A review of P.D.E. models in image processing and image analysis

F. Guichard, L. Moisan, J.-M. Morel

Abstract

The years 1985-2000 have seen the emergence of several nonlinear P.D.E. models in image restoration and image analysis. Before that date, the heat equation and the reverse heat equation had been considered as relevant, one as a model of image smoothing compatible with Shannon conditions, and one as a restoration model proposed by Gabor. We try in this review to organize the P.D.E. models according to their genealogy from the initial heat equation and according to their very diverse use : some are useful for image denoising, some for image deblurring, some for invariant smoothing in view of shape recognition. Some permit to define easily active contours (snakes), some may be used for a nonlinear interpolation of sparse images. We show many experiments illustrating these different applicative aspects.

1 Introduction

This paper adresses a possible theory of image *low level* analysis. Image "low level" analysis aims at extracting reliable, local geometric informations from a digital image. Such informations are often called "features" and they are used in order to compare an image to other images. For instance, these features can be used for motion estimation, or to retrieve shapes, or to build the still hypothetic "high level" vision. The observed image is the result of a smoothing of the original photon flux and is therefore continuous. It is nontheless well admitted that the subjacent "real image", namely the focused photon flux, is either a measure or, for more optimistic authors, a function which presents strong discontinuities. Rudin and de Giorgi proposed independently in 1984 the space BV of functions with bounded variation as the right function space for "real" images. More recently (1999), however, Gousseau and Alvarez [1] used a statistical device on digital images to estimate how their real subjacent images oscillate. They deduced, by geometric measure arguments, that the "real" physical images have in fact unbounded variation. We may therefore accept the idea that the subjacent high resolution image behaves in a strongly oscillatory way. Although the digital images present an averaging of this oscillatory phenomenon, common sense tells us that they must have anyway strong discontinuities at transitions between different observed objects, i.e. on the apparent contours of physical objects. The BV space looked at first well adapted to that aim because it contains functions having step discontinuities.
One of the goals of image analysis has ever been to find such discontinuities in an image. This search is called "edge detection" because early vision research played with images of cubes. Along the edges of the cubes, the light intensity behaved, in a first approximation, as a step function. Unfortunately, the early research in vision led to the sad discovery that one could find edges "everywhere" in a digital image (Marr Marr:1982:V), due to the oscillations remaining in the digital image after the digitization step. As a consequence, the image analysis process was conceived as a smoothing process, permitting to decluster the true "edges" from the inherent noise. As in Distribution theory, a smoothing was necessary before computing any derivative. This is why the heat equation was proposed and a new doctrine proposed : the "scale space". Scale space means that, instead of talking of features of an image at a given location, we talk of them at a given location and at a given scale. The scale quantifies the amount of smoothing performed on the image before computing the feature. We shall therefore see in experiments "edges at scale 4" and "edges at scale 7" as different outcomes of an edge detector.

Which kind of smoothing should be performed ? Three terms associated with image analysis operators arise here, to which we will give a more and more precise meaning.

The first one, "locality", is related to the occlusion problem : most optical images are made of a superposition of different objects partly hiding each other. It is plain that we must avoid mixing them in the analysis, as would do e.g. a wide convolution. Thus, the analysis must be made as local as possible. As we shall see, the heat equation is the worst candidate to the task, since it makes a wide-range mélange of grey levels.
The second key word is "iteration". Indeed, we shall see that it is generally better, from the locality viewpoint, to iterate a very local smoothing operator than to apply it directly at a large scale. This is precisely not true for the heat equation ! Iterating the convolution of small gaussian kernels is stricly the same as convolving directly the image with a big gaussian. Now, iteration of very local filters will bring a significant improvement for some of the most relevant nonlinear filters which we shall consider, namely the median filter and the affine erosion-dilations. At this point, it must be immediately announced that the combination of smoothing, locality and iteration implies that we are talking about parabolic partial differential equations.
Our last key word is "invariance". The invariance requirements play a central role in image analysis because the objects to be recognized have to be recognized under varying conditions of illumination (contrast invariance) and from different points of view (projective invariance). Contrast invariance is one of the key requirements of a famous image analysis theory, the Mathematical Morphology (Matheron [25], Serra [32]). This theory proposed a list of contrast invariant image analysis operators (dilations, erosions, median filters, openings, closings,...) We shall involve this theory, as we shall attempt to localize as much as possible the "morphomath" operators to extract their behaviour at small scale, and then iterate them. As an outcome, we shall prove that several geometric partial differential equations, namely the curvature motions, can be considered as the common asymptotic denominators to many "morphomath" operators. These P.D.E.'s permit therefore to fuse the Scale Space doctrine and Mathematical Morphology. In particular, affine invariant morphomath operators, which looked unpractical, turn out to yield in their local iterated version a very affordable P.D.E., the so called "affine morphological scale space" (A.M.S.S.).
In this paper, we shall make a survey of most P.D.E.'s which have been proposed for image analysis.

2 Image Processing and Analysis

Image processing can be divided in three parts, corresponding to as many different goals. The first one derives from the discrete nature of images and the search of their minimal representation in terms of digital memory. This discipline is called *image compression* (see Figure 1). The second goal is the *restoration* of

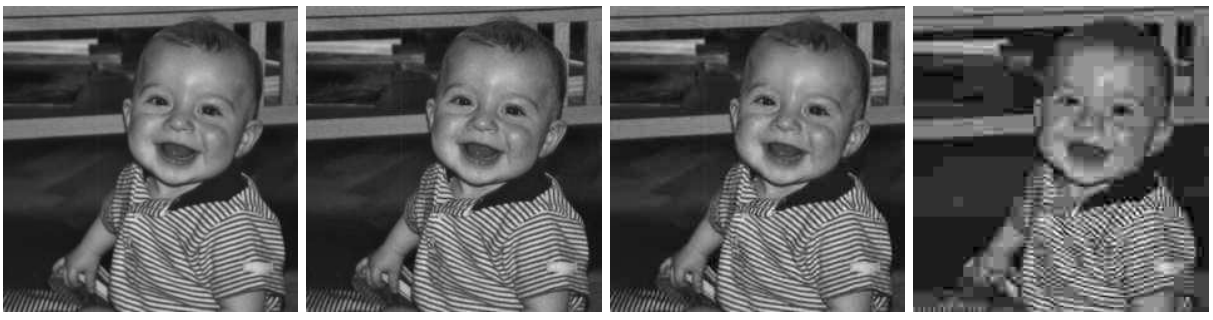

Figure 1: Compression. From left to right : An original image and its more and more compressed versions : compression factor 7, 10 and 25 respectively. One of the first goals of image processing is the definition of algorithms permitting high compression factors without visible alteration. Compression may, however alter the image.

a better version of an image, given a generation model with noise and blur, or other perturbations. This is illustrated in Figure 2. The image on the left is apparently destroyed: more than 75% of its pixels have been put at a random value. We can nontheless restore it significantly: here is, on the right, such a restored version. The third goal is analysis , which means in Greek "breaking into parts". Look at the level curve of Figure 3, extracted from a hand image: it is full with a mix of details and noise. What if we ask for a sketchy version, where, however, all essential details are kept ? The curve on the right

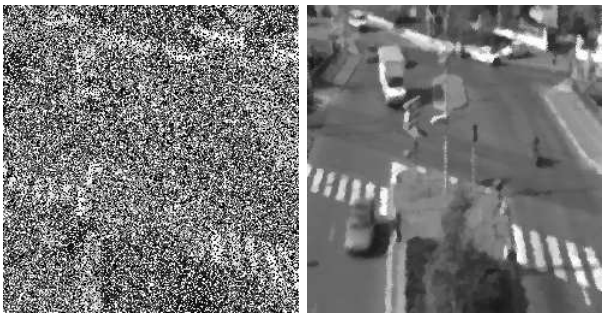

Figure 2: Denoising. A second goal of image processing : the restoration. Left: original noisy image (simulated salt and pepper noise up to 75%), right: denoised version

is such a sketchy version, where most of the spurious details have disappeared, but the main structures are maintained. This is what we shall call *image analysis* . The aim is not denoising or compression : it is to construct an invariant code putting in evidence the main parts (here, for instance, the fingers) and permitting a fast recognition in a large database of shapes.

Figure 3: Analysis of a shape. Left : Original scanned shape, then some simplified versions : the aim here is not restoration, but analysis, that is, to define more and more sketchy versions of the shape. Those sketchy versions may permit a very short and invariant encoding of the shape. Notice how the number of inflexion points of the shape has decreased in the simplification process.

3 The Heat equation

The heat equation arises naturally in the image generation process. Indeed, according to Shannon's theory, an image can be correctly represented by a discrete set of values, the "samples", only if it has been previously smoothed. This is illustrated in Figure 4 : Let us call the original baby image "Victor". If we attempt to reduce the size of Victor by a mere subsampling, that is by taking a point of each sixteen, we obtain a new and smaller image, in which the subsampling has created new and unstable patterns : see how new stripes have been created, with a frequency an direction which has nothing to do with the original ! If, instead of being steady, the camera moved, those newly created patterns would move and flicker in a totally uncontrolled way. This kind of moving pattern appears often in recent commercial DVD's. They have simply been subssampled against the Shannon rule. Let us now comment briefly how the subsampling should be done. According to Shannon's theory, a previous smoothing must to be done before the subsampling. We start with u_0, the original image, a real function defined on a domain of $\mathbb{R}^2$. Then a blur kernel k is applied, i.e. we convolve u_0 with k to obtain a new image $k * u_0$. A subsequent subsampling is thereafter possible, where the distance between samples is related to the band-width of the blur kernel by the Nyquist rule. Stability of the image representation is maintained.
This simple remark, that smoothing is a necessary part of image formation, leads us to our first PDE's. Gabor remarked in 1960 that the difference between the original and the blurred image is roughly proportional to its laplacian. In order to formalize this remark, we have to notice that k is spatially concentrated,

Figure 4: Shannon theory and subsampling. From left to right: original image, smoothed image, subsampling of the original image and subsampling of the smoothed image. In the subsampling, one point of each 4 is taken in the horizontal and vertical directions. In order to make the reduced image still visible, we have zoomed back the subsampled versions by a zoom factor 4. We clearly see that subsampling an image without previous smoothing creates aliasing : high frequencies are projected onto lower frequencies and therefore generate new patterns. Shannon theory tells us how to remove those potentially parasite high frequencies before subsampling. This results in the necessity of smoothing the image before subsampling.

and that we may introduce a scale parameter for k, namely $k_h(\mathbf{x}) = h^{-1}\, k(h^{-1/2}\mathbf{x})$. Then

$$\frac{u_0 * k_h(\mathbf{x}) - u_0(\mathbf{x})}{h} \to \Delta u_0(\mathbf{x}),$$

so that when h gets smaller, the blur process looks more and more like the heat equation

$$\frac{\partial u}{\partial t} = \Delta u, \quad u(0) = u_0.$$

Conversely, Gabor deduced that we can, in some extent, deblur an image by reversing time in the heat

Figure 5: Heat equation and blur. Left : original image, right : the heat equation has been applied to some scale and the resulting image is blurred.

equation :

$$\frac{\partial u}{\partial t} = -\Delta u, \qquad u(0) = u_{observed}.$$

Numerically, this amounts to substracting its laplacian from the observed image :

$$u_{restored} = u_{observed} - h\Delta u_{observed}.$$

This operation can be repeated several times with some small values of h, until it... blows up. Indeed, the reverse heat equation is extremely ill-posed. All the same, this Gabor method is efficient and can be applied with some success to most digital images obtained from an optical device. Let us examine what happens with Victor (Figure 6). We see that the method yields some improvement at the beginning and then blows up. We can also simulate a blur on Victor and try to go back : again, the process blows up but yields a significant improvement at some scale.

Figure 6: Gabor's deblurring. Gabor proposed in 1960 to deblur an image by substracting its laplacian : this means inverting the heat equation ! Left : original image, middle : three iterations of Gabor's algorithm, right : ten iterations. As is well known, and can be observed in the right image, the inverse heat equation blows up. A few iterations can, as we see in the middle, nicely enhance the image.

Figure 7: Gabor's deblurring again. Same deblurring experiment as in Figure 6, but applied on a much more blurred image

We therefore see two directions. One is to improve, to stabilize, the reverse heat equation. We shall see that this is doable by nonlinear models. The second direction is to go on with the heat equation : we can numerically simulate a further blurring of the image. Why should we do so ? Because, first, this leads to the wavelet theory and its applications to optimal multiscale sampling and compression. Second, iterated linear and nonlinear smoothing (that is, nonlinear PDE's) will be relevant to our main goal : image analysis.

4 Inverse heat equation and deblurring

We can indeed improve the time-reverse heat equation. The first example, due to Rudin and Osher in 87 [31] and 92 [30] proposes an pseudoinverse, where the propagation term —Du— (the modulus of the gradient of u) is tuned by the sign of the laplacian.

$$\frac{\partial u}{\partial t} = -\text{sign}\big(\Delta u\big)|Du|.$$

The equation is called "shock filter". As we shall see, this equation propagates, with a constant speed, the level lines of the image in the same direction as the reverse heat equation would do. It therefore enhances the image. The equation is more or less equivalent to a good old nonlinear filter due to Kramer in the seventies. Kramer's filter [21] can be interpreted as a partial differential equation, by the same kind of heuristic arguments which Gabor developed to derive the heat equation. This equation is

$$\frac{\partial u}{\partial t} = -\text{sign}\big(D^2u(Du,Du)\big)|Du|.$$

Thus, the laplacian is replaced by a directional second derivative of the image, $D^2u(Du, Du)$. We shall later on interpret this differential operator as an "edge detector". Kramer's equation yields a slightly better version of shock filter as is illustrated in Figure 8. Both deblurring equations work... to some

Figure 8: Image deblurring by shock filters and by a variational method. From left to right : blurred image, Rudin-Osher shock filter [30] which is a pseudoinverse of the heat equation attaining a steady state, Kramer's [21] improved shock filter, also attaining a steady state and the Rudin, Osher, Fatemi [30] restoration method, obtained by deblurring with a controlled image total variation. This last method is very efficient when the noise and blur models are known. It is currently being used by the French Space Agency (CNES) for satellite image restoration.

extent. They experimentally do not blow up and attain steady states ! The third deblurring method we can mention here is, to our knowledge, the best version. It poses the deblurring problem as an inverse problem. Given the observed image u_0, we try to find a restored version u such that $k * u$ is as close as possible to u_0 and the oscillation of u is nontheless bounded :

$$u_{restored} = \text{Argmin} \left(\int |Du| + \lambda (k * u - u_0)^2 \right).$$

The parameter λ tunes the oscillation we allow for the restored version. If λ is large, the restored version will satisfy accurately the equation $k * u = u_0$, but may be very oscillatory. If instead λ is small, we get a smooth but unaccurate solution. This parameter can be computed in principle as a Lagrange multiplier. The obtained restoration can be remarkable. We display the best result we can obtain with the blurred Victor in Figure 8-right. This total variation restoration method also has fast wavelet packets versions. It recently won a benchmark in satellite image deblurring organised by the French Space Agency (CNES).

The original remark of Gabor, about image generation being related to the Laplacian of the image, leads to the wavelet theory as well. Here is how it works : if we convolve the image with some smoothing kernel and thereafter make the difference, we obtain a new image, actually a laplacian, which turns out to be faded with respect to the original. In Figure 9, the last image on the right shows in black the values of this laplacian image of Victor which differ significantly from zero : in most natural images, as here, this representation is sparse and adapted to compression. This is why one of the first wavelet representations, due to Burt and Adelson in 83 [4], was called "Laplacian pyramid". It boils down to the iteration of a convolution followed by subsampling. We only keep the differences between images smoothed at different scales, i.e. their laplacians. The objective is a compressed representation, but to the price of a loss of invariance due to the multiscale subsampling.

In image analysis, the heat equation has had a very different use: Marr [24], Hildreth [24], Canny [5], Witkin [38], Koenderink [19] proposed in the eighties to analyse an image by applying the heat equation. As Rudin and Florack noticed, this is related to distribution theory. Indeed, details of the images, like boundaries, corners and other singularities cannot be computed without some previous smoothing because they are derivatives of a nonsmooth function. And this smoothing has to be multiscale because the image is multiscale ! The heat equation is easily proved to be the only good candidate to the task if image analysis has to be linear. What derivatives should be computed in an image ? The early research in computer vision proposed "edge detection" as a main tool : it is assumed that the apparent contours of the objects and also the boundaries of the facets of objects, result in step discontinuities in the image, while, inside those boundaries, the image oscillates mildly. The apparent contour points, or "edges points" will be computed

Figure 9: The "laplacian pyramid" of Burt and Adelson [4]. From left to right : original image, image blurred by gaussian convolution, then difference between the original image and the blurred version, which simulates the laplacian of the original image. In black in the last image, points where this laplacian image is large. This experiment simulates the first step of the laplacian pyramid. The laplacian image is, for most digital images, a sparse representation, therefore well adapted to compression.

as points where the gradient is is some sense largest. Two ways to do so : Hildreth and Marr proposed the points where Δu crosses zero. A significant improvement was done by Canny, who proposed to compute the points where $|Du|$ is maximal on the gradient lines. Such points satisfy $D^2u(Du, Du) = 0$. Figure 10 displays what happens when we smooth the image with the heat equation and compute the points where $D^2u(Du, Du) = 0$ and $|Du|$ is large enough. At first, everything in the image is boundary : the image, being a very oscillatory function, has "inflexion points" everywhere ! After some evolution of the heat equation, we can see what happens : we are able to extract some structure.

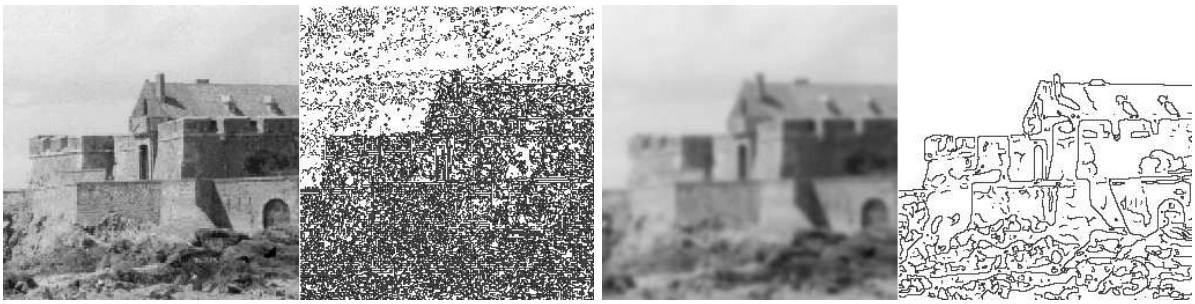

Figure 10: Heat equation and Canny's edge detector. Boundaries, or "edges" of the image can be defined as points where the gradient attains a maximal and large value along the gradient lines. This amounts to say that edge points are points where $D^2u(Du, Du)$ crosses zero and $|Du|$ is large. Canny's edge detector [5] computes those points. On the left, the original image, followed by the edge points found. They make a very dense set, because of the oscillatory character of the image. Next, the image blurred by the Gauss kernel (heat equation) and the Canny edges found. The heat equation has removed the "irrelevant" edges

5 Nonlinear diffusion models

If the heat equation is, under sound invariance requirements, the only good linear smoother, there are instead many nonlinear ways to smooth an image. The first one was proposed by Perona and Malik in 87 [29]. The idea is roughly to smooth out what has to be smoothed, the irrelevant, homogeneous, regions and to enhance instead the boundaries. Thus, the diffusion should look like the heat equation when $|Du|$ is small and an inverse heat equation should instead be applied when $|Du|$ is large. Here is the equation in divergence form.

$$\frac{\partial u}{\partial t} = \text{div}\left(g(|Du|^2)\right),$$

where $g(s) = \frac{1}{1+\lambda s^2}$ decreases when s increases. It is easily checked that we have a diffusion equation when $|Du| \leq \lambda$ and an inverse diffusion equation when $|Du| > \lambda$. In order to do so, we rewrite the equation in

the following way. We consider the second derivative of u in the direction of Du,

$$u_{\eta\eta} = D^2u\left(\frac{Du}{|Du|}, \frac{Du}{|Du|}\right)$$

and the second derivative in the orthogonal direction,

$$u_{\xi\xi} = D^2u\left(\frac{Du}{|Du|}^{\perp}, \frac{Du^{\perp}}{|Du|}\right),$$

where $Du = (u_x, u_y)$ and $Du^{\perp} = (-u_y, u_x)$. The laplacian can be rewritten in the intrinsic coordinates (ξ, η) as $\Delta u = u_{\xi\xi} + u_{\eta\eta}$. The Perona-Malik equation rewrites

$$\frac{\partial u}{\partial t} = \frac{u_{\xi\xi}}{1 + \lambda^2|Du|^2} + \frac{(1 - \lambda^2|Du|^2)u_{\eta\eta}}{(1 + \lambda^2|Du|^2)^2}.$$

So the first term always appears as a one-dimensional heat equation in the direction orthogonal to the

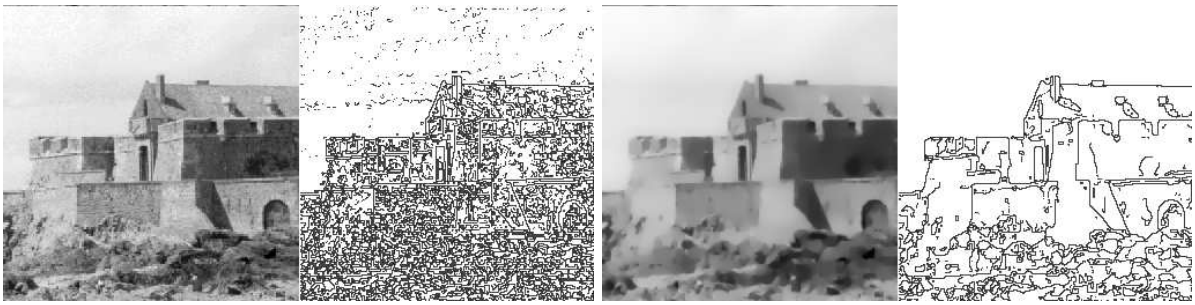

Figure 11: Perona-Malik equation [29] and edge detection : same experiment as in Figure 10, but the heat equation has been replaced by the Perona-Malik equation. Notice that the edge map looks slightly better localized than with the heat equation.

gradient, tuned by the size of the gradient though. The second term can be a directional heat equation, or reverse heat equation in the direction of the gradient. So we indeed mix in this model the heat equation and the reverse heat equation ! We compare in Figure 11 the Perona-Malik with the classical heat equation in terms of accuracy on the boundaries obtained by Canny's edge detector : at a comparable scale of smoothing, we clearly gain some accuracy in the boundaries and get rid of more "spurious" boundaries. The representation is both more sparse and more accurate.

Now, this ambitious model attempts to put in a single operator two very different goals which we already mentionned, namely restoration and analysis. This has a cost : the model contains a "constrast threshold" which can only be fixed manually. Mathematical existence and uniqueness are not guaranteed, despite some attempts by Kichenassamy [17] and Weickert. Let us summarize the involved parameters : we need to fix both λ and the smoothing scale(s) t and the threshold on the gradient in Canny's detector as well. We obviously must take the same gradient threshold λ in Canny's detector and in the Perona-Malik equation. All the same, we have a two parameters game : how will this be dealt with in automatic image analysis ? This question seems to have no general answer for the time being. An interesting attempt based on statistical arguments is made, however, in Black, Sapiro.

If any nonlinear diffusion can be an image analysis model, why not trying them all ? This is exactly what has happened in the past ten years. We can claim that almost all possible nonlinear parabolic equations have been proposed. The logic in this proliferation of models is this : each attempt fixes one intrinsic diffusion direction and tunes the diffusion by the size of the gradient or by the size of a nonlinear estimate of the gradient. Sometimes, the proposed models are even systems of PDE's, but in order to remain concise, we shall focus on the simplest proposed examples. We can start with Rudin-Osher-Fatemi's model

[30], which consists, for the smoothing term, of minimizing the total variation of u. The gradient descent for $\int |Du|$ writes

$$\frac{\partial u}{\partial t} = \text{div}\left(\frac{Du}{|Du|}\right) = \frac{1}{|Du|}u_{\xi\xi}.$$

Written in that way, the method appears as diffusion in the direction orthogonal to the gradient, tuned by the magnitude of the gradient. Caselles et al. proved that this equation is indeed well posed in the space of bounded variation. A variant was proposed independently by Alvarez and al. [2],

$$\frac{\partial u}{\partial t} = \frac{|Du|}{|k * Du|}\text{div}\left(\frac{Du}{|Du|}\right) = \frac{1}{|k * Du|}u_{\xi\xi},$$

where the tuning of the gradient is nonlocal. Kimia, Tannenbaum and Zucker [18] proposed, endowed in a more general shape analysis framework, the simplest equation of the list,

$$\frac{\partial u}{\partial t} = |Du|\text{div}\left(\frac{Du}{|Du|}\right) = D^2u\left(\frac{Du}{|Du|}^{\perp}, \frac{Du^{\perp}}{|Du|}\right) = u_{\xi\xi}.$$

This equation had been proposed some time before in another context by Sethian [33] as a tool for front propagation algorithms. This equation, which we call in continuation "curvature equation", is a "pure" diffusion in the direction orthogonal to the gradient. The Weickert equation is a variant of the curvature equation, with nonlocal estimate of the direction orthogonal to the gradient : the diffusion direction $d = SEigen(k * (Du \otimes Du))$ is computed as the eigenvector of the least eigenvalue of $k * (Du \otimes Du)$: if the convolution kernel is removed, this eigenvector simply is $Du^{\perp}$. The three mentionned models can be interpreted as diffusions in a direction orthogonal to (an estimate of) the gradient, tuned by the magnitude of the gradient (Figure 13). Other diffusions have been considered as well : for interpolation goals, Caselles

Figure 12: A proliferation of diffusion models. From left to right: Original image, Perona and Malik equation 1987 [29], Zhong, Carmona 1998 [6] (diffusion along the least eigenvector of D^2u) and Sochen, Kimmel and Malladi, 1998 [34] (minimization of the image graph area.

et al. [10] proposed a diffusion which may be interpreted as the strongest possible image smoothing,

$$\frac{\partial u}{\partial t} = D^2u(Du, Du).$$

This equation is not used as the other ones as a preprocessing of the image, but a way to interpolate between the level lines an image with sparse level lines (Figure 14). Zhong and Carmona [6] proposed a diffusion in the direction $d = SEigen(D^2u)$ of the eigenvector with least eigenvalue of D^2u (Figure 12). Sochen, Kimmel and Malladi [34] propose instead a nondegenerate diffusion, associated with a minimal surface variational formulation : their idea was to make a gradient descent for the area of the graph of u, $\int \sqrt{1 + |Du|^2}$, which leads to the diffusion equation (Figure 12).

$$\frac{\partial u}{\partial t} = \text{div}\left(\frac{Du}{\sqrt{1 + |Du|^2}}\right).$$

Figure 13: A proliferation of diffusion models (II). From left to right: Osher, Sethian (1988) curvature equation [33], Rudin, Osher and Fatemi(1992) minimization of the image total variation [30], Alvarez, Lions et al. (1992) nonlocal variant of the preceding [2], Weickert (1994) nonlocal variant of the curvature equation [36]. All of these models only diffuse in the direction orthogonal to the gradient, with a more or less local estimate of this direction.

Figure 14: A proliferation of diffusion models (III) Here, the diffusion is made in the direction of the gradient and the model is applied for image interpolation when level lines are sparse. From left to right : original image, quantized image (only 10 levels are kept - 3.32 bits/pixel) and reinterpolated image by the Caselles and Sbert (1998) algorithm [10]. They apply a diffusion on the quantized image, with values on the remaining level lines as boundary conditions.

6 Contrast invariance and shape analysis

Among the mentionned models, only the curvature motion was explicitly proposed by Kimia, Tannenbaum and Zucker [18] as shape analysis tool. We shall now explain why.
In order to do so, we have to give a definition of image analysis. There might be as many ways to define this discipline as they are applicational goals involving digital images. Now, the range of applications is as wide as the human activity, since most of the scientific and technical human activity, including even sound analysis (visual sonagrams), involves the perceptual analysis of images. Fortunately, we have at hand a mathematical shortcut to avoid an endless list of partial and specific requirements : this shortcut, well known in Mechanics, consists of stating invariance requirements. Invariance requirements will be a short list and they will, as we shall see, give a possible classification of models and point out the ones which are the most adequate for all purposes image analysis tools. The first invariance requirement is the Wertheimer principle [37] according to which visual perception (and therefore, may we add, image analysis) should be independent of the image contrast. We state this in the following way :

Contrast invariant classes
u and v are said to be (perceptually) equivalent if there is a continuous increasing function g such that $v = g(u)$.

Contrast invariance requirement
An image analysis operator T must act directly on the equivalence class. As a consequence, we may ask that $T(g(u)) = g(Tu)$, i.e. a commutation of the image analysis operator with contrast changes.

The contrast invariance requirement rules out the heat equation and all models stated before, except the curvature motion. Contrast invariance led Matheron in seventy-five to reduce image analysis to a set analysis, namely the analysis of level sets. We call upper level set with level λ of an image u the set

$$\mathcal{X}_\lambda u = \{\mathbf{x}, u(\mathbf{x}) \geq \lambda\}.$$

We can define in exactly the same way the lower level sets, by changing "$\geq$" into "$\leq$". The main point to retain here is the global invariance of level sets under contrast changes, namely, if g is a continuous increasing contrast change, then,

$$\mathcal{X}_{g(\lambda)} g(u) = \mathcal{X}_\lambda u.$$

According to Mathematical Morphology, this image analysis doctrine founded by Matheron, all of the image shape information is therefore contained in the level sets : it can be proved that an image can be reconstructed, up to a contrast change, from its set of level sets (Figure 15 : an image and some of its level sets).

Figure 15: An image and one of its level sets. Right : level set 140 of the left image. This experiment illustrates Matheron's thesis that the main shape information is contained in the level sets of the image. Level sets are contrast invariant.

The contrast invariance requirement yields powerful and simple denoising operators as the so called "Extrema killer" defined by Vincent [35] and Serra in 1993. This image operator simply removes all connected components of upper and lower level sets with area smaller than some fixed scale. This is not a PDE, actually it's much simpler ! Now, its effect is amazingly good for impulse noise i.e. local destructions of the image, spots. In Figure 16, we see a image degraded up to 75%. Below, its restoration by the extrema killer. Left, result of the same operator applied to the original.

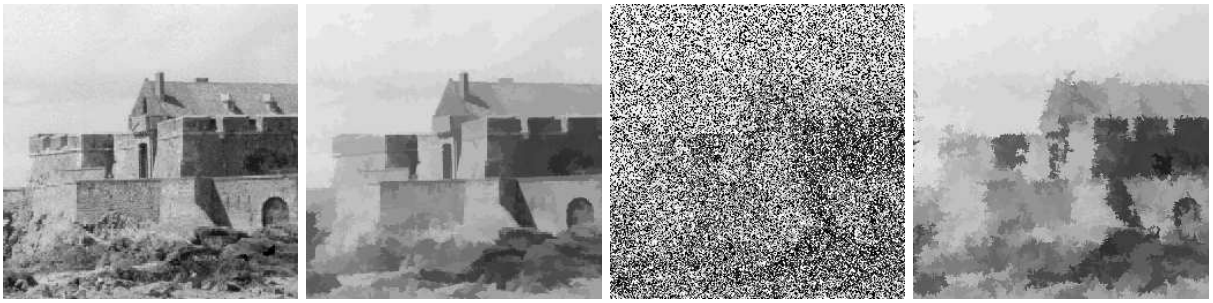

Figure 16: The "extrema killer" filter : all connected components of the upper or lower level sets with small area are removed from the image. From left to right : original image, extrema killer applied with area 80 pixels, then 75% salt and pepper noise added to the original image and the same filter applied.

Caselles and Coll localized farther in 1996 [8] the contrast invariance requirement in image analysis. They

proposed as the main object of analysis the level lines of the images, that is, the boundaries of level sets. This proposition makes sense for a digital image, which is assumed to be a sampling of a very smooth function as the result of the optical smoothing. We can therefore define the level lines if, e.g., the interpolated image is C^1 as is guaranteed by the canonical Shannon interpolation. There may be other interpolation methods, and even interpolations into a discontinuous functions : this is the case if, e.g., we consider the digital image as constant on each pixel. We must then for each interpolation method make clear how the level lines are computed and what their structure is. Two properties are desirable : that the level curves indeed are curves in some affordable sense (Jordan rectifiable curves) and that they are nested, i.e. never cross, so that they make an inclusion tree. A study of Kronrod (1950) shows that if the function u is continuous, then the isolevels sets $\{\mathbf{x}, u(\mathbf{x}) = \lambda\}$ are nested : they build a tree ordered by inclusion. Now, these isolevel sets need not be really curves. Monasse (2000) generalized recently the preceding result to lower semicontinuous or upper semicontinous functions. His result implies that the simplest, piecewise constant, interpolation of an image yields a nested set of Jordan curves bounding the pixels. Thus, we have two good ways to associate with the digital image a set of nested Jordan curves. We call this set "topographic map" [1]. We display in Figure 17 the level lines of a digital image at some fixed level. By the introduction of the topographic map, the search for image smoothing, which we had already reduced to set smoothing, is further reduced to curve smoothing, provided of course this smoothing preserves curve inclusion.

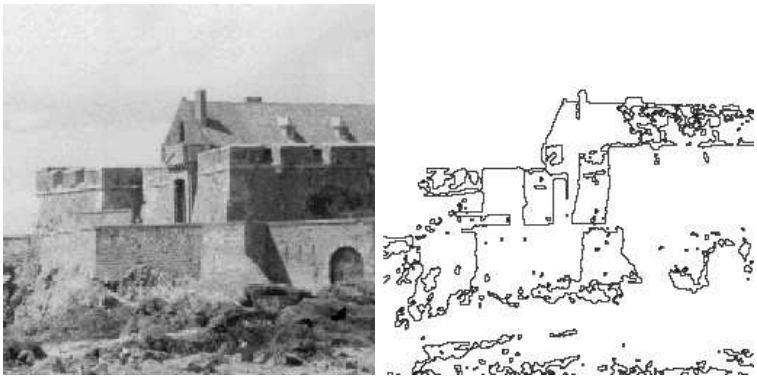

Figure 17: Level lines of an image. Level lines, defined as the boundaries of level sets, can be defined to be a nested set of Jordan curves. They give a contrast invariant representation of the image. Right : level lines with level 183 of the left image.

Chen, Giga and Goto [11] and Alvarez et al. Alvarez:1993:AFE proved that, under the usual invariance requirements for image processing, including the contrast invariance, all image multiscale analyses should have the form of a curvature motion, namely

$$\frac{\partial u}{\partial t} = F(curv(u), t)|Du|,$$

where F is increasing with respect to its first argument. This equation can be interpreted as this : we consider a point $\mathbf{x}$ on a given level curve of $u(t)$, at time t. We call $n(\mathbf{x})$ the normal vector to the level curve and $curv(\mathbf{x})$ its curvature. Then the preceding equation turns out to be associated with the curve motion equation

$$\frac{\partial \mathbf{x}}{\partial t} = F(curv(\mathbf{x}))n(\mathbf{x}),$$

which describes how the point $\mathbf{x}$ moves in the direction of the normal. Not much more can be said at this level of generality on F. Now, two particular cases happen to play a prominent role. First, the case $F(curv(u), t) = curv(u)$, the so called curvature equation which we already met, and second the case

[1] This point of view also is coherent with the "BV assumption" which we mentionned at the beginning of the introduction, according to which the right function space for images should be the space BV of functions with bounded variation. By coarea formula, we can then describe the image by a bunch of Jordan level curves (see Ambrosio et al.) Now, it is in general false for BV functions that boundaries of lower and upper level sets make a nested set of curves : these curves may cross.

$F(curv(u),t) = curv(u)^{\frac{1}{3}}$.

This particular form for the curvature dependence, the power one third, permits to get a very relevant additional invariance, the affine invariance. We would like to have a full projective invariance, but a theorem proved by Alvarez et al. shows that this is impossible. The best we can have is invariance with respect to the so called chinese perspective, which preserves parallelism. Most of the mentionned equations, particularly when F is a power of curvature, have a viscosity solution in the sense of Crandall and Lions [12], as shown by recent works of Ishii and Souganidis [16].

Contrast invariance is also naturally relevant in motion analysis, according to the principle that the light intensity reflected by a same physical point (lying on a Lambertian surface) should not depend on the point of view. In image epipolar geometry, where the x axis represent space and the y axis represent time, this means that level lines correspond to trajectories (see Figure 18). In this non-isotropic context, a diffusion has been proposed [26] that writes

$$u_t = D^2u(v,v), \qquad \text{with} \qquad v = \left(-\frac{u_y}{u_x}, 1\right)^T.$$

Such a diffusion can be interpreted as follows : each level line of u, considered as a (potentially multivalued) graph $y(x)$, is filtered independently by the 1-dimensional heat equation [13]. The smoothing it performs preserve the physical interpretation of the scene, as proven in [28].

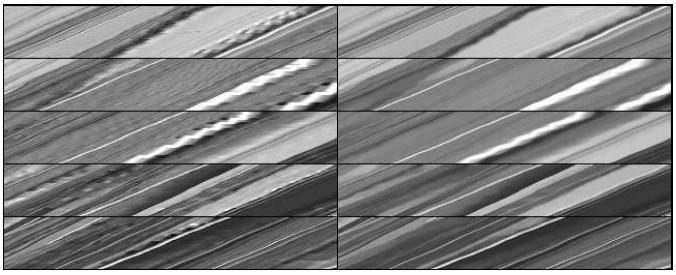

Figure 18: A non-isotropic contrast-invariant diffusion devoted to epipolar motion filtering and depth recovery [26]. The left images are original epipolar images : the x axis represent space, the y axis is time, and the level lines of each image are trajectories of physical points. In order to estimate the slope of each trajectory (from which the geometry of the underlying scene can be deduced), an anisotropic diffusion driven by the motion vector field is applied (right). Such a diffusion preserves the physical interpretation of the image generation process, and thus does not interfere with the scene reconstruction process.

7 Affine Scale Space

As we already mentionned, contrast invariant processing boils down to level set, and finally level curve processing. The above mentionned equations indeed are equivalent to curve evolution models, provided strong existence results are at hand. This is the case for the most important cases, namely the power 1, the so called "curve shortening" and the power 1/3, known as "affine shortening". Grayson [15] proved existence, uniqueness and analycity for the first equation,

$$\frac{\partial \mathbf{x}}{\partial t} = curv(\mathbf{x})n(\mathbf{x})$$

and Angenent, Sapiro and Tannenbaum [3] for the affine shortening

$$\frac{\partial \mathbf{x}}{\partial t} = curv(\mathbf{x})^{\frac{1}{3}}n(\mathbf{x}).$$

Those results are very relevant to image analysis as they ensure that the diffusion process indeed reduces the curve to a more and more sketchy version. We check the affine invariance in Figure 20 . We performed

Figure 19: The Affine and Morphological Scale Space (AMSS model). From left to right : original image, level lines of the images (16 levels only), smoothed image by using the affine and morphological scale space, and its level lines.

he numerical tests according to the fast and fully affine invariant numerical scheme described in [27], [20]. In the middle, the initial shape is an affine transform of the first one ; the shape on the right will be an inverse affine transform of the middle shape. If everything is correct, we can expect that, after processing, the shape on the right will be identical to the shape on the left. We make the experiment with both the the curve shortening and the affine shortening. So, it works !

Evans, Spruck [14] and Chen, Giga, Goto [11] proved in 1991 that a continuous function moves by curva-

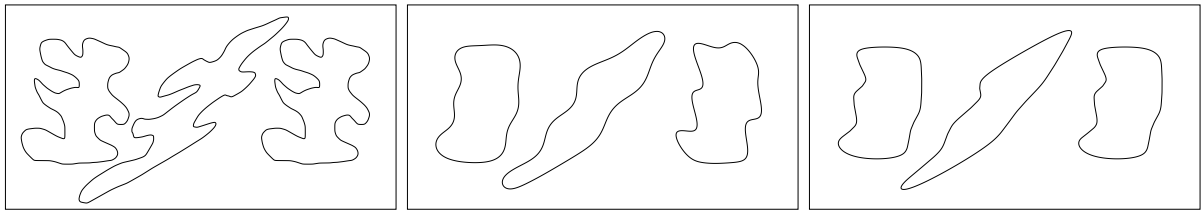

Figure 20: Experimental check of the affine invariance of the affine shortening (AMSS). We display on the left image three shapes. The second one is obtained by applying an affine transform A to the first shape S, yielding a shape $S' = AS$. The third one is obtained from the second by the inverse affine transform. It therefore initially is $A^{-1}S' = S$. On the right image : result after application of AMSS to the two first shapes : for some t, are viewed $S(t)$, $S'(t)$ and $A^{-1}(S'(t))$). If the numerical scheme is affine invariant, this third shape should coincide with $S(t)$, which is indeed the case. Middle : the same procedure applied with the curvature equation, which proves not to be affine invariant, as expected.

ture motion if and only almost all of its level curves move by curve shortening. This yields, in that case, a mathematical justification of the now classical Osher-Sethian numerical method for moving fronts by moving a distance function to the front. The same result is true for the affine invariant curve evolution. The Osher-Sethian point of view is just converse to the point of view adopted here : they associate with some curve C or surface its signed distance function u , so that the curve or surface is handled indirectly as the zero isolevel set of u. Then u is evolved by, say, the curvature motion with a classical numerical difference scheme. In that way, the curve evolution is dealt with efficiently and accurately. From our point view, the image can be viewed as a distance function to all and each of its level sets, since we are interested in all of them.
We show in Figure 19 an application of this numerical method, with both curvature and affine invariant curvature motions. In order to gain visibility, we do not display all level curves, but only for about eighteen levels. Notice that the aim is not here subsampling ; we keep the same resolution. It is not either restoration : the processed image is clearly worse than the original. The aim is invariant simplification leading to shape recognition.

8 Contours

Before proceeding to shape recognition, let us mention that a well adapted variant of curvature equation can be used for shape detection. It's a by now famous method of contour detection in an image, initially

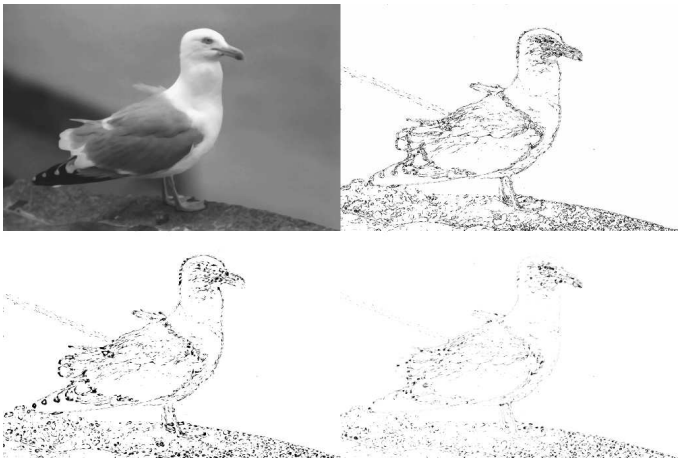

Figure 21: Computation of the curvature of the original sea bird image after it has been smoothed by curvature motion at calibrated scale 1. The first image displays the smoothed version of the sea bird at a small scale. The second image displays the absolute value of the curvature, with the convention that the darkest points have the largest curvature. We have displayed the curvature only at points where the gradient of the image was larger than 6. (The image grey levels range from 0 to 255). In continuation, we display on separate images the positive part of the curvature and the negative part. The curvature motion can be used as a nonlinear means to compute a "multiscale" curvature of the original image. Compare with Figure 22, where the calibrated scale of smoothing is 4 (a calibrated scale t means that at this scale a disk with radius t disappears).

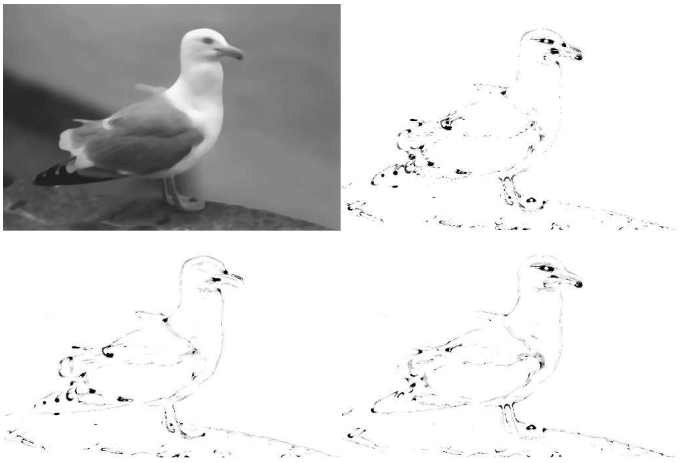

Figure 22: Same as Figure 21, except that now the calibrated scale of smoothing is equal to 4 (instead of 1)

proposed by Kass, Witkin and Terzopoulos. This method was very unstable and the winning method turns out to be a variant of curvature motion proposed by Caselles, Catté, Coll, Dibos [7] and improved simultaneously by Caselles, Kimmel, Sapiro [9], and Malladi, Sethian [23]. Here is how it works. The user draws roughly what are the contours he wants in the image and the algorithm then finds the best possible contour in terms of some variational criterion. This turns out to be very useful in medical imaging. The motion of the contour is a tuned curvature motion which tends to minimize the energy E which we will now explain. Given an original image u_0 containing some circular contours which we wish to approximate, we start with an "edge map"

$$g(\mathbf{x}) = \frac{1}{1 + |Du_0(\mathbf{x})|^2},$$

that is, a function which vanished on the edges of the image. The user then points out the contour he is interested in, by drawing a polygon γ_0 surrounding roughly the desired contour. The "geodesic snake" algorithm then builds a distance function v_0 to this initial contour, so that γ_0 is the zero level set of v_0.

The energy to be minimized is

$$E(\gamma) = \int_{\gamma} g(\mathbf{x}(s))ds,$$

where g is the edge map associated with the original image u_0 and s denotes the length parameter on γ. The motion of the "analysing image" v is governed by

$$\frac{\partial v}{\partial t} = g|Dv|curv(v) - Dv \cdot Dg.$$

We display an example in Figure 23.

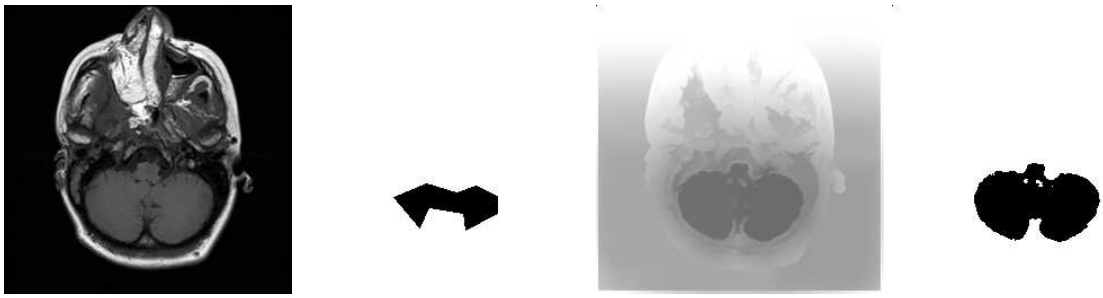

Figure 23: Active contour, or "snake" From left to right : original image, initial contour, evolved distance function, final contour.

The main obvious application of invariant PDE's seems to be shape retrieval in large databases. There are thousands of different definitions of shapes, and of shape recognition algorithms. Now, the real bottleneck has ever been extraction of the relevant shapes. The discussion above points to a brute force strategy : all contrast invariant local elements, are are the level lines of the image, are candidates to be "shape elements". Of course, this name of shape element suggests the contours of some object, but there is no way to give a simple geometric definition of objects. We must give up the hope of jumping from the geometry to the common sense world. We may instead simply ask the question : given two images, can we retrieve all level lines similar in both ? This would give a factual, a posteriori definition of shapes : they would be defined as pieces of level lines common to two different images, no matter what their relationships to real physical objects are. Of course, this brute force strategy would be impossible without the previous invariant fil-

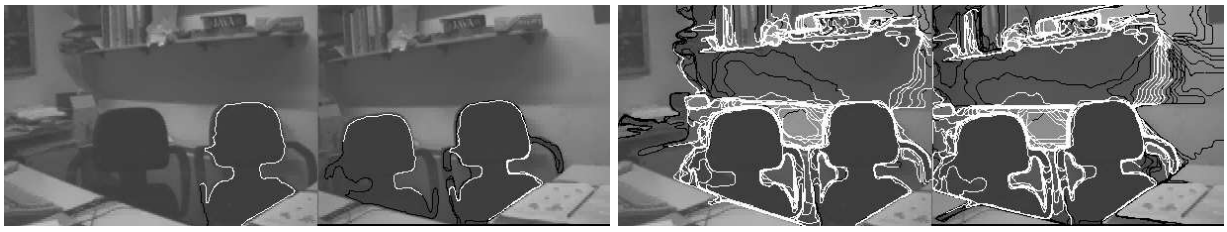

Figure 24: Level lines based shape parser. Shape extraction has ever been the bottleneck of shape recognition algorithms. With the presented algorithm, this problem is solved by a brute force method : it compares all level lines of the images to be compared. Left pair of images : two images of a desk taken from different angles. In the left hand desk image, one level line has been put in white. We display, also in white, in the right image of the pair, the matching level lines. The match is ambiguous, as must be expected when the same object is repeated twice in the scene ! In the right pair of images, we display in white all matching pairs of level lines. (Experiment : J.-L. Lisani).

tering (AMSS). It is instead doable if the level lines have been significantly simplified. This simplification entails the possibility of compressed invariant encoding. In Figure 24, we present an experiment due to Lisani et al. [22]. Two images of a desk are taken from different angles, and then, in white, all of the pieces of level lines in Image 1 and in Image 2 which found a match in the other image. In continuation, we present some of the matches. We notice that several of these matches are doubled : indeed, there are

two similar chairs in each images ! A Gestalt law comes immediately to mind. This law states that human perception tends to group similar shapes. We now see the numerical necessity of this perceptual grouping : a previous self-matching of each image, with grouping of similar shapes, must be performed before we can compare it to other images !

References

[1] L. Alvarez, Y. Gousseau, and J. M. Morel. Scales in natural images and a consequence on their BV norm. In *ScaleSpace99*, pages 247–258, 1999.

[2] L. Alvarez, P-L. Lions, and J-M. Morel. Image selective smoothing and edge detection by nonlinear diffusion (II). *SIAM Journal of numerical analysis*, 29:845–866, 1992.

[3] S. Angenent, G. Sapiro, and A. Tannenbaum. On the affine heat flow for nonconvex curves. *J. of the American Mathematical Society*, July 1998.

[4] P. J. Burt and E. H. Adelson. The laplacian pyramid as a compact image code. *IEEE Transactions on Communication*, 9(4):532–540, 1983.

[5] John F. Canny. A computational approach to edge detection. *PAMI*, 8:679–698, November 1986.

[6] R. A. Carmona and S. F. Zhong. Adaptive smoothing respecting feature directions. *IP*, 7(3):353–358, March 1998.

[7] V. Caselles, F. Catté, T. Coll, and F. Dibos. A geometric model for active contours in image processing. *Numerische Mathematik*, 66:1–31, 1993.

[8] V. Caselles, B. Coll, and J. M. Morel. A kanizsa programme. *Progress in Nonlinear Differential Equs. and their Applications*, 25, 1996.

[9] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 1(22):61–79, 1997.

[10] V. Caselles, J. M. Morel, and C. Sbert. An axiomatic approach to image interpolation. *IP*, 7(3):376–386, March 1998.

[11] Yun-Gang Chen, Yoshikazu Giga, and Shun'ichi Goto. Uniqueness and existence of viscosity solutions of generalized mean curvature flow equations. *J. Differ Geom. 33, No.3. [ISSN 0022-040X]*, 1991.

[12] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Am. Math. Soc., New Ser. 27, No.1. [ISSN 0273-0979]*, 1992.

[13] Evans. A geometric interpretation of the heat equation with multivalued initial data. *SIAM Journal of Mathematical Analysis*, 27:2:932–958, 1996.

[14] L. C. Evans and J. Spruck. Motion of level sets by mean curvature. I. *J. Differ. Geom. 33, No.3, 635-681 (1991). [ISSN 0022-040X]*, 1991.

[15] M. Grayson. The heat equation shrinks embedded plane curves to round points. *J. of Differential Geometry*, 26:285–314, 1987.

[16] Hitoshi Ishii and Panagiotis Souganidis. Generalized motion of noncompact hypersurfaces with velocity having arbitrary growth on the curvature tensor. *Tohoku Math. J., II. Ser. 47, No.2. [ISSN 0040-8735]*, 1995.

[17] S. Kichenassamy. The perona-malik paradox. *SIAM*, 57(5):1328–1342, 1997.

[18] Benjamin B. Kimia, Allen Tannenbaum, and Steven W. Zucker. On the evolution of curves via a function of curvature. I. the classical case. *Journal of Mathematical Analysis and Applications*, 163(2):438–458, 1992.

[19] J. J. Koenderink and A. J. van Doorn. Dynamic shape. *Biological Cybernetics*, 53:383–396, 1986.

[20] G. Koepfler and L. Moisan. Geometric multiscale representation of numerical images. In *ScaleSpace99*, pages 339–350, 1999.

[21] H. P Kramer and J. B Bruckner. Iterations of a non-linear transformation for enhancement of digital images. *Pattern Recognition*, 7, 1975.

[22] J. L. Lisani, L. Moisan, P. Monasse, and J. M. Morel. Affine invariant mathematical morphology applied to a generic shape recognition algorithm. In *Proceedings of International Symposium of Mathematical Morphology*, San Francisco, California, June 2000. available at `http://pascal.monasse.free.fr/index.html`.

[23] R. Malladi, J. A. Sethian, and B. C. Vemuri. Shape modeling with front propagation: a level set approach. *PAMI*, 17(2), 1995.

[24] D. Marr and E. Hildreth. Theory of edge detection. *Proc. Royal Soc. Lond.*, B 207:187–217, 1980.

[25] G. Matheron. *Random Sets and Integral Geometry.* John Wiley, N.Y., 1975.

[26] L. Moisan. Perspective invariant multiscale analysis of movies. In *International Symposium on Optical Science Engineering (SPIE'95, San Diego)*, volume 2567, pages 84–94, July 1995.

[27] L. Moisan. Affine plane curve evolution : a fully consistent scheme. *IEEE Transactions On Image Processing*, 7:411–420, March 1998.

[28] L Moisan. Pde's, motion analysis, and 3d reconstruction from movies. *Partial Differential Equations - Theory and numerical solution, Chapman & Hall/CRC Research Notes in Mathematics*, 406:273–282, 1998.

[29] P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. In *CVWS87*, pages 16–22, 1987.

[30] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D 60, No.1-4, 259-268. [ISSN 0167-2789]*, 1992.

[31] Leonid Iakov Rudin. *Images, Numerical Analysis of Singularities and Shock Filters.* PhD thesis, California Institute of Technology, 1987.

[32] J. Serra. *Image Analysis and Mathematical Morphology.* Academic Press, 1982.

[33] J. Sethian. Curvature and the evolution of fronts. *Comm. Math. Phys.*, 101, 1985.

[34] N. Sochen, R. Kimmel, and R. Malladi. A general framework for low-level vision. *IP*, 7(3):310–318, March 1998.

[35] L. Vincent. Grayscale area openings and closings, their efficient implementation and applications. In J. Serra and Ph. Salembrier, editors, *Proceedings of the* 1^{st} *Workshop on Mathematical Morphology and its Applications to Signal Processing*, pages 22–27, Barcelona, Spain, 1993.

[36] J. Weickert. Anisotropic diffusion filters for image processing based quality control. In A. Fasano and M. Primicerio, editors, *Proc. Seventh European Conf. on Mathematics in Industry*, pages 355–362. Teubner, Stuttgart, 1994.

[37] M. Wertheimer. Untersuchungen zur lehre der gestalt, II. *Psychologische Forschung*, 4:301–350, 1923.

[38] A. P. Witkin. Scale-space filtering. In *Proceedings of the International Joint Conference on Artificial Conference*, pages 1019–1021, 1983.