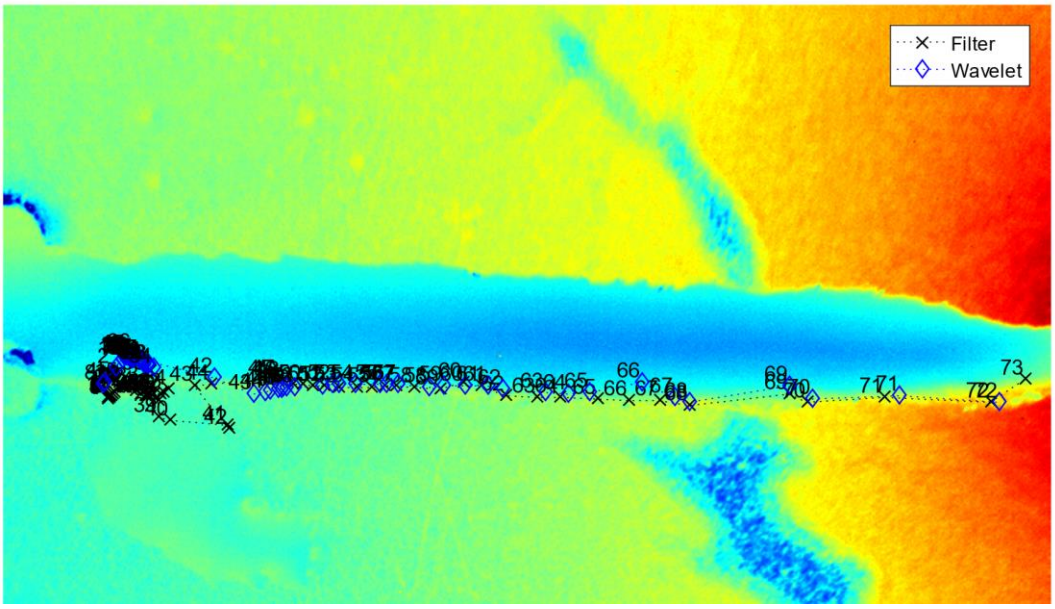


Figure 9-2 Crack growth for F=6500N and R=0.1

For the problems explained above, obtaining a smooth a(N) curve is not possible. The fitting has been made with a 3rd degree polynomial equation to reduce oscillations. In this case it's not enough to obtain a good da/dN curve especially in the first phase.

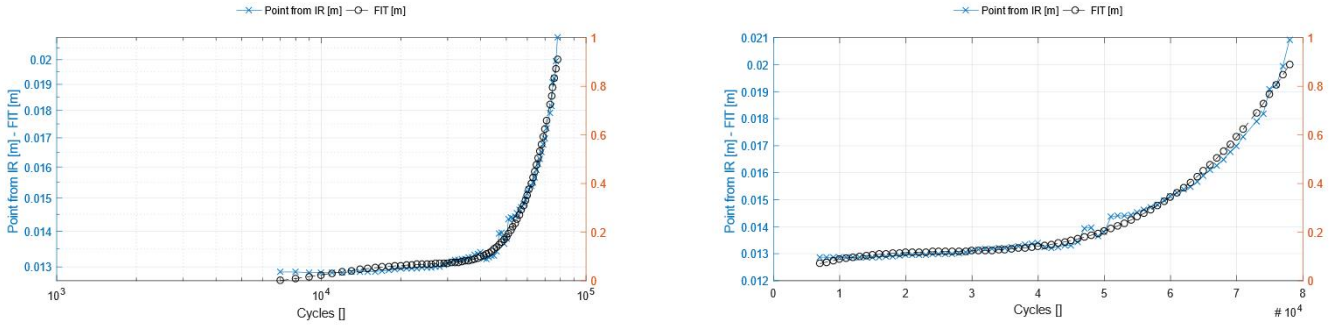
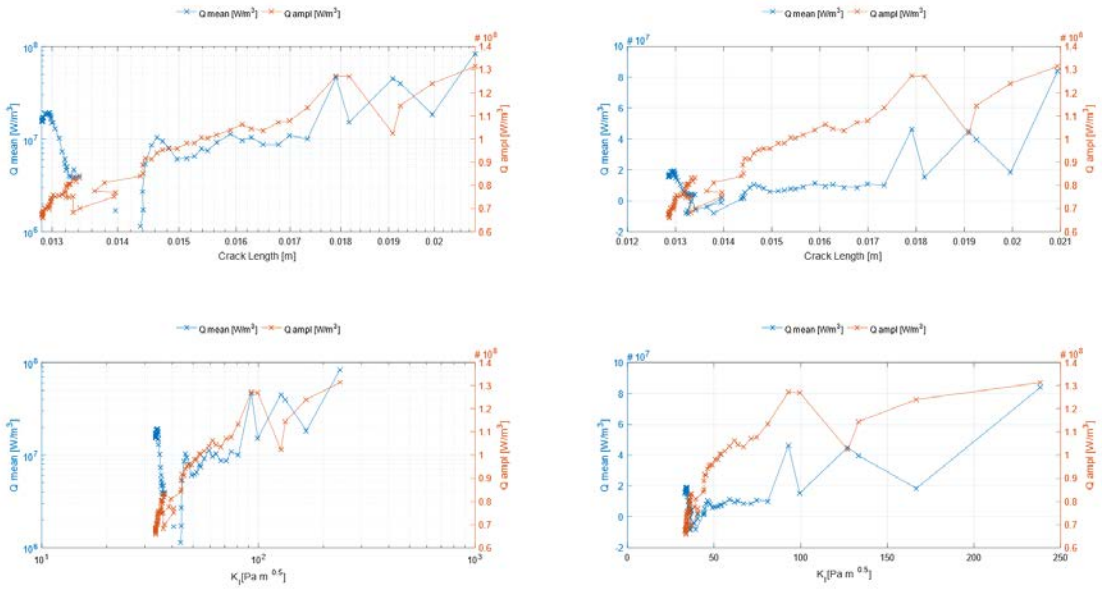


Figure 9-3 Crack length versus numer of cycles. On the left on double logaritmic scale, on the right on normal scale

Two main parameters would likely be representative of the crack growth process. The shape of the heat function is sinosuidal as the loading function, so the mean and the amplitude characterize the wave. The frequency is the same as the loading function.

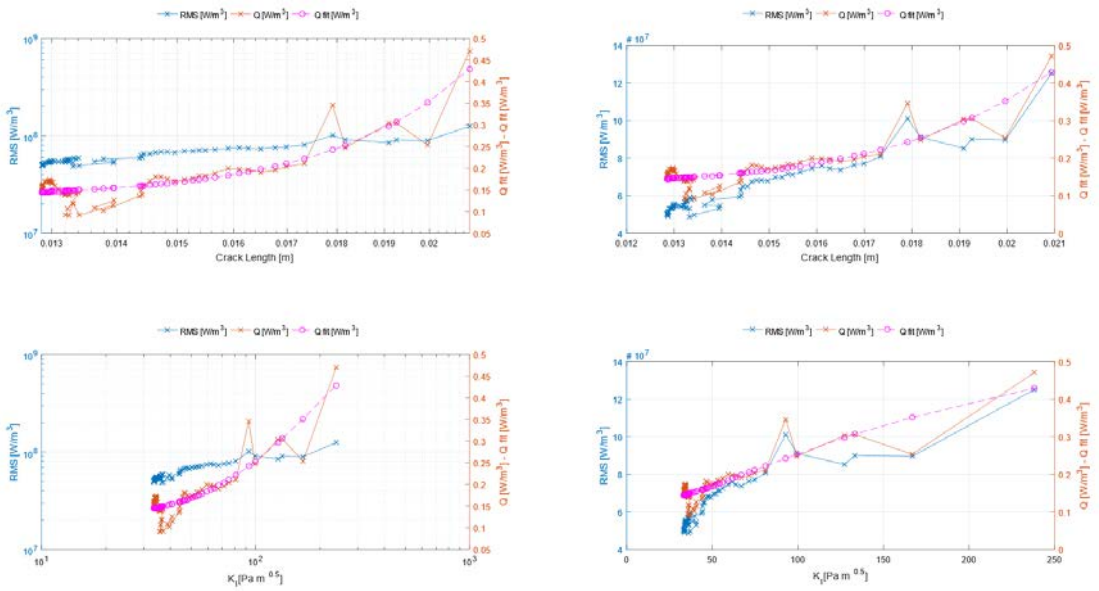
For the problems said above, the first part should not be considered. Moreover, the notch could influence the generated heat. As expected both the amplitude as the mean are increasing as K_1 increases.

The absolute value for the mean is very small, the same order as the mean residual error for the sinusoidal fitting, but still a trend can be read.



A better means of comparison is the RMS (Root Mean Square) of the function because the optimizer behind the fitting algorithm works in minimizing the root mean square error. Also, the maximum value

for the heat prove to be a good index of the crack growth. A power law fit ($a \cdot x^b + c$) seems to describe very well this behaviour.



For this other specimen, the maximum load is of 6500 N and R=0.5. The name of the folders of camera and AE data is 'Specimen 7'. The specimen is marked as #6.

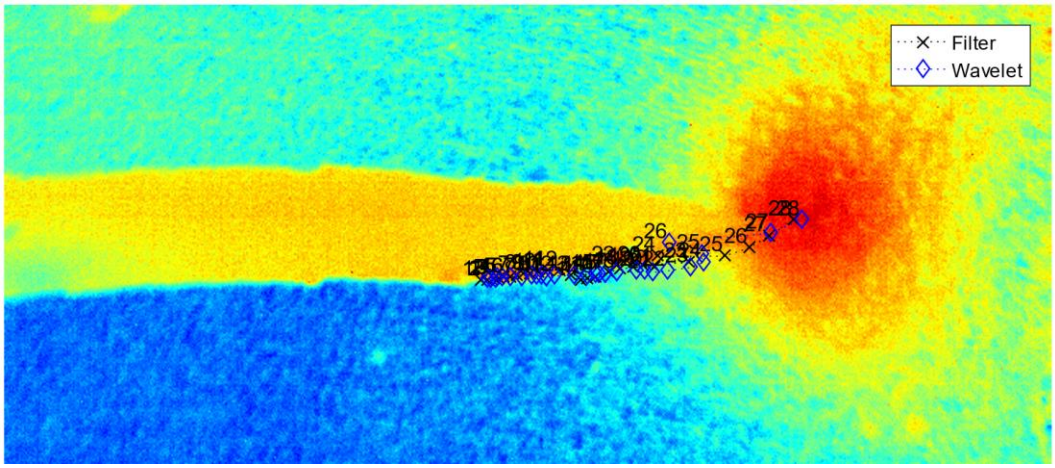
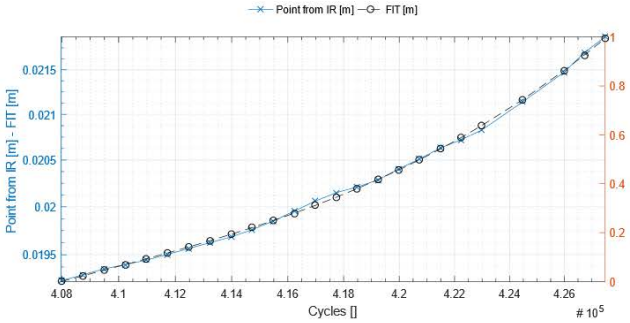
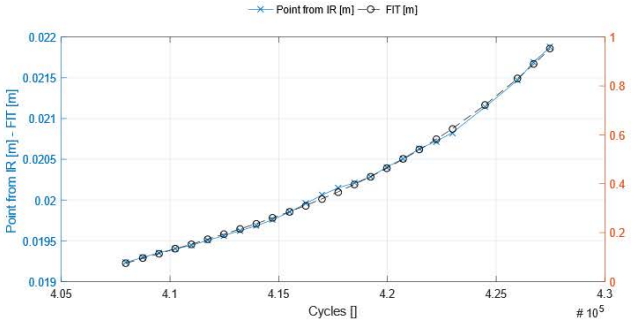


Figure 9-4 Crack growth for F=6500N and R=0.5

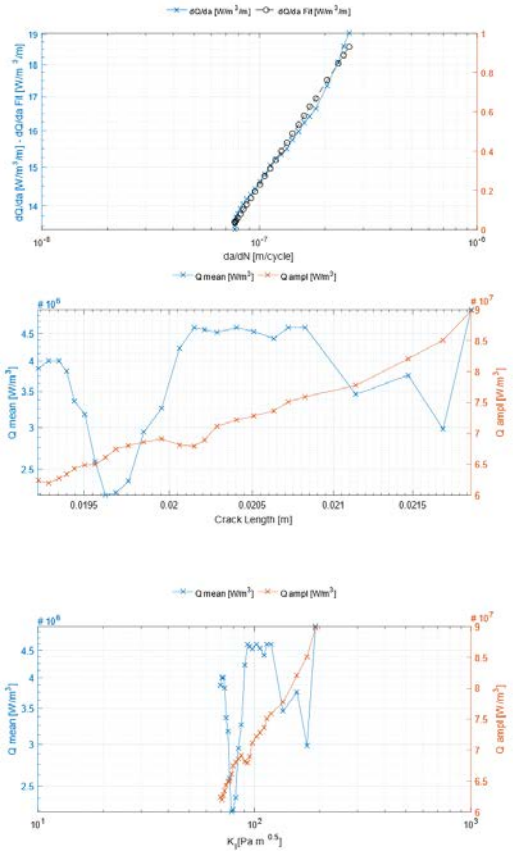
This specimen has data only for the final phase. This region is less prone to errors so the curve a(N) is more smooth and also da/dN.

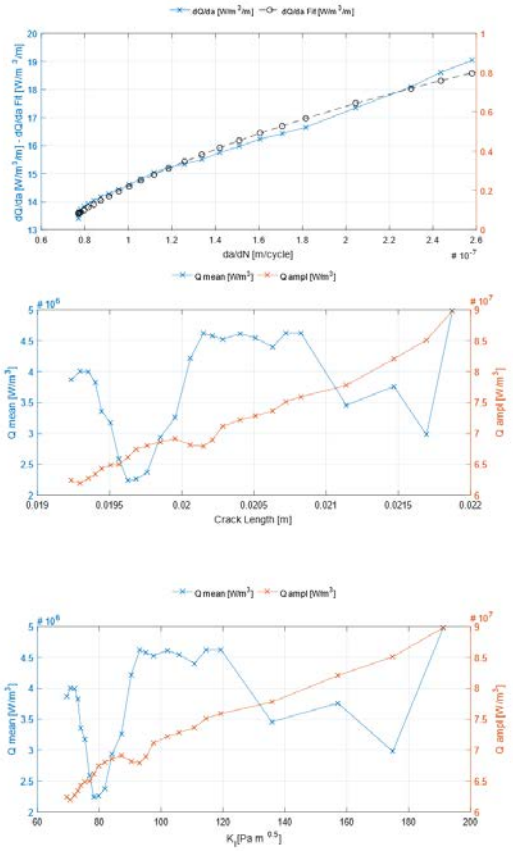




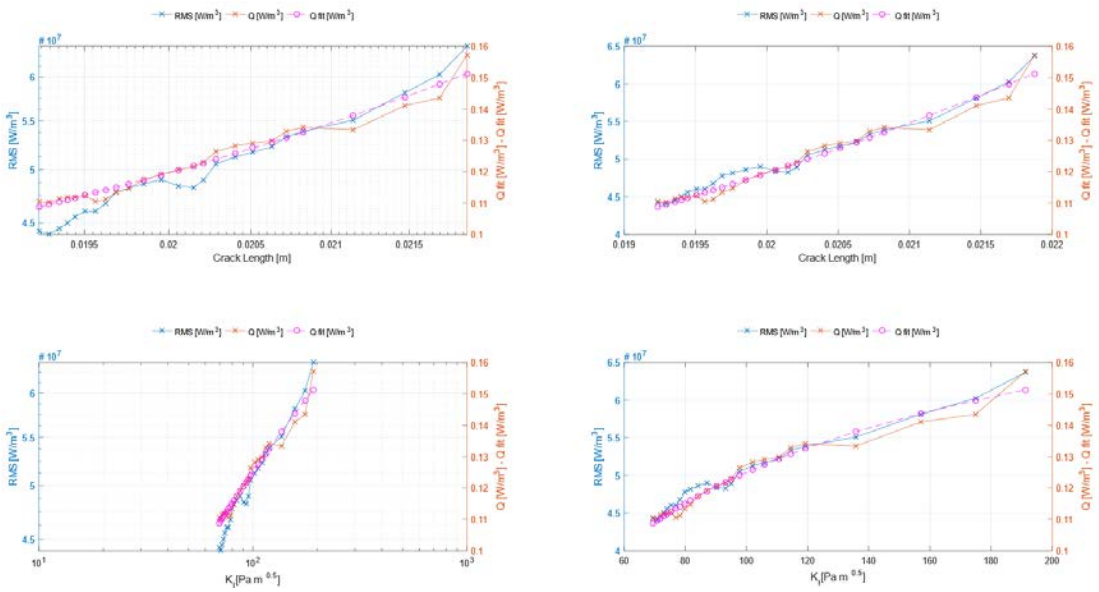
Because of a smooth da/dN curve, a plot of dQ/da over da/dN can be done. A linear relationship in the double logarithmic scale can be seen even outside the validity of the Paris Law.

The Q mean can be considered constant, small oscillation can be seen only because of the scale, which is smaller because data include only the last phase. The Q amplitude is always increasing as expected.





A linearity of the RMS (as the same of the maximum Peak for Q) can be still seen.



The maximum load is of 6500 N and R=0.1. The name of the folders of camera and AE data is 'Specimen 9'. The specimen is marked as #7. This specimen serves to check the repeatability of the configuration.

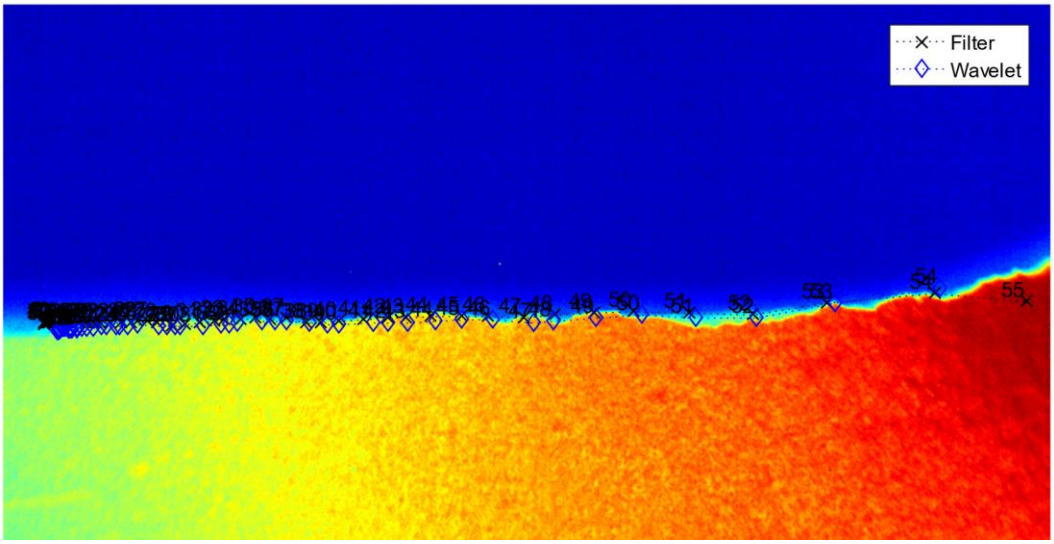
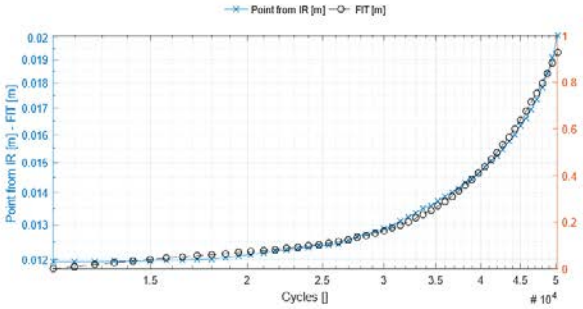
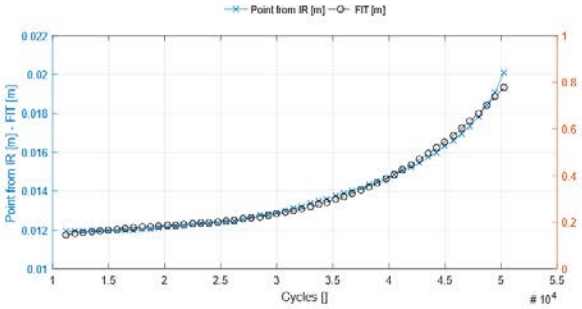


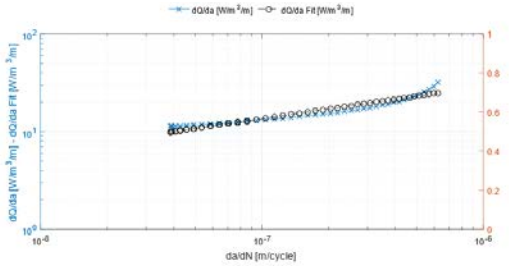
Figure 9-5 Crack growth for F=6500N and R=0.1. Second specimen analysed to check repeteabilty

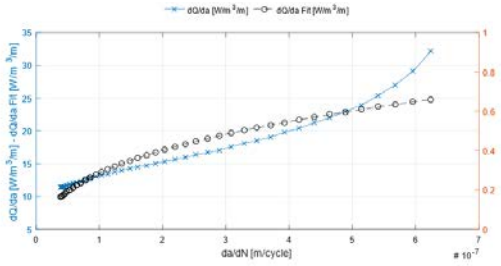
The a(N) curve resulted very smooth.



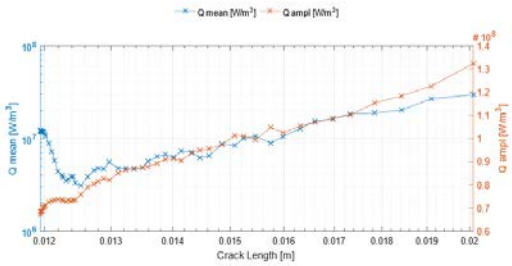


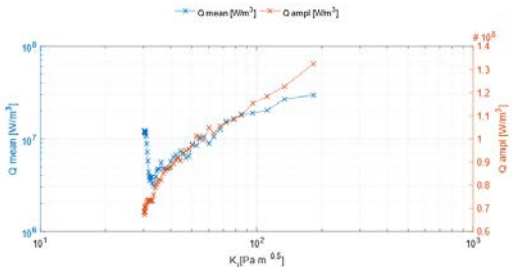
As a results also dQ/da is linear on double logarithmic scale when compared versus da/dN.

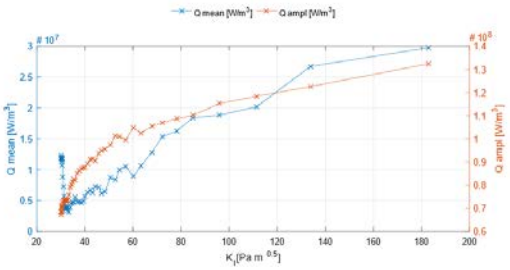




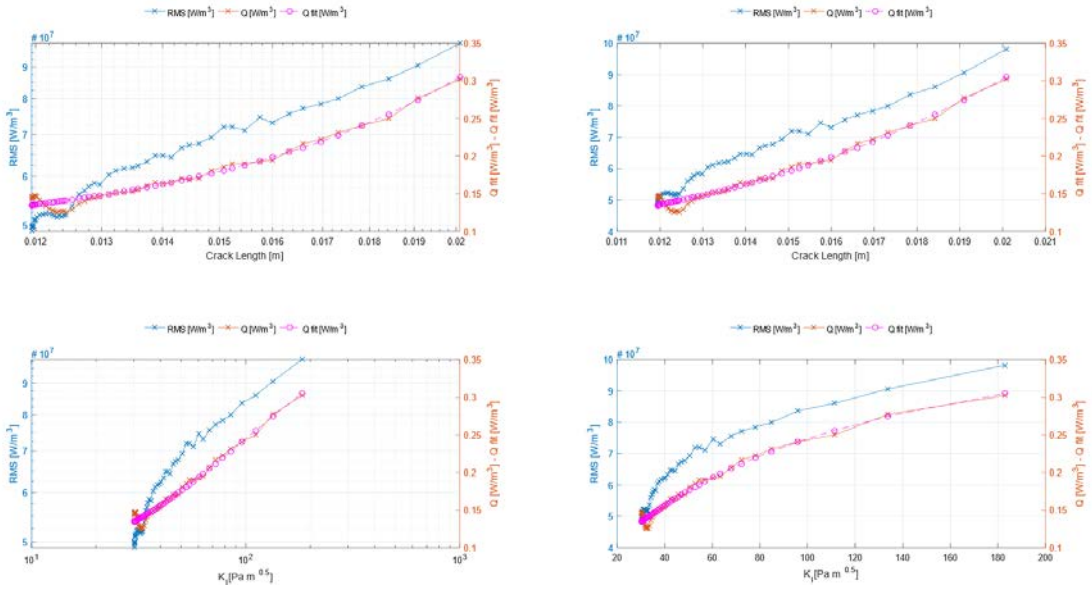
A peak for the first phase in the mean of the Q function is present also in this case. Some of the reasons for this fact are tried to be explained above.







The RMS proves to be always a good parameter to describe to process.



For the next specimen, the maximum load is of 6500 N and R=0.3. The name of the folders of camera and AE data is 'Specimen 13'. The specimen is marked as #11.

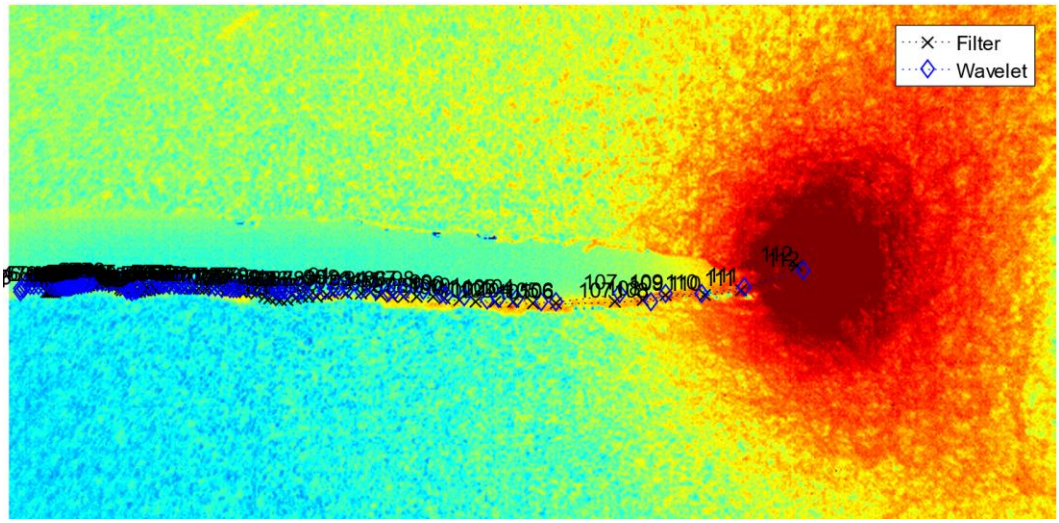
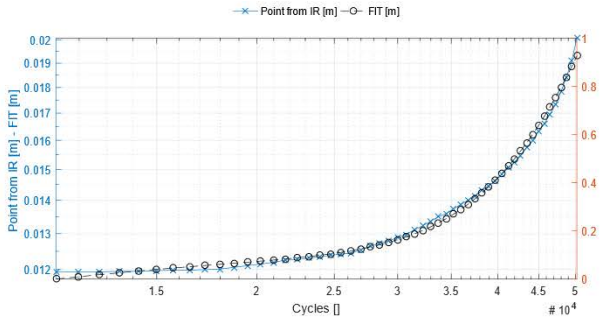
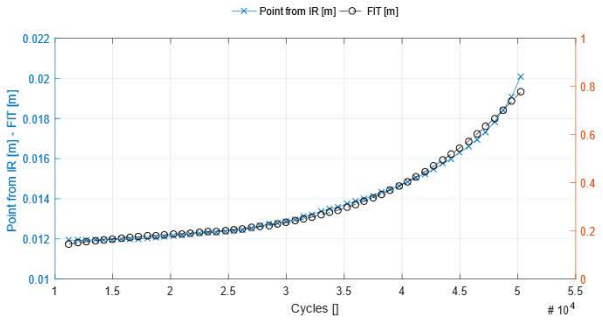


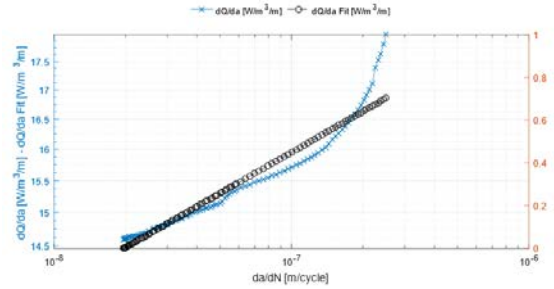
Figure 9-6 Crack growth for F=6500N and R=0.3

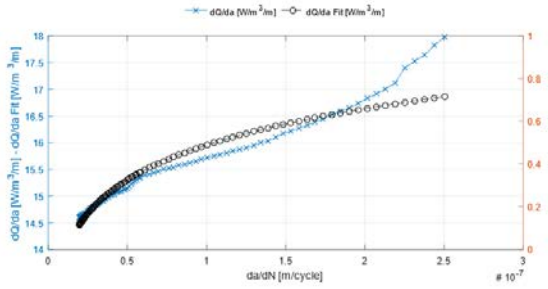
The absence of defects on the surface permits to have a smooth a(N) curve.



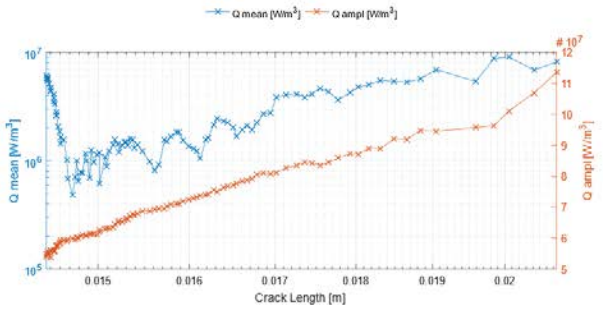


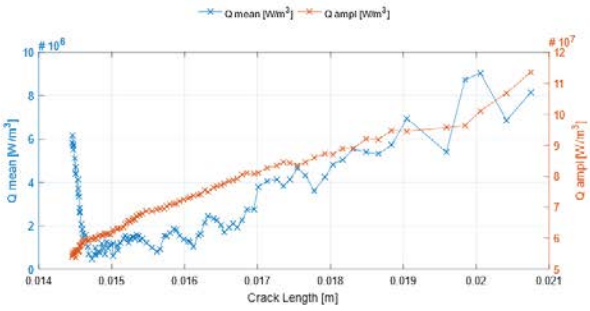
In the double logarithmic scale the relationship isn't anymore linear. This could derive from some errors in the da/dN curve's final phase.

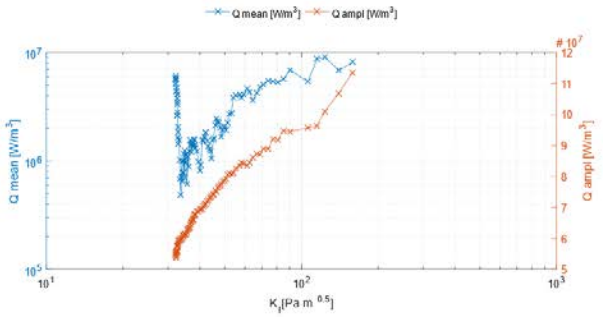


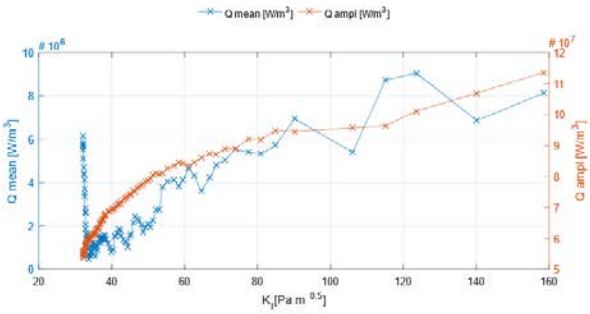


An unexpected peak at the start of the growth is present. One reason could be the influence of the notch present to initiate the crack. The nature of this phenomena could also be in numerical errors in the fitting algorithm. The absolute value of the Q mean is always the same order of magnitude as the mean of the residuals error. A trend could be seen anyway.

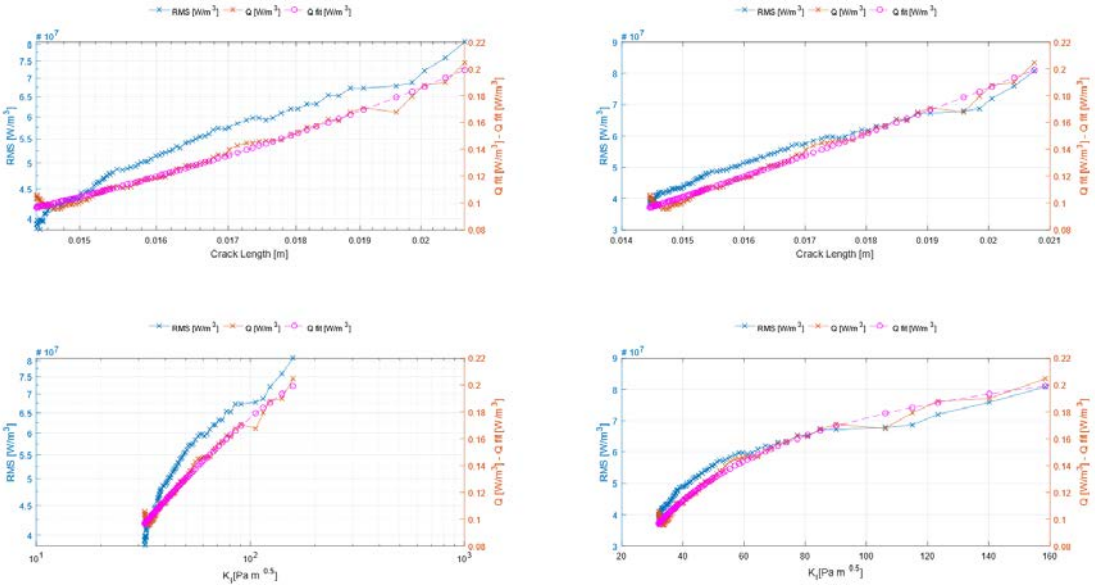








The values of the RMS don't suffer from a peak at the start of the crack growth. A linear relationship can still be observed on the RMS versus ΔK_1.



Comparing all the cases together, a coincidence in the two specimens tested at R=0.1 can be noticed confirming a good repeatability even if one specimen has defects on the surface. Even at a constant ΔK_1 for different R ratios the value for the RMS is very different. The maximum load being the same for every case, an increase in the ratio means increasing the mean of the applied load.

Big errors due to the error on the measurement of the crack length are not present.

At a constant ΔK_1, a very similar RMS was expected even if the loading amplitude is different. That would have correlated energy and stress field directly. An hypothesis for this behaviour could be, according to [14], the thermo-elastic constant's dependence on the effect of the mean stress. A more advanced study on the dependence between these last two parameters should be conduct.

According to this first data, a higher mean stress corresponds to lower heat generation. Still the lower amplitude of the load for higher R ratios can be noticed by a lower RMS. More tests with constant amplitude but different R ratio or constant ΔK_1 trough the specimen should be conduct to explore more the hypothesis.

A change in the R ratio correspond to a different Paris law, which means that different energy processes are involved as a consequence of different microscopic effects. A thermodynamic parameter, like α, could be the link from microstructure to energy and then to fatigue life.

A precise Paris Law could not be find because of a problem with the clip gauge and the crack tip detection algorithm is not so precise for this scope. A prediction on the fatigue life based on thermodynamic data is difficult to perform only with this available data. Moreover, a RMS (or Q amplitude) versus da/dN graph is impossible to draw because of the small accuracy of the a(N) curve.

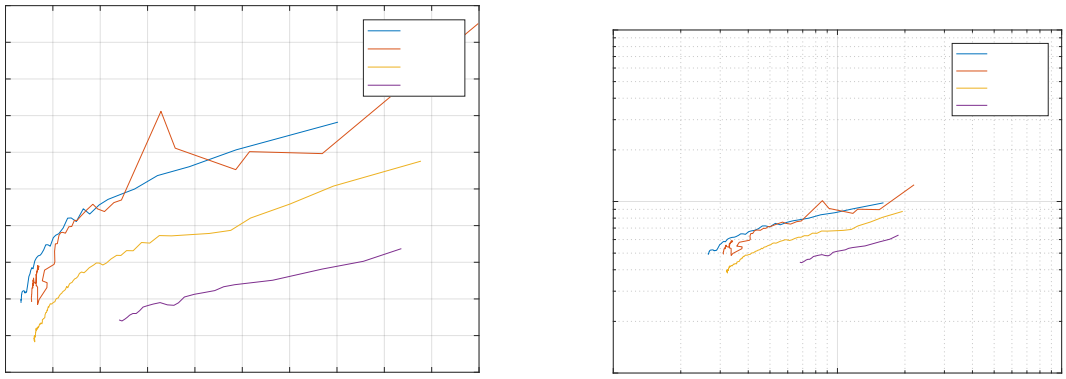


Figure 9-7 Comparisons between different test cases (varying R ratio, maximum load fixed at 6500N). R=0.5 include only the last phase

Similar results have been obtained varying the maximum load and keeping constant the R ratio at R=0.1. Even here similar values for RMS were expected at constant ΔK_1. Because of small variation of the load, the 5.5 kN test is similar to the 6 kN one. However, a different trend than before can be identified: an increase in the mean stress causes an increase in the energy generated.

For varying maximum load the error on ΔK_1 could be more significative because the inaccuracy on the crack length are not compensated by larger differences of R.

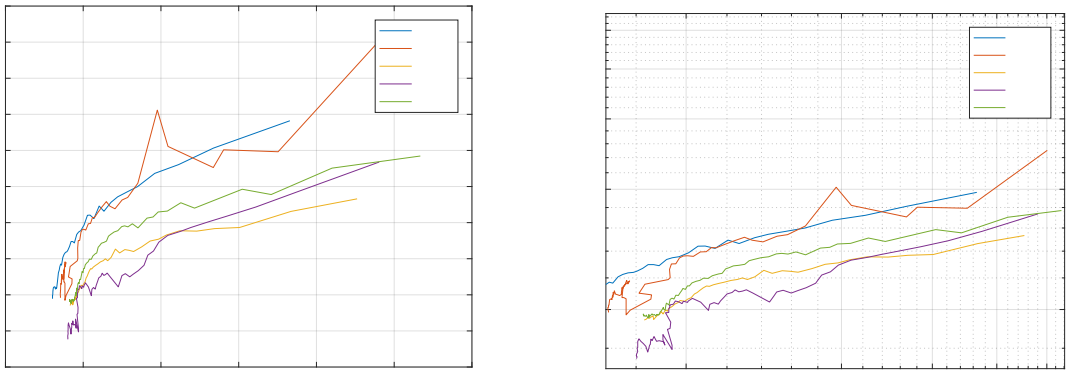


Figure 9-8 Comparisons between different test cases (varying maximum load, ratio R fixed at 0.1)

The mean stress hypotesis can't be the explanation because here, for higher mean load, the RMS of the heat source function is higher. Effects of nonlinearities like density, specific heat and thermal conductivity dependence on temperature could be neglected.

An effect of plasticity could be the explanation but the order of the mean of the heat source function is 10^7 compared to 10^8 of the amplitude. The load amplitude seems to drive the phenomenon and not the first intensity factor.

ΔK_1 is not a parameter that alone can describe the energy behaviour of the crack growth. It would be interesting to study the influence of the mean value for the heat source function but the accuracy has to be improved more. Even if a clear trend can be read, the absolute values can't be trusted so much because are in the range of the normal uncertainties.

A number of new experiments (constant ΔK_1 tests, constant load amplitude and different R) have been proposed to clear some unanswered questions.

On the experimental setup area, some points still can be improved more. The specimen isolated form the environment light and air and more importantly a new way of applying the black paint should be thought to have a more uniform surface compared to literature. In this way the filtering can be reduced and less information be lost.

10 MATERIALS

Temperature measurements are affected by thermal properties of the material: the objective is to find a metal which has a low thermal conductibility so the temperature distribution measured on the surface doesn't smooth out quickly [38].

Metal	Density [kg m^{-3}]	Specific heat [J kg^{-1} $°C^{-1}$]	Conductivity [W m^{-1} $°C^{-1}$]	Diffusivity [10^5 m^2 $°C^{-1}$]
Silver	10500	230	418	17.1
Copper	8940	380	389	11.4
Aluminium	2700	860	200	8.6
Brass	8500	370	100	3.3
Steel (0.1%C)	7850	490	46	1.2
AISI 316L steel	8000	500	15	0.4

The material's thermal diffusivity (ratio of thermal conductivity to volumetric heat capacity) has to be as low as possible so that heat generation is not dissipated too quickly with respect to the frame rate of the IR camera. This is actually essential to be able to capture the thermal field's heterogeneities. Secondly, single phase steel is preferred.

In order to respect as well as possible all the above-mentioned constraints, the chosen material is an AISI 316L austenitic stainless steel since it meets the two first criteria: its diffusivity is low and it is a single phase steel. Moreover, different heating treatments could be performed to change the microstructure and study its influence. For example in [39] different grains sizes has been tested.

10.1 AISI 316L (UNS31603)

Grade 316 is the standard molybdenum-bearing grade, second in importance to 304 amongst the austenitic stainless steels. The molybdenum gives 316 better overall corrosion resistant properties than Grade 304, particularly higher resistance to pitting and crevice corrosion in chloride environments.

Grade 316L, the low carbon version of 316 and is immune from sensitisation (grain boundary carbide precipitation). Thus it is extensively used in heavy gauge welded components (over about 6mm). There is commonly no appreciable price difference between 316 and 316L stainless steel.

The austenitic structure also gives these grades excellent toughness, even down to cryogenic temperatures.

Compared to chromium-nickel austenitic stainless steels, 316L stainless steel offers higher creep, stress to rupture and tensile strength at elevated temperatures.

A possible alternative to 316L is 317L, it has higher resistance to chlorides than 316L, but with similar resistance to stress corrosion cracking.

10.1.1 Composition

Composition ranges for 316L stainless steels.

Grade		C	Mn	Si	P	S	Cr	Mo	Ni	N
316L	Min	-	-	-	-	-	16.0	2.00	10.0	-
	Max	0.03	2.0	0.75	0.045	0.03	18.0	3.00	14.0	0.10

10.1.2 Physical Proprieties

Grade	Density (kg/m^3)	Elastic Modulus (GPa)	Mean Co-eff of Thermal Expansion (µm/m/°C)			Thermal Conductivity (W/m.K)		Specific Heat 0-100°C (J/kg.K)	Elec Resistivity (nΩ.m)
			0-100°C	0-315°C	0-538°C	At 100°C	At 500°C		
316/L/H	8000	193	15.9	16.2	17.5	16.3	21.5	500	740

10.1.3 Stress-strain curve

Based on the work of [44] a non linear model of the steel has been adopted.

The full-range stress-strain curve can be written out in full as follows:

$$\varepsilon = \begin{cases} \dfrac{\sigma}{E_0} + 0.002\left(\dfrac{\sigma}{\sigma_{0.2}}\right)^n & \text{for } \sigma \le \sigma_{0.2} \\ \dfrac{\sigma-\sigma_{0.2}}{E_{0.2}} + \varepsilon_u\left(\dfrac{\sigma-\sigma_{0.2}}{\sigma_u-\sigma_{0.2}}\right)^m + \varepsilon_{0.2} & \text{for } \sigma > \sigma_{0.2} \end{cases}$$

Where:

$$E_{0.2} = \frac{E_0}{1+0.002\,n/e};\ \varepsilon_{0.2} = \frac{\sigma_{0.2}}{E_0} + 0.002; m = 1 + 3.5\frac{\sigma_{0.2}}{\sigma_u}$$

And the coefficient for the AISI 316L are:

E_0 [GPa]	$\sigma_{0.01}$ [MPa]	$\sigma_{0.2}$ [MPa]	σ_u [MPa]	ε_u	e	n	m
190	190	316	616	0.51	0.0017	5.88	2.8

10.1.4 Corrosion Resistance

Excellent in a range of atmospheric environments and many corrosive media (generally more resistant than AISI 304). Subject to pitting and crevice corrosion in warm chloride environments, and to stress corrosion cracking above about 60°C but it's considered resistant to potable water at ambient temperatures, reducing to about and up to 60°C.

316 is usually regarded as the standard “marine grade stainless steel”, but it is not resistant to warm sea water. In many marine environments 316 does exhibit surface corrosion, usually visible as brown staining. This is particularly associated with crevices and rough surface finish.

10.1.5 Plastic Heat Dissipation

Oliferuk [45] measured experimentally the stored energy of 316L steel. When a material deforms plastically, a part of mechanical energy w_p expended in the plastic deformation is converted into heat q_d while the remainder e_s is stored in the material. Thus:

$$e_s = w_p + q_d$$

The stored energy is an essential feature of the cold-worked state and represents the change in internal energy of the material. The measure of energy conversion at each instant of the deformation process is the rate of energy storage:

$$\frac{de_s}{dw_p} = 1 - \beta(w_p) = 1 - \beta(\varepsilon)$$

Where β is the Taylor Quinney coefficient defined as $\beta = \frac{q_d}{w_p}$.

The experiment for determining β [45] has been carried out at a constant deformation rate of $\dot{\varepsilon} = 4.3\ 10^{-3}\ s^{-1}$.

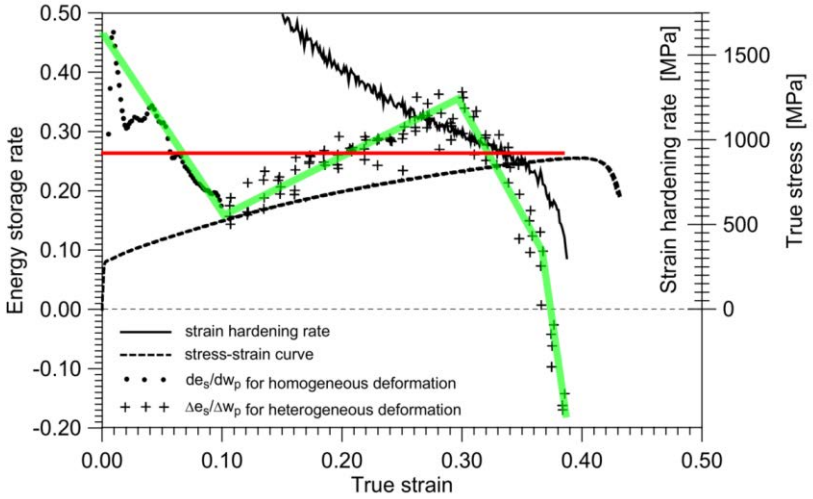


Figure 10-1 One minus Taylor Quinney coefficient for AISI 316L as a function of true strain. It is approximated by the green line, the red line is the mean

In Abaqus β versus true strain proprieties can be implemented.

10.1.6 Heat Resistance

Good oxidation resistance in intermittent service to 870°C and in continuous service to 925°C. Continuous use of 316 in the 425-860°C range is not recommended if subsequent aqueous corrosion resistance is important. Grade 316L is more resistant to carbide precipitation and can be used in the

above temperature range. Grade 316H has higher strength at elevated temperatures and is sometimes used for structural and pressure-containing applications at temperatures above about 500°C.

10.1.7 Heat Treatment

Solution Treatment (Annealing) - Heat to 1010-1120°C and cool rapidly. These grades cannot be hardened by thermal treatment.

10.1.8 Welding

Excellent weldability by all standard fusion and resistance methods, both with and without filler metals. Heavy welded sections in Grade 316 require post-weld annealing for maximum corrosion resistance. This is not required for 316L.

316L stainless steel is not generally weldable using oxyacetylene welding methods.

10.1.9 Machining

316L stainless steel tends to work harden if machined too quickly. For this reason low speeds and constant feed rates are recommended.

316L stainless steel is also easier to machine compared to 316 stainless steel due its lower carbon content.

10.1.10 Hot and Cold Working

316L stainless steel can be hot worked using most common hot working techniques. Optimal hot working temperatures should be in the range 1150-1260°C, and certainly should not be less than 930°C. Post work annealing should be carried out to induce maximum corrosion resistance.

Most common cold working operations such as shearing, drawing and stamping can be performed on 316L stainless steel. Post work annealing should be carried out to remove internal stresses.

10.1.11 Hardening and Work Hardening

316L stainless steel does not harden in response to heat treatments. It can be hardened by cold working, which can also result in increased strength.

10.1.12 Crack Growth Rate

At 24°C the Paris Law coefficients, derived from the graph in absence of a table, are C=1.023e-12 and m=5.052.

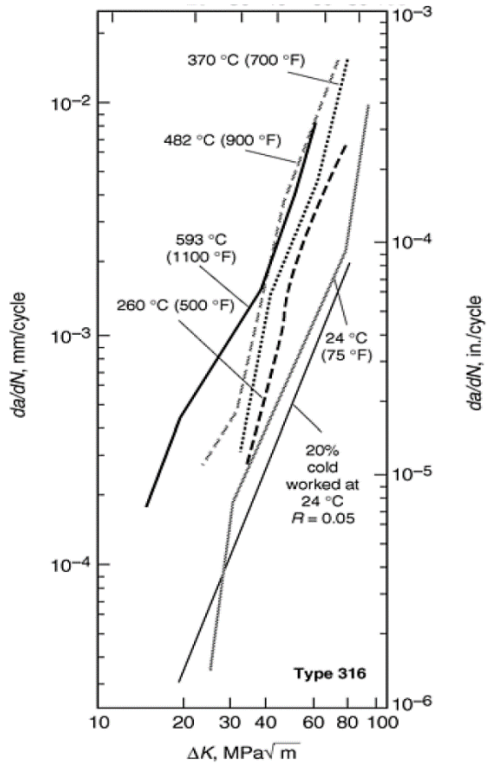


Figure 10-2 Crack Growth Rate for 316L, R=0

10.1.13 Applications

Typical applications include [46]:

- Food preparation equipment particularly in chloride environments.
- Pharmaceuticals
- Marine applications
- Architectural applications
- Medical implants, including pins, screws and orthopaedic implants like total hip and knee replacements
- Fasteners
- Nuclear power plants components

10.1.14 Source of errors and noise in IR-image

During the first stages of an applied mechanical loading, according to their orientation with respect to the loading axis, some grains are submitted to a more important deformation than others, leading to local heterogeneous plasticity even with a homogeneous load (uniform field of stresses). Dissipation and strain fields in a polycristal are not homogeneous so the specimen is submitted to instable and local phenomena.

10.2 OTHER MATERIALS

Other materials have been taken into account for the experiment. One of the possibilities of extending the test is to repeat it at different temperatures up to 500°C (the testing machine has a hoven to heat the specimen).

It's preferable that the materials have this type of data:

- Thermal conductivity at high temperatures
- Stress and strain curve at differente temperature
- A Taylor Quinney coefficient
- Different microstructure deriving from different heat treatment or working

10.2.1 Aluminium

They are not very indicated because their thermal conductivity is very high but it's easily available.

10.2.1.1 *6000 series*

6000 series are alloyed with magnesium and silicon. They are easy to machine, are weldable, and can be precipitation hardened, but not to the high strengths that 2000 and 7000 can reach. 6061 alloy is one of the most commonly used general-purpose aluminium alloys.

The 6061 and 6063 are commonly used in aircraft and aerospace structures, cycling frames, boat building and shipbuilding, and other marine and salt-water sensitive shore applications [47].

Material Name	Thermal Conductivity [W/m-K @ °C range]
Aluminum 6061-O	180 @ 20.0 - 25.0
Aluminum 6061-T4; 6061-T451	154 @ 20.0 - 25.0
Aluminum 6061-T6; 6061-T651	167 @ 20.0 - 25.0
Aluminum 6061-T8	170 @ 20.0 - 25.0
Aluminum 6061-T91	170 @ 20.0 - 25.0
Aluminum 6061-T913	170 @ 20.0 - 25.0
Aluminum 6063-O	218 @ 20.0 - 25.0
Aluminum 6063-T1	193 @ 20.0 - 25.0
Aluminum 6063-T4	200 @ 20.0 - 25.0
Aluminum 6063-T5	209 @ 20.0 - 25.0
Aluminum 6063-T6	200 @ 20.0 - 25.0
Aluminum 6063-T83	200 @ 20.0 - 25.0
Aluminum 6063-T831	200 @ 20.0 - 25.0
Aluminum 6063-T832	200 @ 20.0 - 25.0
Aluminum 6063-T835	200 @ 20.0 - 25.0
Aluminum 6101-H111	220 @ 20.0 - 25.0
Aluminum 6101-T6	218 @ 20.0 - 25.0
Aluminum 6101-T61	222 @ 20.0 - 25.0
Aluminum 6101-T63	218 @ 20.0 - 25.0
Aluminum 6101-T64	226 @ 20.0 - 25.0

Aluminum 6101-T65	218 @ 20.0 - 25.0
Aluminum 6463-O	200 @ 20.0 - 25.0
Aluminum 6463-T1	193 @ 20.0 - 25.0
Aluminum 6463-T4	190 @ 20.0 - 25.0
Aluminum 6463-T5	209 @ 20.0 - 25.0
Aluminum 6463-T6	200 @ 20.0 - 25.0

10.2.1.2 *7000 series*

7000 series are alloyed with zinc, and can be precipitation hardened to the highest strengths of any aluminium alloy (ultimate tensile strength up to 700 MPa for the 7068 alloy).

The 7050 and 7075 (and the 7475 less widely) are commonly used in aircraft and aerospace structures. The 7075 for cycling frames and components.

Material Name	Thermal Conductivity [W/m-K @ °C range]
Aluminum 7039-O	140 @ 20.0 - 25.0
Aluminum 7039-T61	140 @ 20.0 - 25.0
Aluminum 7039-T64	140 @ 20.0 - 25.0
Aluminum 7050-T74	157 @ 20.0 - 25.0
Aluminum 7050-T7451	157 @ 20.0 - 25.0
Aluminum 7050-T7651	153 @ 20.0 - 25.0
Aluminum 7075-O	173 @ 20.0 - 25.0
Aluminum 7075-T6; 7075-T651	130 @ 20.0 - 25.0
Aluminum 7075-T73	155 @ 20.0 - 25.0
Aluminum 7075-T76; 7075-T7651	150 @ 20.0 - 25.0
Aluminum 7178-O	150 @ 20.0 - 25.0
Aluminum 7178-T6; 7178-T651	125 @ 20.0 - 25.0
Aluminum 7475-T61	138 @ 20.0 - 25.0
Aluminum 7475-T651	138 @ 20.0 - 25.0
Aluminum 7475-T7351	163 @ 20.0 - 25.0
Aluminum 7475-T76	147 @ 20.0 - 25.0
Aluminum 7475-T761	147 @ 20.0 - 25.0
Aluminum 7475-T7651	163 @ 20.0 - 25.0

10.2.2 Tianium Alloy

Titanium alloys are metals that contain a mixture of titanium and other chemical elements. Such alloys have very high tensile strength and toughness (even at extreme temperatures). They are light in weight, have extraordinary corrosion resistance and the ability to withstand extreme temperatures. However, the high cost of both raw materials and processing limit their use.

The microstructure influences greatly titanium's alloy proprieties. It can be modified opportunely choosing heating treatments and mechanical processes.

Alloys may be supplied in the following conditions: Grades 5, 23, 24, 25, 29, 35, or 36 annealed or aged; Grades 9, 18, 28, or 38 cold-worked and stress-relieved or annealed; Grades 9, 18, 23, 28, or 29 transformed-beta condition; and Grades 19, 20, or 21 solution-treated or solution-treated and aged.

Ti-6Al-4V cover more than 50% usage of titanium alloys, it's the most used and studied alloy. Pure titanium counts for 20% of the market.

Generally, resistance to crack initiation increase when the grains are bigger. Alpha-beta alloys have an anisotropic behaviour because of a compact hexagonal lattice system.

10.2.2.1 *Alpha alloys and Near Alpha alloys*

The alpha group is formed by alloys including only α stabilising elements or neuter elements. If a small presence of β stabilising elements is present, near-α is the name to be utilized. Alpha group does not respond to heat, instead near-alpha can.

Ti-3Al-2.5V has a medium resistance (50% than pure titanium), cannot be aged, and offer the best resistence to pressure according to ASME Pressure Vessel Code. It also has good weldability and cold forming capabilities. Resistant from mildly reducing to slightly oxidizing solutions.

Ti-5Al-2.5Sn (6 grade) weldable, cannot be aged but with high resistance. It has very good stability and resistance to oxidation and creep at high temperatures (up to 480 °C).

Material Name	Thermal Conductivity [W/m-K @ °C range]
Titanium Ti-3Al-2.5V (Grade 9), alpha annealed	8.30 - 11.8 @ 23.0 - 315
Titanium Ti-3Al-2.5V (Grade 9), ST 925°C, Aged 480°C	8.30 - 11.8 @ 23.0 - 315
Titanium Ti-3Al-2.5V (Grade 9), Alpha-Beta Anneal, Quenched	8.30 - 11.8 @ 23.0 - 315
Titanium Ti-3Al-2.5V (Grade 9), Beta-Anneal 950°C	11.8 @ 23.0 - 315
Titanium Ti-5Al-2.5Sn (Grade 6)	15 7.80 - 10.9 @ 20.0 - 315

10.2.2.2 *Alpha-Beta alloys*

Alpha-beta alloys contain, at room temperature, phase β that oscillates approximately between 5 and 50% due to the presence of α and β stabilizing elements. They are the most used alloys. This group can respond to heat treatment and encounter martensitic type transformations.

Ti-6Al-4V is the most commonly used alloy. It has a chemical composition of 6% aluminium, 4% vanadium, small percentage of iron and oxygen and the remainder titanium. It is significantly stronger than commercially pure titanium while having the same stiffness and thermal properties (excluding thermal conductivity, which is about 60% lower in Grade 5 Ti. Among its many advantages, it is heat treatable. This grade is an excellent combination of strength, corrosion resistance, weld and fabricability. Applications include: blades, discs, rings, airframes, fasteners, vessels, cases, hubs, forgings, biomedical implants.

Ti-4.5Al-3V-2Mo-2Fe has been developed to offer very high superplasticity properties. It has good ductility and fatigue resistance, and greater resistance to Ti-6-4. For aerospace components, golf club heads, metal spheres, work tools, wrist watch cases, automotive components, climbing equipment is formed exploiting superplasticity.

Ti-4Al-4Mo-2Sn-0.5Si has high resistance (1100 MPa of rupture stress) for forging and exhibits good resistance and creep behavior up to 400 °C. It can be aged for thicknesses up to 150 mm. it's used for aeronautical structures and motor components.

Material Name	Thermal Conductivity [W/m-K @ °C range]
Titanium Ti-4.5Al-3V-2Mo-2Fe (SP-700)	7.00 - 10.0 @ 23.0 - 400
Titanium Ti-4Al-4Mo-2Sn-0.5Si (IMI550), STA	7.50 - 9.50 @ 20.0 - 200
Titanium Ti-6Al-4V (Grade 5), annealed	6.70 @ 20 - 25
Titanium Ti-6Al-6V-2Sn (Ti-6-6-2) Annealed	6.60 - 11.9 @ 20.0 - 425
Titanium Ti-6Al-6V-2Sn (Ti-6-6-2) STA 870°C/565°C	6.60 - 11.9 @ 20.0 - 425
Titanium Ti-6Al-6V-2Sn (Ti-6-6-2) STA 910°C/540°C	6.60 - 11.9 @ 20.0 - 425

10.2.2.3 *Beta Alloys*

Beta alloys contains α and β at room temperature, but even after hardening there is no martensitic structure (metastable β alloys) or they contain only phase β (for β single-phase alloys).

Beta alloys can be subjected to thermal treatments but do not undergo martensitic transformations. The choice of thermal treatment influences, in addition to static strength and ductility of the material, also its fracture toughness and fatigue resistance.

Ti-13V-11Cr-3Al, developed in the half of the 20th century, for years has been the only commercially available beta-alloy, while Ti-15Mo-5Zr-3Al can be cold worked, aged and it has high mechanical resistance and to reducing solutions. Moreover, aluminium increase toughness.

They aren't common as the Ti-10V-2Fe-3Al, Ti-3Al-8V-6Cr-4Zr-4Mo (grade 19) and Ti-3Al-8V-6Cr-4Zr-4Mo-0.05Pd (grade 20) but no high temperature data can be found for this ones.

Material Name	Thermal Conductivity [W/m-K @ °C range]
Titanium Ti-13V-11Cr-3Al (Ti-13-11-3) Solution Treated	6.90 - 17.1 @ 20.0 - 425
Titanium Ti-13V-11Cr-3Al (Ti-13-11-3) SolnTreat; Age 450°C	6.90 - 17.1 @ 20.0 - 425
Titanium Ti-13V-11Cr-3Al (Ti-13-11-3) Aged 400°C	6.90 - 17.1 @ 20.0 - 425
Titanium Ti-13V-11Cr-3Al (Ti-13-11-3) Aged 490°C	6.90 - 17.1 @ 20.0 - 425
Titanium Ti-15Mo-5Zr-3Al ST 850°C (1560°F), Aged 500°C	7.50 - 10.5 @ 20.0 - 200
Titanium Ti-15Mo-5Zr-3Al ST 735°C, Aged 500°C	7.50 - 10.5 @ 20.0 - 200
Titanium Ti-15Mo-5Zr-3Al ST 785°C, Aged 500°C	7.50 - 10.5 @ 20.0 - 200
Titanium Ti-15Mo-5Zr-3Al Solution Treated 735°C	7.50 - 10.5 @ 20.0 - 200
Titanium Ti-15Mo-5Zr-3Al Solution Treated 785°C	7.50 - 10.5 @ 20.0 - 200
Titanium Ti-15Mo-5Zr-3Al Duplex Aged	7.50 - 10.5 @ 20.0 – 200

10.2.3 Others Stainless Steel

Other types of Stainless Steel have been considered because of the low thermal conductivity and of its data availably at high temperatures. Moreover, the possibility of having mechanical and thermal data

for different heat treatments or cold processing is considered a plus to study the influence of the different microstructure on fatigue.

10.2.3.1 *AISI 300 Series*

300 Series stainless steels are classified as austenitic, and are hardenable only by cold working methods. These grades of stainless have chromium (approx. 18 to 30%) and nickel (approx. 6 to 20%) as their major alloying additions. 300 Series Stainless steel alloys resist corrosion, maintain their strength at high temperatures and are easy to maintain.

302 Stainless Steel is non-magnetic, extremely tough and ductile. It is one of the more common chrome-nickel stainless and heat-resisting steels. Cold working will dramatically increase its hardness, and applications range from the stamping, spinning and wire forming industry to food and beverage, sanitary, cryogenic and pressure-containing. It is also formed into all types of washers, springs, screens and cables.

Type 304 is the most widely used alloy of all stainless steels. This non-magnetic alloy is the most versatile and the most widely used of all stainless steels. 304 Stainless Steel has lower carbon to minimize carbide precipitation and is used in high-temperature applications. It's commonly used to process equipment in the mining, chemical, cryogenic, food, dairy and pharmaceutical industries. Its resistance to corrosive acids also makes 304 Stainless Steel ideal for cookware, appliances, sinks and tabletops.

316 Stainless Steel (the second most common grade) is recommended for welding because it has a carbon content lower than 302 to avoid carbide precipitation in welding applications. The addition of molybdenum and a slightly higher nickel content make 316 Stainless Steel suitable for architectural applications in severe settings, from polluted marine environments to areas with sub-zero temperatures. Equipment in the chemical, food, paper, mining, pharmaceutical and petroleum industries often includes 316 Stainless Steel.

Material Name	Thermal Conductivity [W/m-K @ °C range]
302 Stainless Steel (from 0% to 60% cold reduction)	16.2 - 21.5 @ 100 - 500
302 Stainless Steel, annealed	16.2 - 21.5 @ 0.000 - 500
304 Stainless Steel	16.2 - 21.5 @ 0.000 - 500
304 Stainless Steel, annealed	16.2 - 21.5 @ 0.000 - 500
305 Stainless Steel	16.3 - 21.5 @ 100 - 500
305 Stainless Steel, annealed	16.3 - 21.5 @ 100 - 500
305 Stainless Steel (from 0% to 60% cold reduction)	16.3 - 21.5 @ 100 - 500
316 Stainless Steel, annealed	16.3 @ 100
316L Stainless Steel, annealed	14.0 - 15.9 @ 20.0 - 25.

Moreover, for 304L, like 316L, the Taylor Quinney coefficient depending on true strain can be found in literature [45].

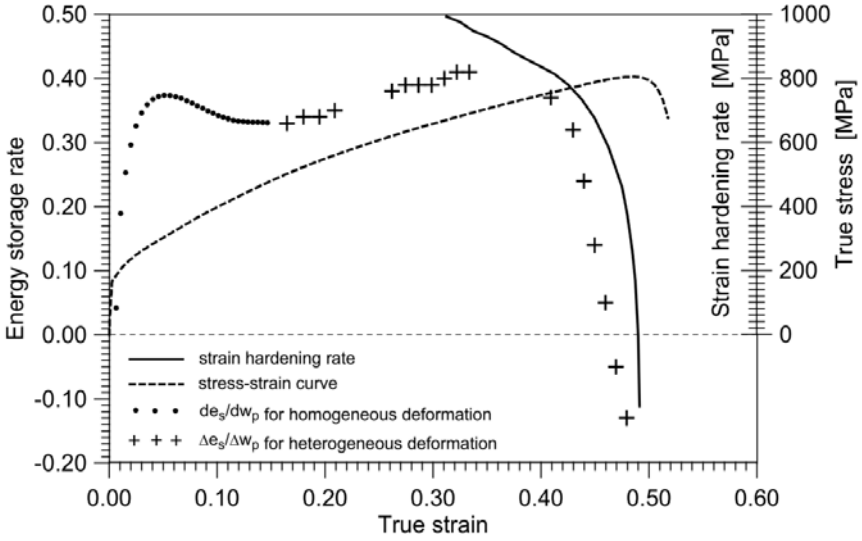


Figure 10-3 Energy Storage rate (1- β) for AISI 304L

10.2.3.2 *AISI 400 Series*

Ferritic and martensitic chromium alloys. The 400 series is susceptible to rust and corrosion under some conditions. Heat-treating will harden the 400 series.

Grade 410 stainless steels are general-purpose martensitic stainless steels containing 11.5% chromium, which provide good corrosion resistance properties. However, the corrosion resistance of grade 410 steels can be further enhanced by a series of processes such as hardening, tempering and polishing. Quenching and tempering can harden grade 410 steels. They are generally used for applications involving mild corrosion, heat resistance and high strength.

AISI 414 is not very used but we can find thermal data at high temperatures.

Grade 416 steel is a free-machining stainless steel with a machinability of 85%, highest of all stainless steels. But this addition of sulphur reduces the formability, weldability and corrosion resistance of 416 steels to below that of grade 410. Because of their high machinability and low cost, grade 416 steels are available in highly tempered, hardened or unhardened forms.

Stainless steel 422 is an appealing alternative to Stainless 403 because of the better strength at higher temperatures. It offers ductility and strength from room temperature to 649°C and resists scaling and oxidation to 760°C.

Material Name	Thermal Conductivity [W/m-K @ °C range]
410 Stainless Steel	24.9 - 28.7 @ 100 - 500
410 Stainless Steel, tempered at 233°C	24.9 - 28.7 @ 100 - 500
410 Stainless Steel, tempered at 343°C	24.9 - 28.7 @ 100 - 500
410 Stainless Steel, tempered at 453°C	24.9 - 28.7 @ 100 – 500
410 Stainless Steel, tempered at 508°C	24.9 - 28.7 @ 100 – 500
410 Stainless Steel, tempered at 568°C	24.9 - 28.7 @ 100 – 500
410 Stainless Steel, tempered at 623°C	24.9 - 28.7 @ 100 – 500

410 Stainless Steel, tempered at 678°C	24.9 - 28.7 @ 100 – 500
410 Stainless Steel, tempered at 733°C	24.9 - 28.7 @ 100 – 500
410 Stainless Steel, tempered at 650°C	24.9 - 28.7 @ 100 – 500
414 Stainless Steel	24.9 - 28.7 @ 100 – 500
414 Stainless Steel, tempered at 233°C	24.9 - 28.7 @ 100 – 500
414 Stainless Steel, tempered at 343°C	24.9 - 28.7 @ 100 – 500
414 Stainless Steel, tempered at 453°C	24.9 - 28.7 @ 100 – 500
414 Stainless Steel, tempered at 508°C	24.9 - 28.7 @ 100 – 500
414 Stainless Steel, tempered at 568°C	24.9 - 28.7 @ 100 – 500
414 Stainless Steel, tempered at 623°C	24.9 - 28.7 @ 100 – 500
414 Stainless Steel, tempered at 678°C	24.9 - 28.7 @ 100 – 500
414 Stainless Steel, tempered at 733°C	24.9 - 28.7 @ 100 – 500
416 Stainless Steel	24.9 - 28.7 @ 100 - 500
422 Stainless Steel 315°C (600°F) temper	23.9 - 27.3 @ 100 - 500
422 Stainless Steel 425°C (800°F) temper	23.9 - 27.3 @ 100 - 500
422 Stainless Steel 480°C (900°F) temper	23.9 - 27.3 @ 100 - 500
422 Stainless Steel 540°C (1000°F) temper	23.9 - 27.3 @ 100 - 500
422 Stainless Steel 650°C (1200°F) temper	23.9 - 27.3 @ 100 - 500
422 Stainless Steel 370°C temper	23.9 - 27.3 @ 100 - 500
422 Stainless Steel 525°C (980°F) temper	23.9 - 27.3 @ 100 - 500
422 Stainless Steel 650°C double temper	23.9 - 27.3 @ 100 - 500
422 Stainless Steel 675°C double temper	23.9 - 27.3 @ 100 - 500
422 Stainless Steel 705°C double temper	23.9 - 27.3 @ 100 - 500
422 Stainless Steel 730°C double temper	23.9 - 27.3 @ 100 - 500
422 Stainless Steel 760°C temper	23.9 - 27.3 @ 100 - 500

10.3 STAINLESS STEEL METALLOGRAPHY

In order to check for an anisotropy of the steel plate from which the specimen will be made, a microscopy is necessary. Small pieces are taken from the plate from which the Ct specimen have been produced.

To examine grain boundary of the metallic surface, it is first necessary to grind and polish the surface to a mirror-like finish. Then, the surface must be etched.

10.3.1 Grinding

The purpose of the grinding step is to remove damage from cutting, planarize the specimen, and to remove material approaching the area of interest.

The most common metallographic abrasive used is Silicon Carbide – SiC. It is an ideal abrasive for grinding because of its hardness and sharp edges. For metallographic preparation, SiC abrasives are used in coated abrasive grinding papers ranging from very coarse 60 grit to very fine 1200 grit sizes.

10.3.2 Polishing

Polishing is the most important step in preparing a specimen for microstructural analysis. It is the step which is required to completely eliminate previous damage.

Ideally the amount of damage produced during cutting and grinding was minimized through proper blade and abrasive grinding so that polishing can be minimized.

To remove deformation from fine grinding and obtain a surface that is highly reflective, the specimens must be polished before they can be examined under the microscope. Polishing is a complex activity in which factors such as quality and suitability for the cloth, abrasive, polishing pressure, polishing speed and duration need to be taken into account. The quality of the surface obtained after the final polishing depends on all these factors and the finish of the surface on completion of each of the previous stages.

Polishing Cloths are used in this process and final polishing abrasives are selected based upon specimen hardness and chemical reactivity. The most common polishing abrasives is alumina. Alumina abrasives are primarily used as mechanical abrasives because of their high hardness and durability.

10.3.3 Etching

The principle of all etchants is to attack the crystal at crystal grain boundaries, the specific choice of etchant will depend on the nature of the material. The attack produces a faceted surface and the orientation of these facets will vary from grain to grain. As a result, the amount of reflected light reflected into the objective lens of a microscope will vary between grains and grain boundaries.

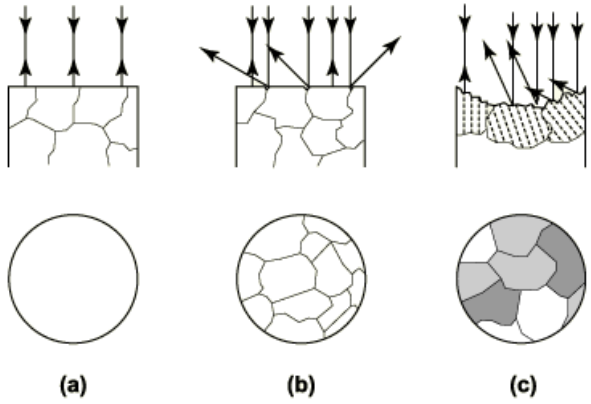


Figure 10-4(a) exhibits a polished metal surface (unetched) with no microstructural detail revealed, Figure 4(b) shows an etched surface and reveals grain boundaries, while Figure 4(c) reveals grain shading and grain boundaries as a result of further etching. In this last example, the amount of reflected light captured by the objective lens of the microscope will vary from one grain to another. This will lead to some grains appearing to be light in colour and others appearing dark – even though all the grains may be identical in composition and type.

Three different types of etching liquid, that can be used on stainless steel, has been found in literature:

- Adlers 300 Series Stainless Steel Etchant is a very effective etchant for 300 series, austenitic, duplex stainless steels. The surface has to be immerse for several seconds. The chemical composition is:

    - Ferric chloride, 45 g
    - Copper ammonium chloride, 9 g
    - Hydrochloric acid, 150 mL
    - Distilled water, 75 mL
- Carpenter 300 Series Stainless Steel Etchant is a nice etchant for 300 series, austenitic, duplex stainless steels. The surface has to be immerse for several seconds. The chemical composition is:
    - Ferric chloride, 8.5 g
    - Cupric chloride, 2.4 g
    - Alcohol, 122mL
    - Hydrochloric acid, 122 mL
    - Nitric acid, 6 mL
- Third alternative if Ferric chloride is not available:
    - Hydrochloric acid, 15 mL
    - Nitric acid, 5 mL
    - Distilled water, 100 mL
- Stronger variant of third solution, few second are enough, the reaction is esothermic. The chemical composition is the following:
    - Hydrochloric acid, 30 mL
    - Nitric acid, 30 mL
    - Distilled water, 30 mL

After application wash off with ethanol and water.

11 IMAGE ACQUISITION AND PROCESSING FRAMEWORK

11.1 ACQUISITION OF IMAGES IN MATLAB

Two main script exist for importing the frames as matrix. One is made for .ASC files (Flir's camera files) and the other for .HCC files (Telops file format). The latter depends on TelopsToolbox v12 and especially on the function readIRCam. Both the files depends on the external function natsortfiles written by Stephen Cobeldick.

Given a folder, all the files found with the specific extension will be converted to matrix and saved in .mat files in the same folder. These files will after be opened to be processed.

The structure of the matrix is A(y,x,1,t) where 1 represent the first colour channel in case a multispectral image can be took.

11.2 IMAGE PROCESSING

All the image processing is done in MATLAB, but the script can be linked also to FIJI to use some of its functionalities. Fiji is an image processing package—a "batteries-included" distribution of ImageJ, bundling a lot of plugins which facilitate scientific image analysis. Like ImageJ itself, Fiji is an open source project.

Figure 11-1 FIJI logo

If options.fiji is set to 1 then the program will be abilitated and the folder added to the MATLAB path. Because of memory limitation of Java the frames has to be processed one by one and not in batch mode, for this reason after every frame the command MIJ.run('Close') has to be executed.

11.2.1 Folder Structure

The three main files needed are in the principal directory. The other will be automatically added when MATLAB needs them. Inside Fiji.app is present the FIJI software that contains also the library to communicate with MATLAB. In Filters are present all the filtering function that aren't implemented

nativilty, in the same manner in the Other folder are present all the function needed that had to be written for this specific application.

11.2.2 Parameters needed for calculation

The user must set the material proprieties: density, heat capacity and heat conductivity (options.mat.ro, options.mat.c and options.mat.k respectivilty) and the thickness of the specimen (options.mat.thck) all in SI units.

Also the camera parameters are needed: pixel size (options.camera.dy), time interval between frames (options.dt=1/freq) and it's best to specify also the horizontal and vertical pixels dimensions.

Also the load frequency of the experiment is needed if one wants to fix this value in the heat function's curve fitting.

11.2.3 Parameter needed for Defective Pixel Removal

The first parameter to be defined is the threshold for defective pixels. It has to be chosen the maximum number of standard deviation of the pixels for every area (for example options.BAD.sigma=3 to keep only the pixel in the range $T_{mean} \pm 3\sigma$). The size (in pixel) of the square area in which calculate mean and standard deviation is specified in options.BAD.boxSize, it has not to be too small to not increase the computational time too much. The mean and standard deviation is not calculated on the entire frame because of the notch area.

The size of the small area, in which calculate the mean to be substituted in the defective pixel, is defined by options.BAD.width.

The possibility to plot a matrix in which the bad pixels, of the first frame of the set, are highlighted is implemented if options.BAD.plot is set to 1 for debugging porpuses.

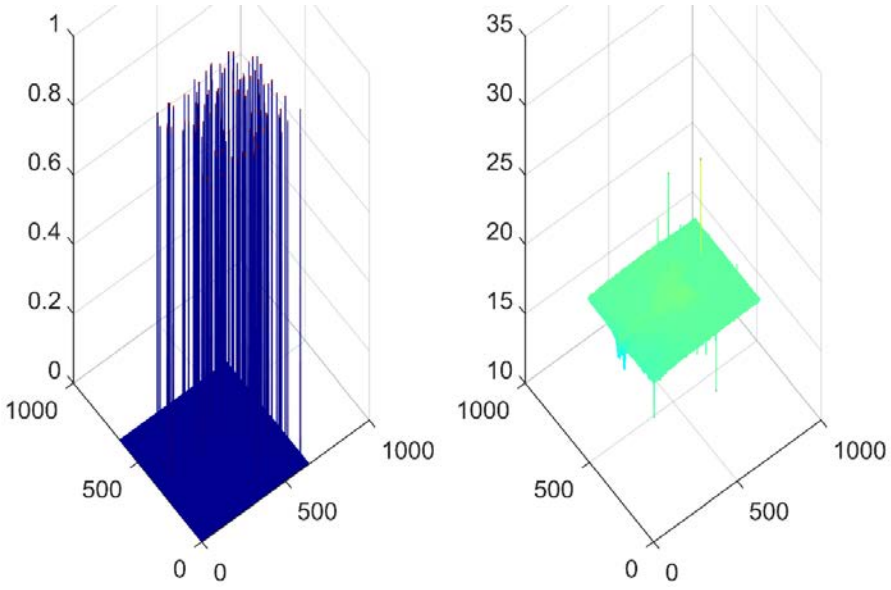


Figura 11-1 On the left the defective pixels assume value of 1, on the right the 3D plot of the image's temperature still in raw data

11.2.4 Parameters needed for fitting the Heat Function

All the parameters needed for the fitting are inside the structured vector options.FIT. This values will be read with the function createFit inside the folder Other. The structure of the options' vector is the following:

- excl: time (in seconds) after which start the fitting, as an example 4*options.dt, it's not very useful.
- freq: logical value, if 0 the frequency is a variable to be fitted, if 1 the value will be the one specified in options.LoadFreq.
- a_0: logical value, if 0 the mean value of the fitting function will be free, if 1 will be set to zero.
- outliers: logical value, if 0 all the point will be kept, if 1 the extreme will be excluded.
    - minTR: number from 0 to 1, minimum of the normalized range of values.
    - maxTR: number from 0 to 1, maximum of the normalized range of values.
- type: 'fourier1' or 'fourier8', chose if the fourier series has to be trouncated to the first term (to obtain a curve like $a_0 + Asin(\omega t + \phi)$) or to the eigth term to approximate better a triangular wave load function. In the latter case the amplitude A and the Qmax derived from fitting has no meaning because of how they have been implemented in matlab. Only the root mean square (RMS) has significance.

  Considering the Fourier series

$$Q(t) = a_0 + \sum_{n=1}^{\infty} (a_n \cos(\mathrm{n} \cdot \mathrm{wt}) + b_n \sin(\mathrm{n} \cdot \mathrm{wt}))$$

  The RMS is defined as:

$$RMS = \sqrt{{a_0}^2 + \frac{1}{2}\left(\sum_{n=1}^{N} {a_n}^2 + \sum_{n=1}^{N} {b_n}^2\right)}$$

11.2.5 Image cropping

Especially for the wavelet filter, it's necessary to remove from the image the notch area. With options.b and options.h are defined the dimension in pixel of the area. To remove 40 pixels on the horizontal dimension it's necessary to specifiy options.b=640-40 (assuming the horizontal pixel length is 640 pixel) and then specify options.Ox=options.pixelX/2+(options.pixelX-options.b)/2 to remove the pixel on the left part of the image. With options.Oy the center of the cropping area is defined. Options.Ox can, of course be a different value, if one want to choose manually the center of the cropping area.

If the latter parameter is defined as options.Ox=[], the cropping area is centered both in x and y. Moreover if one region of the cropping area go outside the original image then the cropped region will automatically be limited in the original data only.

11.2.6 Cropping of the Crack Tip area

If options.cropHT is set to 1 the integration of the heat source would be performed only in a small area and not on the entire specimen. The dimensions of the area are defined with options.bHT and

options.hHT, also an offset in the x direction (positive if in the right way) can be specified with options.HT.offX.

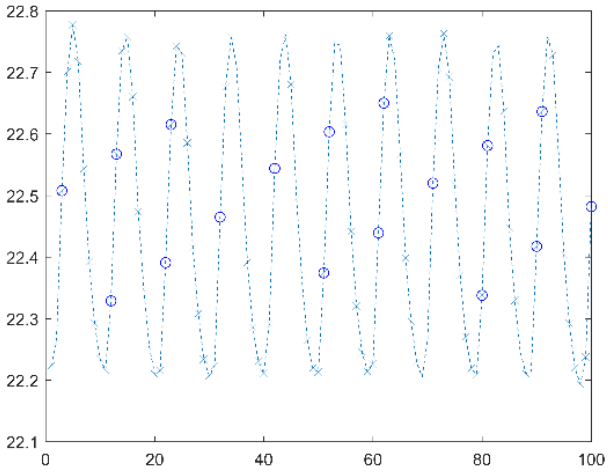
The dimensions of the crack tip area can be increased as the crack growth, the parameter is options.HT.incr and it is set to 1 if one want to maintain the area constant. In the opposite case this value will be the parameter of the law that can be specified in the script after the image processing.

Some parameters are dedicated to the algorithm to find the position of the crack tip for every set of data (recording correspondent to a number of cycles).

- options.HT.trshld: value beween 0 and 1. It's the percentile of the maximum temperature of the frame considered to calculate the instantaneous position of the crack tip. We suggest to keep 1, so only the maximum value is considered. If it was 0.99, the position of all the pixels with temperature values higher than 0.99 of the maxiumum will be searched and then the mean calculated.
- options.HT.minTR and options.HT.maxTR: values between 0 and 1. They are the normalized values of the maximum temperature of all the frames, only the points in the range chosen will be kept to calc the crack tip position.
- options.HT.thr_x and options.HT.thr_y: values between 0 and 1. Normalized range in the spatial domain from which exclude outliers. For example if options.HT.thr_x=0.99, all the point which position is outsie of the range (1-0.99)-0.99 will be discarded.

For debugging purposes some visualization capabilities are also implemented.

- options.HT.plot: logical value. If set to 1 the maximum temperature for every frame is plotted. The points that are kept to calculate the final crack position are highlighted.



- options.crackTip: logical value. If set to 1 when using the function plotData in the form plotData(data,data_w,options,image) every crack tip point used for the calculation will be plotted with his frame number next to it. The point can so be linked to the previous graph to calibrate the threshold values.

If there is more than one set of data points with options.set then a precise set number can be chosen. For this set the crack tip integration area is plotted both in plotData(data,data_w,options,image) and also in the function movie(original,filtered,options,data). If also options.crackTip=1 in plotData(data,data_w,options,image) all the crack tip point non excluded will we plotted.

11.3 FUNCTIONS USED

11.3.1 plotImage

Only if options.plot is set to 1 the figures are created. This can be done only if the sets of data are small otherwise the program will crash because of the too many figures. The function cannot be called after the execution of the START.m script because all the original and filtered recordings (both filtered with conventional methods or wavelets) are deleted to save space on RAM and hard drive.

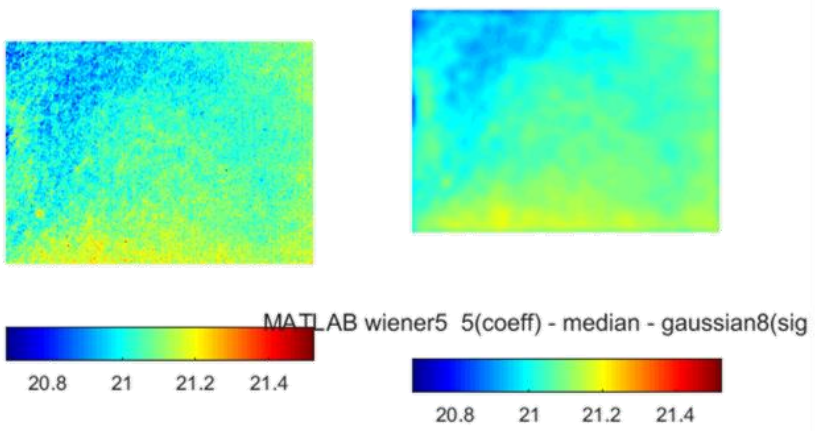


Figure 11-2 The first figure is a comparison between a raw frame and the filtered version

The limit for the colormap are taken from options.minT and options.maxT which are respectively the minimum and maximum values of the raw data. The function has as a input plotImage(original,filtered,options,t_plot) where original and filtered are the 4D matrix of the temperature recording, respectively on the left and right of all the figures. t_plot is a number indicating which frame to plot, if not specified the last frame will be plotted by default. Even if options is not specified the range of temperatures plotted will be between 21°C and 23°C and no description on the method will be displayed.

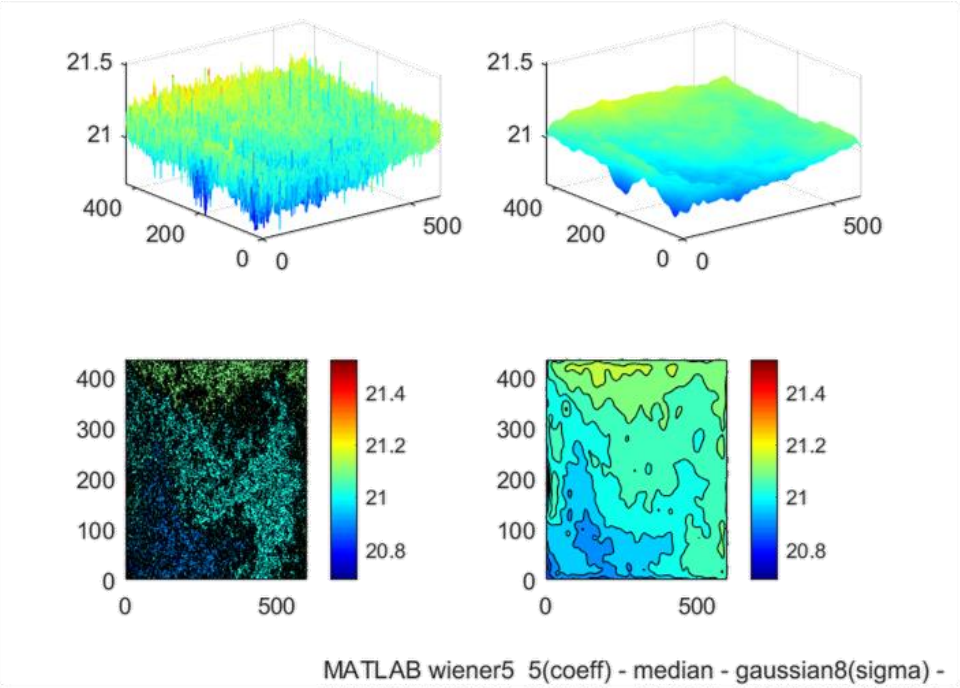


Figure 11-3 The second figure is a meshed surface of the level of temperature (above) and a counter plot of the same distribution (below)

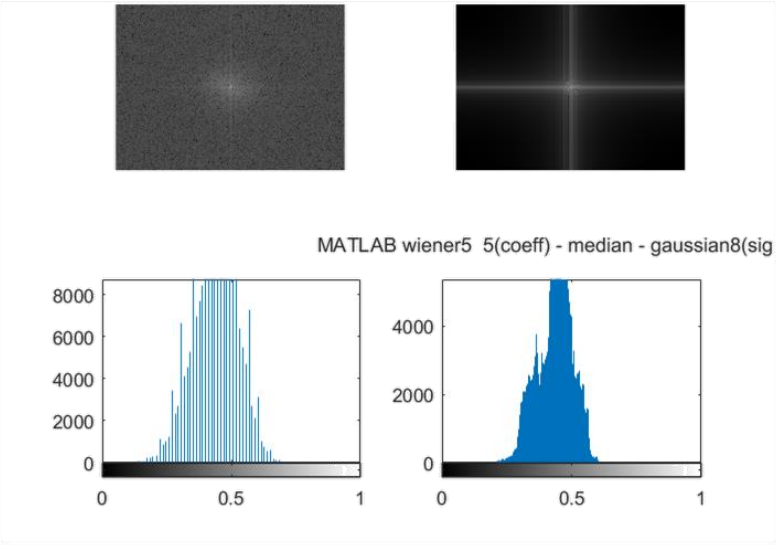


Figure 11-4 The third figure is the spectrum of the orginal and filtered image. Below a color histogram of the color (temperature) distribution of the frame is plotted

If options.plot is set to 1, every figure will be saved in a .png file in the main folder.

11.3.2 tempVStime

The function is called by tempVStime(original,filtered,options), all the parameters are necessary. For the pixel position defined by options.pX and options.pY (if the image is cropped to eliminate the notch

area, as it should, there is a new coordinate system) the temperature versus time (in second) is plotted. The lines with circles and crosses are the maximum temperature value for every frame (raw data if green circle, filtered value if red cross).

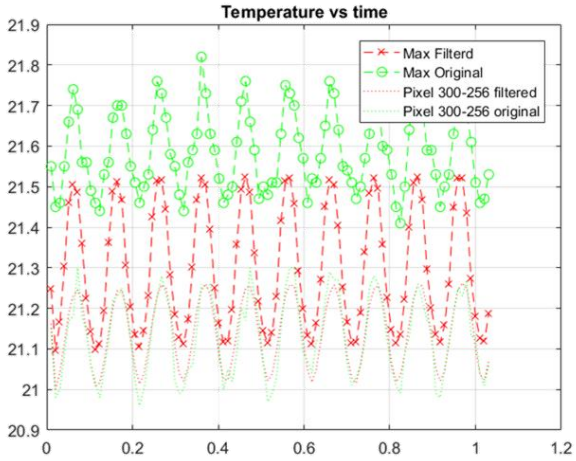


Figure 11-5 All the green lines indicate raw data, the red lines the filtered ones

The effect of filtering is clear especially on the maximum values for every frame.

11.3.3 plotData

Two modalities of use of the functions exists. The first one with plotData(data,data_w,options,original) plots a frame from the set of image 'original' and the evolution of the crack tip's position. If the integration is performed only on the crack tip area (options.crackTip=1) then the area of integration is drawn with a green rectangle. The number of set of which draw the area is defined by options.set='# of set'.

For this set, also, the position of the crack tip for every frame is plotted. Black points correspond to conventional filtering, blue dots to wavelet filtering. If options.plotCrackFrame is set to 1 then also the correspondent number of frame for every black dot (conventional filtering) is displayed.

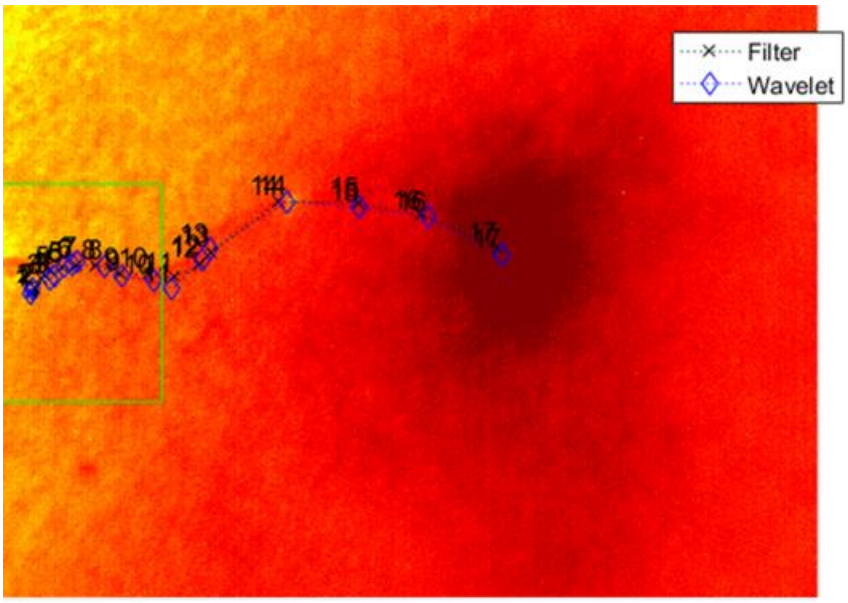


Figure 11-6 Example of result from plotData(data,data_w,options,original)

The second modality is called by the command plotData(data,data_w,options) where data and data_w are structured vectors saved in the .mat file after a complete analysis. Data contains the results from conventional filtering instead data_w the results from wavelet filtering.

Six figures are generated from this function. If options.crackTip is set to 1 more data concerning only the process zone is plotted.

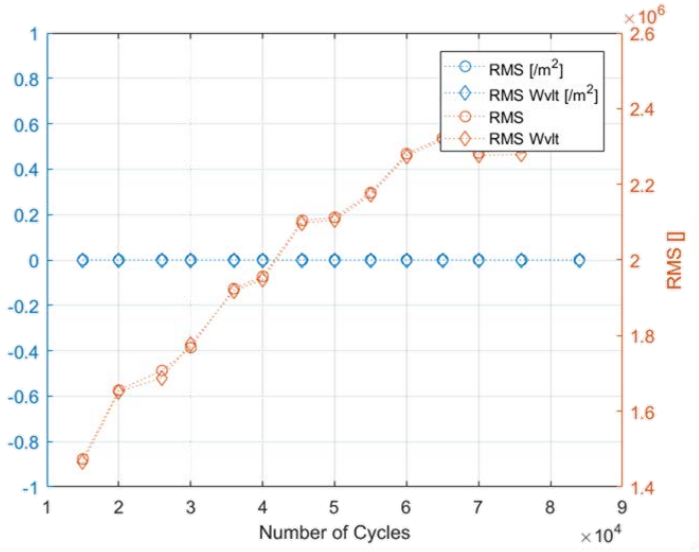


Figure 11-7 The first figure is the plot of the Root Mean Square of the heat source function's fit versus number of cycles. In blue, on the axis on the left, is plotted the value is divided by the area of integration zone of the process zone (zero if the integration is performed on the whole area). In orange, on axis on the right, the RMS is plotted. These values are calculated on the entire cropped image or on the process zone if options.crackTip is set to 1

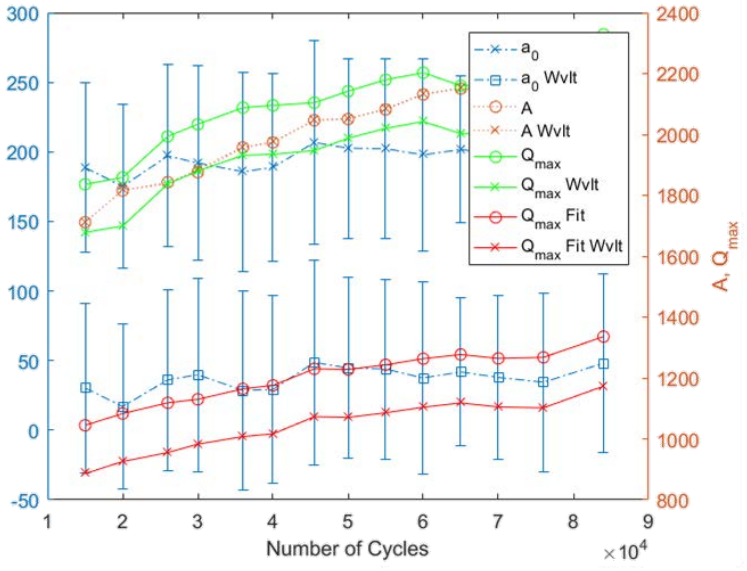


Figure 11-8 On the left axis (blue data) are plotted the mean values of the serie (a_0). On the right axis is plotted the amplitude A, the value of Qmax from raw data and the values of Qmax calculated from the coefficients of the fitting function. if a fourier series of 8 terms is used the amplitude A has no more significance. Crossed points are values derived from wavelet filtering, instead circle points are from conventional filtering. If options.crackTip=1 then all these values are refered to the process zone only

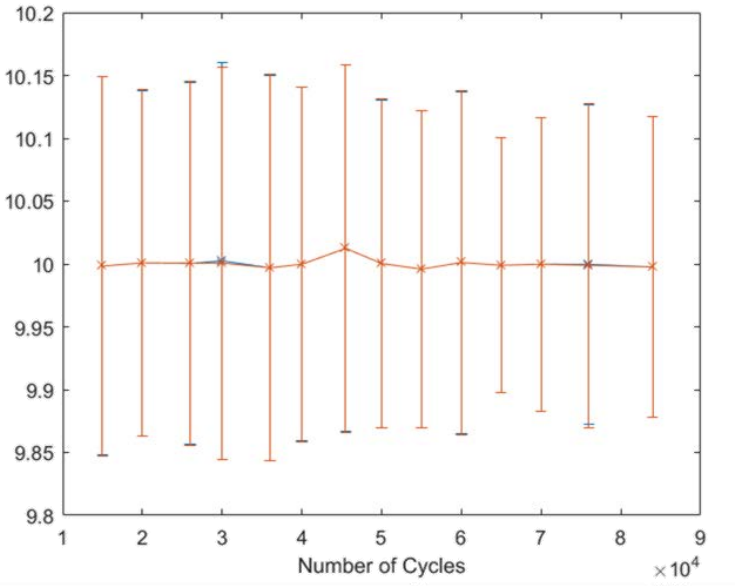


Figure 11-9 The third figure is the plot of frequency [Hz] versus number of cycles

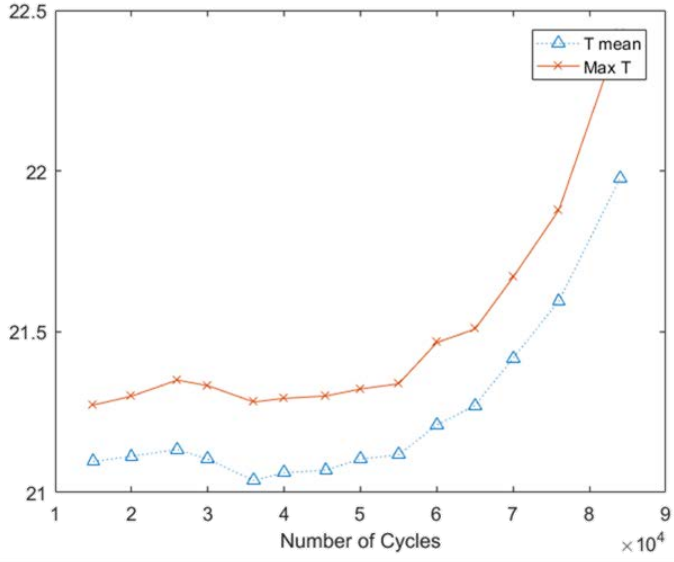


Figure 11-10 the fourth figure is the plot of temperature [°C] versus number of cycles. If option.crackTip=1 then also values for the process zone will be plotted

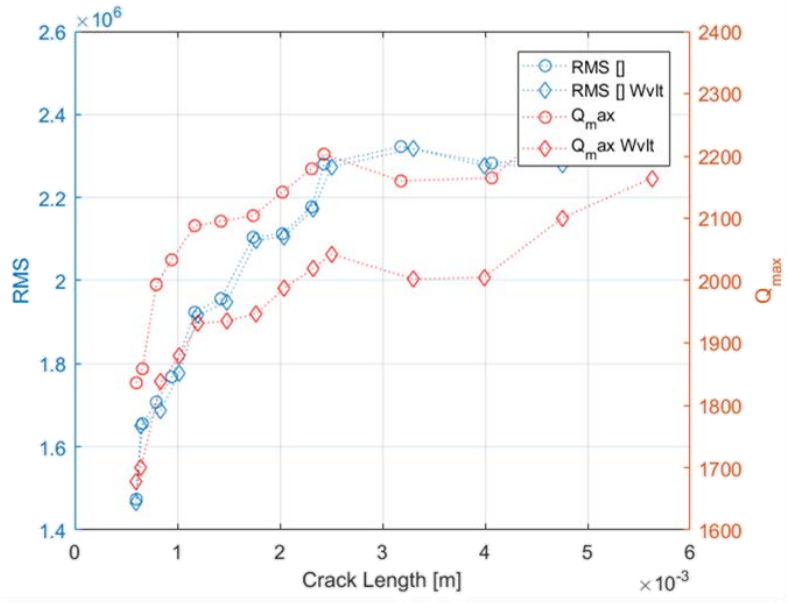


Figure 11-11 The fifth figure shows values for Root Mean Square and of Qmax versus the crack length calculated by the images

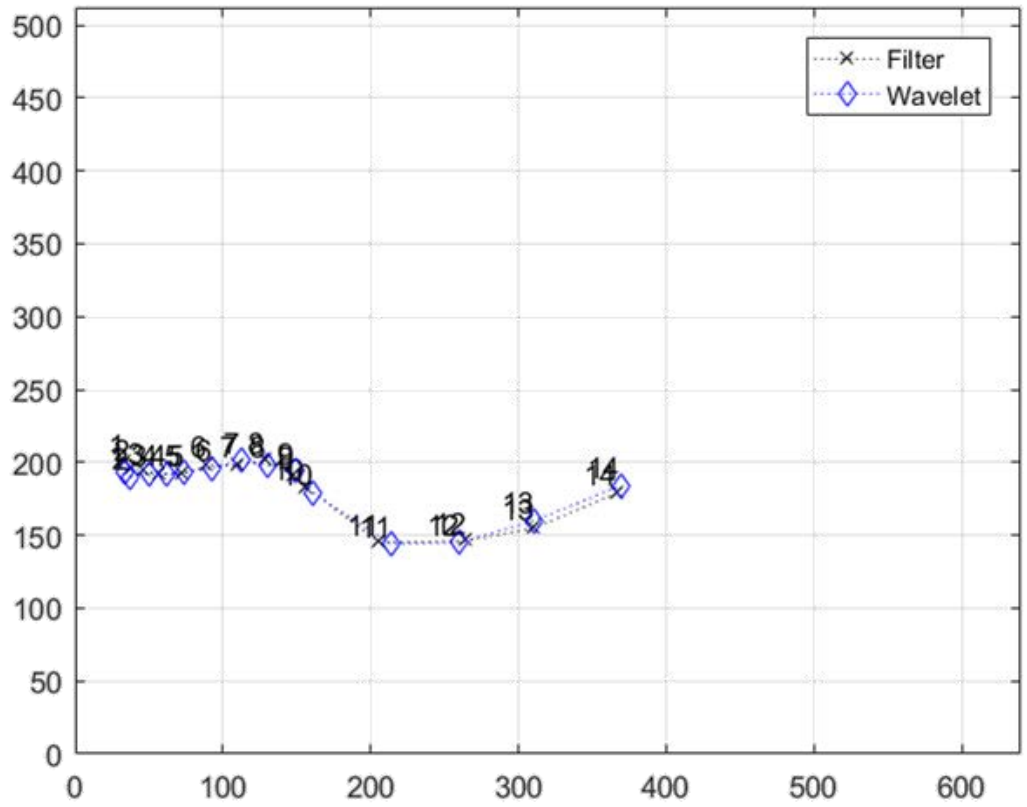


Figure 11-12 The sixth and last figure shows the position of the crack tip for every set of data. The values on the horizontal and veritical axis are pixels positions of the cropped image

11.3.4 movie

The function is called by movie(original,filtered,options,data). The variable data is not necessary, if present and options.crackTip is set to 1 the integration area is displayed in a black rectangle. The number of the set from which the integration area is plotted is defined by options.set='# of set'.

11.3.5 plotFour

The function accepts two structured variables as inputs, the structured vectors ToPlot and Text. All the variable in Text are strings, none of them are compulsory but it's recommended to use it to avoid confusion. The name of figure is specified in Text.name, the name of the two horizontal axis are defined in Text.X1 and Text.X2.

The vector data for horizontal axis are specified with ToPlot.X1 and ToPlot.X2. The Y values are specified with ToPlot.Y1, ToPlot.Y2 and so on. Obviously, the dimensions must agree and for every X point one Y point has to exist.

Y1	*log*	Y2	Y1		Y2
Y5	*log*	Y6	Y5		Y6
X1			**X1**		
Y3	*log*	Y4	Y3		Y4

Y7	*log*	Y8	Y7		Y8
X2			**X2**		

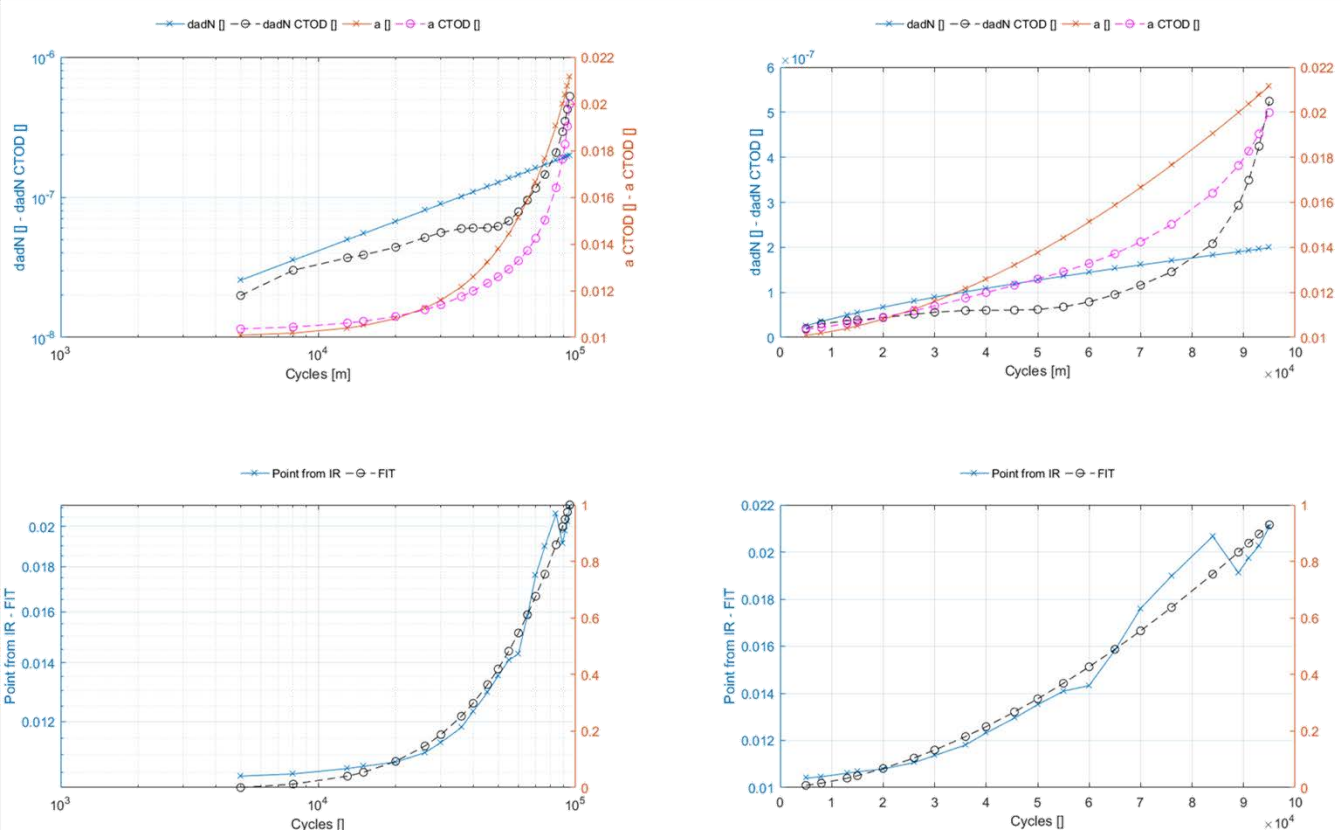


Figure 11-13 Example of plot deriving from plotFour(ToPlot,Text)

Not every Y data (from 1 to 8) is mandatory. On the upper part of the graphs all the points with X1 as abscise are plotted while on the lower part an other abscise is defined. The graphs on the left are the same of the ones on the right but they use a double logarithmic scale.

11.3.6 fitFinalModel

This script generates the final graphs of the analysed data using the plotFour function.

11.3.7 INSTRON_start

This script converts the .csv file coming from the log of the INSTRON fatigue machine to a .mat file easily readable by Matlab. Graphical picking of the file has been utilized.

11.4 FUNCTIONS IMPLEMENTED BUT NOT USED

11.4.1 tmpl_match2

This function has to be inserted inside a loop to be applied for all the recording. In fact the function operates only on two consecutives frames.

The inputs are (first,second,options) where second is the image to be motion compensated and first is the reference (actually the the template will be taken from the first frame and the reference will be the second image but the results are the same). The outputs are the second frame translated (variable new) and a matrix containing the values of the offset in pixels (variable pad).

11.5 EXTERNAL FUNCTIONS

11.5.1 Kalman_Stack_Filters(imageStack,gain,percentvar)

Implements a predictive Kalman like filter in the time domain. It has been created by C.P. Mauer and modified to analyse the new data structure.

11.5.2 Neigh SURE Shrink

Coming from the work Zhou Dengwen and Chen Wengand

11.5.3 plotwavelet2

This function has been written by Benjamin Tremoulheac. This function plots the result of wavedec2 matlab function in two different modes. The first one called 'tree' displays all approximations and details coefficients (horizontal, vertical, diagonal), the second one called 'square' displays the classical "squared" representation of a wavelet image decomposition.

The command is plotwavelet2(C,S,level,wavelet,rv,mode), so given the coefficients (and other options that can be found on the help of the function) as input, the classical representation of wavelet decomposition is displayed.

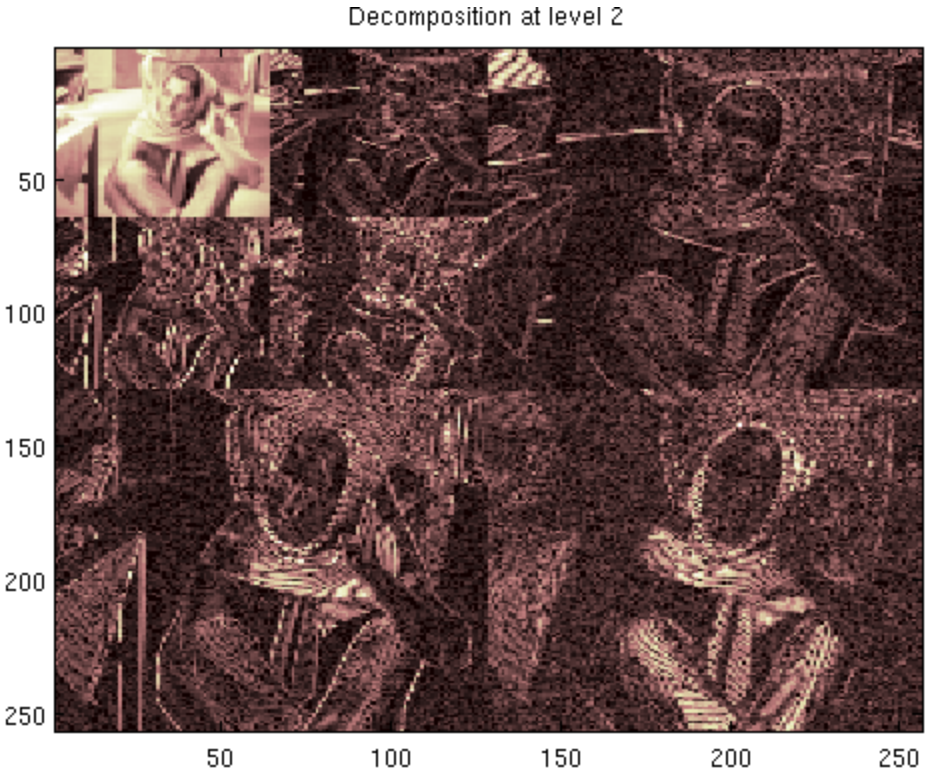


Figure 11-14 Square (classical) rapresentation for a Wavelet decomposition

11.5.4 gauss3filt

The anisotropic equation for Gaussian filtering is used. This permits to use a different sigma for time dimension. The input has to be converted from a 4D-matrix to a three dimensional one using the command ‘squeeze’ to remove singleton dimensions. The output has to be converted in double digit (command ‘double’) and finally reset to a 4D matrix with the command ‘reshape’.

12 TELOPS CAMERA GUIDE

On the HyperIR software, first of all select the folder for the experiment. It has to be on a SSD disk.

If the ambient temperature is 50°C (the maximum) the camera will take 10 minutes for the cooldown. It's better to leave it on for 1 hour to reach steady conditions for best accuracy.

To estimate the size of the recording this formula can be used:

$$WIDTH_{pixel} \cdot \left(HEIGHT_{pixel} + 2\right) \cdot 2 \cdot fps \quad \left[\frac{bytes}{s}\right]$$

$$WIDTH_{pixel} \cdot \left(HEIGHT_{pixel} + 2\right) \cdot 2 \cdot fps \cdot t_{rec} \quad [bytes]$$

The camera has a fast memory on board of 16 Gb. There is a limit of 255 sequences that can be stored. So when working with triggering a sequence has to be a multiple of the number of frames recorded after the rigger signal. In this way more sequences can be stored on the camera and they can be divided after in Matlab.

Connect the camera, if both CameraLink and GigE are present choose which one to use. To have better GigE transfer speed, in Device Info then Transfer Layer put 8164 in GEV Scps Packet Size. At full frame with a GigE connection and writing directly on the pc's hard drive the maximum frequency is 150 Hz. The memory buffer is needed to obtain higher fps.

Decreasing the height of the frame permits to obtain higher fps (to change the height the memory buffer has to be disabled). The width doesn't influence the fps because the complete row of pixel is read anyway rom the camera. Image dimension is influenced also by the width.

The 25mm lens can focus object from 20cm to infinitive. It can be used also with the lens extender. There are four different extenders: ¼ in, ½ in, ¾ in and 1 inch. Longer the extender, higher the magnification but quality at the borders of the image is lost because of vignetting.

Extender ring	Minimum working distance	Maximum working distance	Pixel resolution	Magnification
inch	mm	mm	mm	
0.25	88	122	0.045	0.33
0.5	64	74	0.027	0.556
0.75	53	57	0.021	0.714
1	47	49	0.015	1

The 1 in extender ring produce the same magnification as the 1x lens but the quality is much lower.

When switching lens a new Calibration Block has to be selected (0 for the 25mm lens and 1 for the 1x lens), the press Load Collection and wait until the camera stop blinking.

To remove vignetting stop the acquisition. Put a black body at uniform temperature in front of the lens in a way that all the frame (full frame resolution) is covered. Then switch to expert mode in the control panel. Go to Acquisition Control and insert the Black Body temperature. Press Image Correction to remove the vignetting effect, the camera will be blinking yellow during this phase. This has to be done every time the camera is switched off or it has been reset. Image correction has to be performed at the same focus.

It is important to check for frame dropping, especially when using GigE and writing to the PC's hard drive directly. Click on the sequence (bottom of the screen) and then Proprieties.

Automatic Exposure Control is turned on by default. The exposure time is self-adjusted to have a 60% of pixels' Well Filling (the maximum energy that can fill each pixel is 58 MeV). This value can be changed. Turning off the AEC a manual exposure time can be selected. Lower the exposure time higher fps can be obtain bu also noise increases. With a high exposure time the noises decrease but pixels saturation may happen. A reference of the best exposure time depending on object temperature can be see in the test report. It is suggested to work higher than 10% Well Filling. Every lens has got it's own graph. Moreover, at higher temperatures the frame rate can be higher (and the exposure time lower) because more energy is emitted and noise become less a problem.

Varius calibration mode can be chosen, the most useful are Radiometric Temperaure (°C), In-Band Radiance ($\frac{W}{m^2 sr}$) and In-Band Irradiance ($\frac{W}{m^2}$).

For high fps the Frame Buffer has to be used. There is a limitation of 255 sequences and 16 GB.

12.1.1 Triggering

A 3.3V signal is needed to start the acquisition when triggering is enabled. If triggering is enabled by the Memory Buffer a pre Moment Of Interest (MOI) can be set. In this way the limit of 255 sequences cannot be avoided if recording only small sequences.

12.1.2 Emissivity of the Object

To set the emissivity go to Calibration, Spectral Responses and Import data from the usb key. Then apply radiometric correction. With the ε icon the emissivity can be selected. Atmosphere can be neglected.

12.1.3 Image analysing

Create on the figure a point, rectangle,… then right click on min, max or mean to create a time profile of the selected value. It can be helpful to estimate noise measuring the range on a object on thermal equilibrium.

The sequence can also be exported in a video file.

12.1.4 Lens Cleaning

Manipulate the lenses with gloves, the skin has not to touch the glass' lens. It's better to have always the lens' cover on. Cleaning has to be performed as little as possible.

The first step is to remove dust with compressed air. Then use lens tissue with (in order of increasing cleaning power):

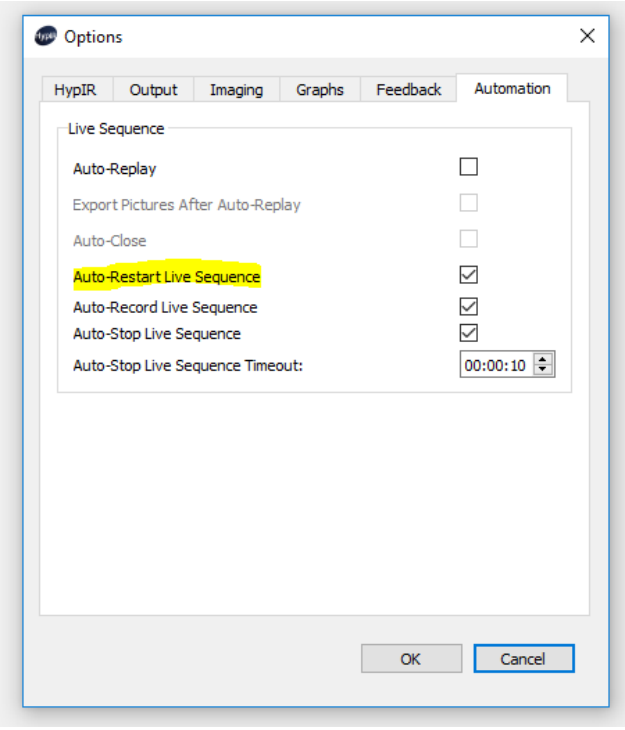
1. Isopropial alcohol
2. De-ionised water
3. Water and salt

12.1.5 Use of CameraLink

After having connected the camera to the PC by the use of the camera link cables, choose this type of connection before connecting the camera.

When working in trigger mode these tree options must be enabled to continuously record the experiment: Auto-Restart Live Sequence, Auto-Record Live Sequence and Auto-Stop Live Sequence. The timeout must be higher than 10s.

Be careful with this options (especially if triggering is not enabled) because the camera could full the hard drive and the OS would not boot again.



Windows 10 Drivers are not installed yet, they are running in compatibility mode for Windows 7. For this reason, some lags are present in the recording process.

13 BIBLIOGRAPHY

[1] C. E. Feltner and J. D. Morrow, “Microplastic Strain Hysteresis Energy as a Criterion for Fatigue Fracture,” *J. Basic Eng.*, vol. 83, no. 1, p. 15, 1961.

[2] A. M. Korsunsky, D. Dini, F. P. E. Dunne, and M. J. Walsh, “Comparative assessment of dissipated energy and other fatigue criteria,” *Int. J. Fatigue*, vol. 29, no. 9–11, pp. 1990–1995, 2007.

[3] M. V Bannikov and O. A. Plekhov, “A study of the stored energy in titanium under deformation and failure using infrared data,” vol. 24, pp. 81–88, 2013.

[4] E. Z. Stowell, “A Study of the Energy Criterion For Fatigue,” *Nucl. Eng. Des.*, vol. 3, pp. 32–40, 1966.

[5] N. W. Klingbeil, “A total dissipated energy theory of fatigue crack growth in ductile solids,” *Int. J. Fatigue*, vol. 25, no. 2, pp. 117–128, 2003.

[6] R. P. Skelton, T. Vilhelmsen, and G. A. Webster, “Energy criteria and cumulative damage during fatigue crack growth,” *Int. J. Fatigue*, vol. 20, no. 9, pp. 641–649, 1998.

[7] J. S. Short, D. W. Hoeppner, and S. L. City, “a Global / Local Theory of Fatigue Crack Propagation,” vol. 33, no. 2, pp. 175–184, 1989.

[8] A. Vshivkov, A. Iziumova, and O. Plekhov, “Experimental study of heat dissipation at the crack tip during fatigue crack propagation,” vol. 35, pp. 57–63, 2016.

[9] J. R. Rice and N. Levy, “Local Heating by Plastic Deformation at a Crack Tip,” pp. 277–293, 1969.

[10] N. Ranc, T. Palin-luc, P. Paris, N. Ranc, T. Palin-luc, and P. Paris, “Thermal effect of plastic dissipation at the crack tip on the stress intensity factor under cyclic loading Science Arts & Métiers (SAM),” 2014.

[11] P. J. Loos and F. R. Brotzen, “Localized Heat Generation during Fracture of Cyclically Loaded Steel,” vol. 14, no. July, 1983.

[12] G. Pitarresi and E. A. Patterson, “A review of the general theory of thermoelastic stress,” no. July, pp. 405–417, 2003.

[13] G. Pitarresi and E. A. Patterson, “A review of the general theory of thermoelastic stress analysis,” *J. Strain Anal. Eng. Des.*, 2003.

[14] A. S. Machin, J. G. Sparrow, and M. G. Stimson, “Mean stress dependence of the thermoelastic constant,” *Strain*, vol. 23, no. 1, pp. 27–30, Feb. 1987.

[15] W. Minkina and S. Dudzik, *Infrared Thermography: Errors and Uncertainties*. 2009.

[16] X. P. Maldague, *Nondestructive evaluation of materials by infrared thermography*. 1996.

[17] H. Louche and A. Chrysochoos, “Thermal and dissipative effects of accompanying L??ders band

propagation," *Mater. Sci. Eng. A*, vol. 307, no. 1–2, pp. 15–22, 2001.

[18] T. Boulanger, A. Chrysochoos, C. Mabru, and A. Galtier, "Calorimetric analysis of dissipative and thermoelastic effects associated with the fatigue behavior of steels," *Int. J. Fatigue*, vol. 26, no. 3, pp. 221–229, 2004.

[19] A. Iziumova, A. Vshivkov, A. Prokhorov, A. Kostina, and O. Plekhov, "The study of energy balance in metals under deformation and failure process," *Quant. Infrared Thermogr. J.*, vol. 13, no. 2, pp. 242–256, 2016.

[20] F. Maquin and F. Pierron, "Heat dissipation measurements in low stress cyclic loading of metallic materials: From internal friction to micro-plasticity," *Mech. Mater.*, vol. 41, no. 8, pp. 928–942, Aug. 2009.

[21] C. Ibarra-Castanedo, D. González, M. Klein, M. Pilla, S. Vallerand, and X. Maldague, "Infrared image processing and data analysis," *Infrared Phys. Technol.*, vol. 46, no. 1–2, pp. 75–83, 2004.

[22] A. Vandone, "Algorithms for infrared image processing," 2011.

[23] J. Semmlow, *Biosignal and biomedical image processing: MATLAB-based applications. 2004*. 2004.

[24] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using MATLAB*. 2004.

[25] M. W. K. Law and A. C. S. Chung, "Efficient Implementation for Spherical Flux Computation and Its Application to Vascular Segmentation," vol. 18, no. 3, pp. 596–612, 2009.

[26] "https://en.wikipedia.org/wiki/Kalman_filter." .

[27] "http://www.bzarg.com/p/how-a-kalman-filter-works-in-pictures/." .

[28] "http://imagej.net/plugins/kalman.html."

[29] T. F. Chan and J. Shen, "Image processing and analysis: variational, PDE, wavelet, and stochastic methods," *Book*, 2005.

[30] J.-P. Antoine, R. Murenzi, P. Vandergheynst, and S. T. Ali, *Two-Dimensional Wavelets and their Relatives*. 2008.

[31] S. Sutha, E. J. Leavline, and D. A. A. G. Singh, "A comprehensive study on wavelet based shrinkage methods for denoising natural images," *WSEAS Trans. Signal Process.*, vol. 9, no. 4, pp. 203–215, 2013.

[32] A. David, R. Girón, H. L. Correa, A. David, R. Girón, and H. L. Correa, "Nuevo algoritmo de detección y corrección de píxeles anómalos en imágenes A new algorithm for detecting and correcting bad pixels in infrared images," vol. 30, no. 2, pp. 197–207, 2010.

[33] J. P. Lewis, "Fast Normalized Cross-Correlation Template Matching by Cross-," vol. 1995, no. 1, 1995.

[34] T. L. Anderson, *Fracture Mechanics: Fundamentals and Applications, 3rd Edition*. 2005.

[35] K. S. Bhalla, "Thermomechanics of slow stable crack growth : closing the loop between

experiments and computational modeling," vol. 70, pp. 2439–2458, 2003.

[36] I. Chasiotis, "Springer Handbook of Experimental Solid Mechanics," in *Springer Handbook of Experimental Solid Mechanics*, 2008.

[37] J. D. Skinner, Jr. A and J. S. Jr, "Finite Element Predictions of Plasticity–Induced Fatigue Crack Closure in Three-Dimensional Cracked Geometries," *Statew. Agric. L. Use Baseline 2015*, vol. 1, no. August, 2001.

[38] A. Mechanics, *Residual Stress , Thermomechanics & Infrared Imaging , Hybrid Techniques and Inverse Problems , Volume 9*, vol. 9. 2016.

[39] L. Bodelot, L. Sabatier, E. Charkaluk, and P. Dufrénoy, "Experimental setup for fully coupled kinematic and thermal measurements at the microstructure scale of an AISI 316L steel," vol. 501, pp. 52–60, 2009.

[40] A. Fedorova and M. Bannikov, "Infrared thermography study of the fatigue crack propagation," *Frat. ed Integrità Strutt.*, vol. 21, pp. 46–53, 2012.

[41] A. Iziumova and O. Plekhov, "Calculation of the energy J-integral in plastic zone ahead of a crack tip by infrared scanning," *Fatigue Fract. Eng. Mater. Struct.*, vol. 37, no. 12, pp. 1330–1337, 2014.

[42] G. Meneghetti and M. Ricotta, "Evaluating the heat energy dissipated in a small volume surrounding the tip of a fatigue crack," *Int. J. Fatigue*, vol. 92, pp. 605–615, 2016.

[43] K. S. Bhalla, A. T. Zehnder, and X. Han, "Thermomechanics of slow stable crack growth: closing the loop between experiments and computational modeling."

[44] A. S. Engineering, "Full-range Stress-strain Curves for Stainless Steel Alloys," no. March 2002, 2001.

[45] W. Oliferuk, "Energy storage rate and Considere criterion in uniaxial deformation of austenitic steel," pp. 0–1.

[46] "http://www.azom.com/article.aspx?ArticleID=2382." .

[47] "http://www.matweb.com/."