MORSE DESCRIPTION AND MORPHOLOGICAL ENCODING OF CONTINUOUS DATA*

VICENT CASELLES, GUILLERMO SAPIRO, ANDRÉS SOLÉ, AND COLOMA BALLESTER †

Abstract. A geometric representation for images is studied in this work. This is based on two complementary geometric structures for the topographic representation of an image. The first one computes a description of the Morse structure, while the second one computes a simplified version of drainage structures. The topographic significance of the Morse and drainage structures of Digital Elevation Maps (DEM) suggests that they can been used as the basis of an efficient encoding scheme. As an application we then combine this geometric representation with a consistent interpolation algorithm and lossless data compression schemes to develop an efficient compression algorithm for DEM. This coding scheme controls the L^∞ error in the decoded elevation map, a property that is necessary for the majority of applications dealing with DEM. We present the underlying theory and some compression results for standard DEM data.

Key words. Mathematical Morphology, Morse theory, drainage structures, connected components, interpolation, encoding, compression

AMS subject classifications. 68U10, 54C30, 35J25

1. Introduction. A geometric approach to representing and compressing Digital Elevation Maps (DEM) is proposed in this paper. Our approach is based on Morse theory and the computation of drainage structures, which lead to an efficient representation of the topographic structures of these images. This paper presents the underlying theoretical results and some experimental examples, which are further developed in [48].

DEM data consist of a discrete digital representation of a surface terrain. Each cell in a DEM corresponds to a point (x, y, z) in 3D space. We can think of (x, y) as the coordinates in the image domain and the height z as the gray value of the image (see Fig. 1.1). The acquisition systems used to obtain a DEM have been improved during the last years in order to obtain a better resolution both in the coordinate plane and in height. Typically a DEM image from a small terrain has 1200×1200 points, that is 1440000 bytes ($1.4Mb$) when using 8 bits for the height, or 2880000 ($2.8Mb$) when using 16 bits. If we note that for a complete terrain description of a country we need thousands of these images, storing and transmitting them requires efficient encoding and compression.

Many algorithms exist for lossless and lossy data compression [40, 56, 55]. Typically, lossy compression algorithms control the L^2 norm of the error (the root mean square error), but it is not so easy to find algorithms which allow a control on the L^∞

* This paper is in part in the PhD Dissertation of Andrés Solé. The paper was under review when Andrés departed from this world. We (VC, GS, and CB) want to dedicate to him our contribution to this work. We would like to thank the reviewers for their very useful suggestions. The first, third and fourth authors acknowledge partial support by the Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya and by PNPGC project, reference BFM2000-0962-C02-01. G. Sapiro is partially supported by a grant from the Office of Naval Research ONR-N00014-97-1-0509, the Presidential Early Career Award for Scientists and Engineers (PECASE), and a National Science Foundation CAREER Award. G. Sapiro thanks Dr. Wen Masters from ONR and Dr. Carey Schwartz from DARPA for introducing him to the problem of DEM compression.

†VC,AS (deceased), and CB are with the Dept. de Tecnologia, University of Pompeu-Fabra, Passeig de Circumvalacio, 8, 08003 Barcelona, Spain (`(vicent.caselles,coloma.ballester)@upf.edu`). GS is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455, USA (`guille@ece.umn.edu`).

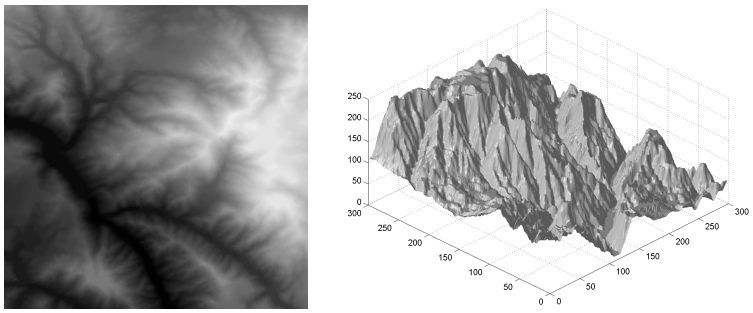

FIG. 1.1. *Left: original DEM images. Right: 3D representation of the image on the left.*

norm of the error (that is the sup error). This is fundamental for DEM applications. Without an L^∞ error control the error in individual pixels may be of the same order of magnitude as the image gray value resolution. For DEM applications, e.g., navigation and landing, this leads to an error in terrain height that makes the algorithm forbidden. A standard algorithm allowing the desired L^∞ control is JPEG Lossless (JPEG-LS) [56, 55, 35], which has a near lossless mode where one can impose the maximum allowed error. The use of JPEG-LS for DEM data has been studied in [35]. Let us mention that if you compress an image with JPEG-2000 [1] and then encode the errors greater than a given bound this algorithm can also be used to compress with control of the maximum error (being this not a fully standard approach).

To represent and compress DEM data it seems reasonable to store only those geometric structures which are of special relevance such as its Morse and drainage structures. Morse theory describes the topological change of the isocontours of scalar data or height function as the height varies, and relates these topological changes to the criticalities of the function. Critical levels permit to decompose the graph of a function into (truncated) cone like pieces. This suggests the idea of storing the level curves corresponding to critical levels, and use a consistent interpolation algorithm to recover the corresponding cone-like structures. Applying this strategy directly would lead to high errors between the original image and its interpolated reconstruction, and the coding of these errors (or a quantized version of them) would be too expensive. Thus we propose to add other topographically relevant information, a simplified drainage structure network. Then, for DEM representation we propose to first encode the level curves corresponding to critical levels, together with a simplified version of drainage structures. We then use this geometric sampling to compute an approximation of the data by interpolating them, compute the errors, and encode a quantized version of them so that we control the L^∞ norm of the error. Then these series of data structures are compressed by an arithmetic coder. This paper presents the underlying theoretical and computational justification of this method, together with some experimental results (for additional experiments, see [48]).

Let us mention that the use of a topographic description of images, surfaces, or $3D$ data has been used in different areas of research, including image processing, computer graphics, and geographic information systems (GIS), e.g., [7, 8, 9, 6, 5, 11, 15, 17, 20, 27, 52, 28, 33, 32, 42, 46]. The motivations for such a description differ depending on the field of application. In all cases these descriptions aim to achieve an efficient description of the basic shapes in the given image and their topological changes as a function of a physical quantity that depends on the type of data (height

in data elevation models, intensity in images, etc.). We shall comment on related work in Section 5 after the description of our main contributions.

Let us the structure of the paper. Section 2 introduces some notation to be used in the paper. Section 3 is the theoretical core of the paper. First we introduce the notion of monotone and maximal monotone section, which represent the truncated cone-like structures of the image without criticalities. Related to them is the notion of singular value, the values where a maximal monotone section begins or ends. Next, to justify the algorithm that will compute the singular values, we define in Subsection 3.2 the notion of critical value and we state the result that both notions are equivalent (the proof is included in Appendix 11). As a main consequence, we propose a simple combinatorial algorithm computing the maximal monotone sections of an image. This is the object of Section 4. Section 5 contains a brief description of related work. Section 6 will be devoted to the computation of a simplified version of *drainage* structures. In Section 7 we discuss some properties of the interpolation algorithms used to recover an approximation of the original data from the sampled one. In Sections 8 and 9 we collect the previously developed algorithms and we apply them to build-up an algorithm for compression of DEM data. We also briefly compare it with existing compression algorithms, mainly with JPEG Lossless (JLS) and JPEG-2000. Finally, in Section 10 we summarize the main conclusions of this work. In Sections 11 and 12 we include as appendix the proofs of the main results stated throughout the text.

2. Some notation. Let $\overline{\Omega}$ be a set homeomorphic to the closed unit ball of R^N ($N \geq 2$), $\left\{x \in R^N, \|x\| \leq 1\right\}$, and Ω be the interior of $\overline{\Omega}$. Note that, in particular, $\overline{\Omega}$ is compact, connected and locally connected. Even though some of the results in this paper could be proved for more general sets, we shall assume that $\overline{\Omega}$ is of this form.

Let $u : \overline{\Omega} \to R$ be a function. We call upper (lower) level set of u any set of the form $[u \geq \lambda] := \{x \in \overline{\Omega} : u(x) \geq \lambda\}$ or $[u > \lambda] := \{x \in \overline{\Omega} : u(x) > \lambda\}$ ($[u \leq \lambda] := \{x \in \overline{\Omega} : u(x) \leq \lambda\}$ or $[u < \lambda] := \{x \in \overline{\Omega} : u(x) < \lambda\}$), where $\lambda \in R$. For each $\lambda, \mu \in R$, $\lambda \leq \mu$ we define $[\lambda \leq u \leq \mu] = \{x \in \overline{\Omega} : \lambda \leq u(x) \leq \mu\}$, and we write $[u = \lambda] = [\lambda \leq u \leq \lambda]$. The connected components of a set $X \subseteq R^N$ will be denoted by $\mathcal{CC}(X)$. If Y is a connected set contained in a set X, the connected component of X containing Y will be denoted by $\mathcal{CC}(X, Y)$. In particular, if $Y = \{x\}$ is reduced to a point, we shall use the notation $\mathcal{CC}(X, x)$.

The space of continuous functions in a compact subset K of R^N will be denoted by $C(K)$. We shall denote by $\overline{X}$ and $int(X)$ the closure and interior, respectively, of a set $X \subseteq R^N$. A continuum is a compact connected set. We shall say that two sets $A, B \subseteq X$ are connected inside X by a continuum if there is a continuum $C \subseteq X$ such that $A \cup B \cup C$ is connected. The Lebesgue measure of a set $X \subseteq R^N$ will be denoted by $|X|$.

3. Morse theory: Monotone sections, singular and critical values. As we explained in the Introduction, the aim of Morse theory is to describe the topological changes of the (iso)level sets of a real valued function in terms of its critical points. Our purpose in this Section is to describe two different notions of critical values (one of them will be called critical value, the other singular value) and prove that they are equivalent. One of those notions, the one of singular value, is intuitively related to the classical notion of critical value for a smooth function (points where the gradient vanishes), the other notion called critical value will permit us to describe a simple and efficient algorithm to compute them. When $N = 2$, this algorithm computes the Morse structure of the image from its upper and lower level sets.

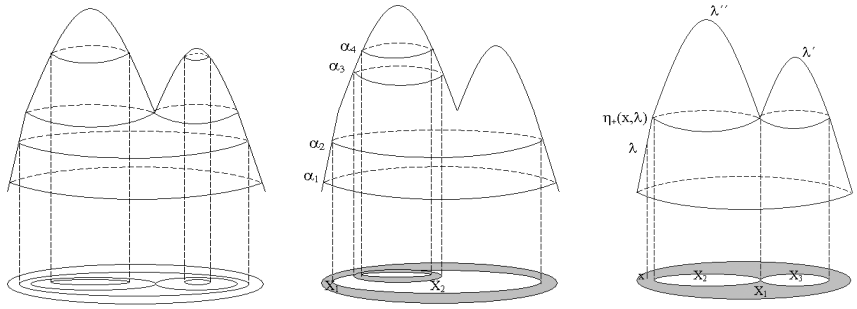

FIG. 3.1. *a) As λ decreases and crosses a maximum a level curve appears. When crossing a saddle point, the isolevel curves $[u = \lambda]$ merge. b) The sets $[\alpha_1 \leq u \leq \alpha_2]$ and $[\alpha_3 \leq u \leq \alpha_4]$ are monotone sections. c) X_1,X_2,X_3 are three maximal monotone sections. X_1 is the maximal monotone section that contains the point x such that $u(x) = \lambda$, and it extends from $\eta_+(x, \lambda)$ until $\eta_-(x, \lambda)$ (we have not displayed this value). In this case, the isolevel set $[u = \eta_+(x, \lambda)]$ is contained in X_1 but it is not contained in X_2 or X_3.*

3.1. Monotone sections and singular values. We shall refer to the topographic map of u as the family of connected components of the isolevel sets of u, $[u = \lambda]$, $\lambda \in R$. The idea of topographic map (though in a slightly different form) was introduced in [11]. In [7],[8], using different approaches, the authors initiated the study of the Morse structure of the topographic map for continuous functions, and bounded upper semicontinuous functions, respectively. They defined the notion of monotone section as a region of the topographic map between two levels containing no topological changes of the topographic structure. We shall follow the approach in [7].

DEFINITION 3.1. *Let $u : \overline{\Omega} \to R$ be a continuous function. A monotone section of the topographic map of u is a set of the form*

$$X_{\lambda,\mu} \in \mathcal{CC}([\lambda \leq u \leq \mu]), \tag{3.1}$$

for some $\lambda, \mu \in R$ with $\lambda \leq \mu$, such that for any $\alpha \in [\lambda, \mu]$, the set $X_{\lambda,\mu} \cap [u = \alpha]$ is a connected set.

Let us explain this definition. The notion of monotone section tries to capture a region of the topographic map of the image u where there is no change of topology, meaning no change in the number of connected components of its isolevel sets $[u = \alpha]$, $\alpha \in R$. To get a better idea, let us assume that $u : \overline{\Omega} \to [a, b]$, $\overline{\Omega} \subseteq R^2$, $a < b$, is smooth and each isolevel set $[u = \lambda]$ is a family of curves. The critical points of u are its maxima, minima and saddle points. If the function is smooth, we may identify these critical points using the usual rules of differential calculus. But there is also a topological description of them, which can be called its Morse description. We look at the isolevel sets $[u = \alpha]$ as α increases from a to b. Notice that if there is a minimum (resp. a maximum) at level α then a small curve appears (resp. disappears), and if there is a saddle point at level α there is some bifurcation in the curve, i.e., two curves merge, or a single curve splits into two (see Figure 3.1.a). Thus, if we see the sets $[u = \alpha]$ as a family of moving curves, at the critical points, one of such curves appears, disappears, splits or merges. Then a connected components X of a set $[\lambda < u < \mu]$, $\lambda < \mu$, could be called monotone section if X contains no critical point of u (see Figure 3.1.b). Notice that, if X is such a monotone section, then $X \cap [u = \alpha]$ is connected (a

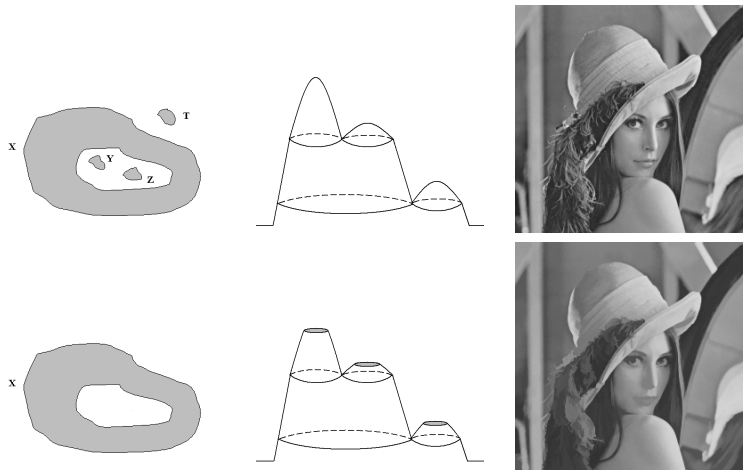

FIG. 3.2. *a) Top left: Original image with a connected component X with area $\geq \epsilon$, and three connected components Y, Z, T of area $< \epsilon$. b) Bottom left: Result of the extrema filter T_ϵ^+ applied to the image in a). c) Top middle: The graph of a simple image, d) Bottom middle: the effect of the extrema filter T_ϵ^+ on the image in c). We see that local maxima are decreased until the area of the connected component containing it attains an area of ϵ or greater. The effect of the filter T_ϵ^- on the local minima of the an image is similar, and we do not display it here. e) Top right: original image. f) Bottom right: result of filtering the image in e) using $T_\epsilon^- T_\epsilon^+$ with $\epsilon = 100$.*

connected curve). Our definition of monotone section is inspired by this observation. Our purpose now is to find a partition of the image domain into its largest monotone sections, and the boundaries of these regions will be the curves corresponding to the critical values.

Due to small oscillations in the image, its Morse structure is too complex, i.e., there are too many criticalities. To simplify the structure of the topographic map while preserving its main features we filter the image with the Vicent-Serra filters [54],[53], also called extrema filters [19]. Extrema filters eliminate the small connected components of upper and lower level sets of the given image [7], [12]. The resulting image has a simplified topographic map structure. Let us recall the definition of such a filter, first on sets (which is equivalent to binary images), then on functions. If X is a set in R^N, and $\epsilon > 0$, we define $T_\epsilon(X) = \cup_{Y \in \mathcal{CC}(X):|Y| \geq \epsilon} Y$, i.e., the union of connected components of X of measure $\geq \epsilon$ (see Figure 3.2.a,b). With this definition we may define the following operators on images: $T_\epsilon^+ u(x) = \sup\{\lambda : x \in T_\epsilon([u \geq \lambda])\}$, $T_\epsilon^- u(x) = \inf\{\lambda \in R : x \in T_\epsilon([u < \lambda])\}$, $x \in D$. The effect of T_ϵ^+ (resp. T_ϵ^-) is to eliminate the connected components of the upper level sets (resp. of the lower level sets) of measure $< \epsilon$, and, as a consequence, the effect on resulting image is that the local maxima (resp. minima) are filtered as reflected in Figure 3.2.c,d. To filter images, we apply both filters iteratively $T_\epsilon^+ T_\epsilon^- u$, or $T_\epsilon^- T_\epsilon^+ u$ (they do not commute, see [12], for instance). The effect of the filter on an image is displayed in Figure 3.2.e,f. If there is some $v \in C(\overline{\Omega})$ such that $u = T_\epsilon^+ T_\epsilon^- v$, or $u = T_\epsilon^- T_\epsilon^+ v$, and X is a connected component of $[u \geq \lambda]$, or $[u < \lambda]$, then $|X| \geq \epsilon$ [7],[12]. This implies that, for any $\lambda \leq \mu$, there are a finite number of connected components of $[\lambda \leq u \leq \mu]$. In particular, there are a finite number of connected components of $[u \geq \lambda]$ and $[u < \lambda]$ [7],[12]. In all what follows and for the purposes of mathematical justification we shall assume that

(**H**) $u \in C(\overline{\Omega})$ *and it has been filtered with the extrema filters, i.e., there is some*

$v \in C(\overline{\Omega})$ *such that* $u = T_\epsilon^+ T_\epsilon^- v$, *or* $u = T_\epsilon^- T_\epsilon^+ v$.

The definition we have given of monotone section is slightly different from the one given in [7], but, according to next Proposition, equivalent to it.

PROPOSITION 3.2. *Let* $u : \overline{\Omega} \to R$ *be a continuous function such that for each* $\lambda, \mu \in R$ *with* $\lambda \leq \mu$ *the set* $[\lambda \leq u \leq \mu]$ *has a finite number of connected components. Let* $\lambda < \mu$, $X \in \mathcal{CC}([\lambda \leq u \leq \mu])$. *Then* X *is a monotone section if and only if for any* $\lambda', \mu' \in [\lambda, \mu]$, $\lambda' \leq \mu'$, *the set* $\{x \in X : \lambda' \leq u(x) \leq \mu'\}$ *is a connected component of* $[\lambda' \leq u \leq \mu']$.

Proof. Let us denote by $(*)$ the assertion that for any $\lambda', \mu' \in [\lambda, \mu]$, $\lambda' \leq \mu'$, the set $\{x \in X : \lambda' \leq u(x) \leq \mu'\}$ is a connected component of $[\lambda' \leq u \leq \mu']$. Obviously, if X satisfies $(*)$, then for any $\alpha \in [\lambda, \mu]$, the set $X \cap [u = \alpha]$ is connected. To prove the converse statement, suppose that X is a monotone section but $(*)$ does not hold. Then there are values $\lambda \leq \alpha \leq \beta \leq \mu$ such that $X \cap [\alpha \leq u \leq \beta]$ is not connected. Observe that $X \cap [\alpha \leq u \leq \beta] \neq \emptyset$ because $\emptyset$ is connected. Let $c = \frac{\alpha+\beta}{2}$. Since $X \cap [\alpha \leq u \leq \beta] = (X \cap [\alpha \leq u \leq c]) \cup (X \cap [c \leq u \leq \beta])$ and $(X \cap [\alpha \leq u \leq c]) \cap (X \cap [c \leq u \leq \beta]) \neq \emptyset$, we conclude that either $X \cap [\alpha \leq u \leq c]$ or $X \cap [c \leq u \leq \beta]$, or both, cannot be connected (since the union of intersecting connected sets is also connected). Let us choose one of the above non connected sets and denote it by $X \cap [\alpha_1 \leq u \leq \beta_1]$. Proceeding iteratively in this way we find a decreasing sequence of intervals $[\alpha_n, \beta_n]$ such that $\cap_n [\alpha_n, \beta_n] = \{\gamma\}$ and $X \cap [\alpha_n \leq u \leq \beta_n]$ are not connected. Now, observe that for all $Y_n \in \mathcal{CC}(X \cap [\alpha_n \leq u \leq \beta_n])$ there exists $Y_{n-1} \in \mathcal{CC}(X \cap [\alpha_{n-1} \leq u \leq \beta_{n-1}])$ so that $Y_n \subseteq Y_{n-1}$. Thus, there are at least two different decreasing sequences of continua (compact connected sets) contained in $X \cap [\alpha_n \leq u \leq \beta_n]$. Since the intersection of a decreasing sequence of continua is a continuum ([26]), we conclude that $X \cap [u = \gamma] = \cap_n (X \cap [\alpha_n \leq u \leq \beta_n])$ is not connected. This contradiction proves the proposition. □

The following result which was proved in [7] permits us to define a monotone section which is maximal with respect to inclusion. Those sets are the non singular sets we mentioned above.

PROPOSITION 3.3. *([7], Proposition 1) Assume that* $u : \overline{\Omega} \to R$ *is a continuous function such that for each* $\lambda, \mu \in R$ *with* $\lambda \leq \mu$ *the set* $[\lambda \leq u \leq \mu]$ *has a finite number of connected components. Let* $\lambda_1, \lambda_2, \mu_1, \mu_2 \in R$. *Then, if* $X_{\lambda_1,\lambda_2}, X_{\mu_1,\mu_2}$ *are monotone sections such that* $X_{\lambda_1,\lambda_2} \cap X_{\mu_1,\mu_2} \neq \emptyset$, *then* $X_{\lambda_1,\lambda_2} \cup X_{\mu_1,\mu_2}$ *is also a monotone section. In other words, the union of intersecting monotone sections is a monotone section.*

As a consequence, assuming property (**H**), the union of monotone sections which intersect is a monotone section. This permits to define the notion of maximal monotone section containing a given point. Indeed, let $x \in D$ and $\lambda = u(x)$, and take $X_{\lambda,\lambda} = \mathcal{CC}([u = \lambda], x)$ (intuitively, the level curve through x). Then we try to increase X_λ going upwards (resp. downwards) until we find a critical level, call it $\eta_+(x, \lambda)$ (resp. $\eta_-(x, \lambda)$). With this we shall find two monotone sections containing x, and whose union is the largest monotone section containing x (see Figure 3.1.c). Let us precise this. For each $\eta \geq \lambda$, let $X_{\lambda,\eta} = \mathcal{CC}([\lambda \leq u \leq \eta], x)$. We define

$$\eta_+(x, \lambda) = \sup\{\eta : \eta \geq \lambda, \text{ s. t. } X_{\lambda,\eta} \text{ is a monotone section}\}.$$

Similarly, we define

$$\eta_-(x, \lambda) = \inf\{\eta : \eta \leq \lambda, \text{ s. t. } X_{\eta,\lambda} \text{ is a monotone section}\}.$$

Note that both numbers are well defined since $X_{\lambda,\lambda}$ is always a monotone section. Note also that, by definition, $\eta_-(x,\lambda) \leq \eta_+(x,\lambda)$. Those numbers determine an (open, closed, half-open, half-closed) interval $I(x,\lambda)$ containing λ whose end-points are $\eta_-(x,\lambda)$, $\eta_+(x,\lambda)$, which gives a monotone section containing x maximal with respect to inclusion (see [7]).

At least, when $N = 2$, maximal monotone sections represent the largest sections of the topographic map containing no topological changes (see Fig. 3.1.c), and our purpose is to compute them. Intuitively, monotone sections are topologically equivalent to truncated cones and the maximal monotone ones are the largest truncated cones contained in the graph of the image. This property is well adapted to the interpolation algorithms described in Section 7, which are able to re-interpolate truncated cones from the curves bounding them.

DEFINITION 3.4. *Let $M \subseteq \overline{\Omega}$. We say that M is a zonal maximum (resp., minimum) of u at height λ if M is a connected component of $[u = \lambda]$ and, for all $\epsilon > 0$, the set $[\lambda - \epsilon < u \leq \lambda]$ (resp., $[\lambda \leq u < \lambda + \epsilon]$) is a neighborhood of M.*

DEFINITION 3.5. *We say that $\lambda \in R$ is a singular value if it corresponds to a zonal maximum, minimum, or it corresponds to a level where it begins or ends a maximal monotone section, i.e., there is a point $x \in \overline{\Omega}$ such that $\eta_+(x,\lambda) = \lambda$ or $\eta_-(x,\lambda) = \lambda$.*

Remark. Observe that the definition of singular value is self-dual, in the sense that, λ is a singular value of u if and only if $-\lambda$ is a singular value of $-u$.

3.2. Critical values. This section is devoted to the definition of critical values. Before going into the formal definition, let us explain the idea behind it. Intuitively, critical levels are levels where a connected component of an isolevel set $[u = \alpha]$ appears, disappears, splits, or merges. Those critical levels can be identified by looking at $[u = \alpha]$ as the common boundary of two sets $[u \geq \alpha]$ and $[u < \alpha]$, and describing the topology of $[u = \alpha]$ in terms of the connected components of the sets $[u \geq \alpha]$ and $[u < \alpha]$. Indeed, if we increase α and we cross a level with a minimum (resp. maximum), then a connected component of $[u < \alpha]$ appears (resp. a connected component of $[u \geq \alpha]$ disappears). If we cross a saddle point then either two connected components of $[u < \alpha]$, or of $[u \geq \alpha]$, merge. This will give us a simple way to compute the critical levels (see Section 4). Then we shall prove the result that both notions of critical and singular values are equivalent, and this will prove that we are indeed computing the maximal monotone sections of the image.

Let us introduce the notion of critical level of u. For that we first need the following definition.

DEFINITION 3.6. *Let $\lambda \in R$. Assume that* (**H**) *holds. A signature of the level set $[u \geq \lambda]$ (resp. $[u < \lambda]$) consists of a finite family of marker points $\{p_i : i = 1, ..., r\}$ (resp. $\{q_j : j = 1, ..., s\}$) such that each p_i (resp. q_j) is a point in a different connected component $X^{\lambda,i}$ (resp. $X_{\lambda,j}$) of $[u \geq \lambda]$ (resp. $[u < \lambda]$). The points p_i, q_j are selected as maximum points of u in $X^{\lambda,i}$, and minima of u in $X_{\lambda,j}$, i.e.,*

$$u(p_i) = \sup_{x \in X^{\lambda,i}} u(x), \text{ and } \quad u(q_j) = \inf_{x \in X_{\lambda,j}} u(x).$$

Having chosen the marker points, we denote the signature of $[u \geq \lambda]$ by $sig([u \geq \lambda])$, the signature of $[u < \lambda]$ by $sig([u < \lambda])$. We define $sig([u \geq \lambda], [u < \lambda]) = sig([u \geq \lambda]) \cup sig([u < \lambda])$.

It is important to notice that the signature of an upper (resp., lower) level set is not unique since there may be many different points in a connected component

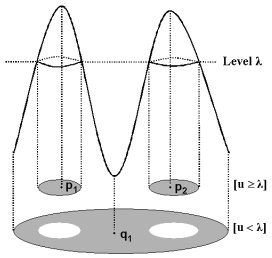

FIG. 3.3. *A function u and its upper and lower level sets at level λ with its assigned signature. The set $[u \geq \lambda]$ has two connected components depicted in gray and its signature consists of two points $\{p_1, p_2\}$. The set $[u < \lambda]$ has only one connected component with two holes, and is depicted as the circular region below with the two white holes. Its signature consists of the point q_1.*

of it where the function u attains its maximum (resp., its minimum). The problems posed by this can be avoided using the following criterion which will be used in the computational algorithm introduced in Section 4: We define the signature of $[u \geq \lambda]$ (resp. $[u < \lambda]$) in a decreasing (resp., increasing) order in the value of $\lambda \in R$. For the moment being we shall consider two signatures S_1 and S_2 of $[u \geq \lambda]$ (resp. $[u < \lambda]$) equal it they have the same number of marker points and for any point $p \in S_1$ there is a point $q \in S_2$ which is a maximum (resp. a minimum) of u in $\mathcal{CC}([u \geq \lambda], p)$ (resp., in $\mathcal{CC}([u < \lambda], p)$), and conversely, interchanging S_2 and S_1. In this way we can give a unique sense to $sig([u \geq \lambda], [u < \lambda])$ for any $\lambda \in R$. The definition of signature is illustrated in Figure 3.3.

Our next Lemma proves that the signature may only change from above. This justifies the definition of critical value that follows it.

LEMMA 3.7. *Let $\lambda \in R$. There is $\epsilon > 0$ such that $sig([u \geq \mu], [u < \mu])$ is constant for all $\mu \in (\lambda - \epsilon, \lambda]$.*

Proof. Let $X^{\lambda,i}$, $X_{\lambda,j}$, $i = 1, ..., r$, $j = 1, ..., s$, be the the family of connected components of $[u \geq \lambda]$, resp. $[u < \lambda]$ with the markers p_i, q_j defined above. Let $i \in \{1, ..., r\}$. For each $\mu < \lambda$, let $X^{\mu,i}$ be the connected component of $[u \geq \mu]$ containing $X^{\lambda,i}$. Then, obviously, we have $X^{\lambda,i} \subseteq \cap_{\mu<\lambda} X^{\mu,i}$. Now, since as $\mu \uparrow \lambda$ $X^{\mu,i}$ is a decreasing family of continua, their intersection is also a continuum [26]. Moreover, it is contained in $[u \geq \lambda]$. Therefore,

$$\cap_{\mu<\lambda} X^{\mu,i} \subseteq \mathcal{CC}([u \geq \lambda], p_i) = X^{\lambda,i},$$

and we have the equality of both sets. As a consequence, there is an $\epsilon > 0$ such that for each $\mu \in (\lambda - \epsilon, \lambda]$, the sets $X^{\lambda,i}$, $i = 1, ..., r$, are contained in different connected components of $[u \geq \mu]$. Moreover, it cannot exist a sequence of values $\mu_n \uparrow \lambda$ such that $[u \geq \mu_n]$ has a connected component Q_n disjoint to $[u \geq \lambda]$. In that case $[u \geq \mu_n] \supseteq [u \geq \lambda] \cup Q_n$, and, thus $|[u \geq \mu_n]| \geq |[u \geq \lambda]| + \delta$ for some $\delta > 0$. Since $[u \geq \mu_n] \downarrow [u \geq \lambda]$, also $|[u \geq \mu_n]| \downarrow |[u \geq \lambda]|$. This contradiction proves that there is an $\epsilon > 0$ such that for each $\mu \in (\lambda - \epsilon, \lambda]$ the set $[u \geq \mu]$ consists of r connected components, each one of them containing a different component of $[u \geq \lambda]$.

Let $\mu_n \uparrow \lambda$. Again, using that $\cup_n [u < \mu_n] = [u < \lambda]$, for n large enough, we have that $[u < \mu_n] \cap X_{\lambda,j}$, $j = 1, ..., s$, are the connected components of $[u < \mu_n]$. We conclude that there is an $\epsilon > 0$ such that $sig([u \geq \mu], [u < \mu])$ is constant for each $\mu \in (\lambda - \epsilon, \lambda]$.□

DEFINITION 3.8. *We say that $\lambda \in R$ is a critical value for u if there is a sequence $\mu_n \downarrow \lambda$ such that $sig([u \geq \mu_n], [u < \mu_n]) \neq sig([u \geq \lambda], [u < \lambda])$ for each $n = 1, 2, ...$*

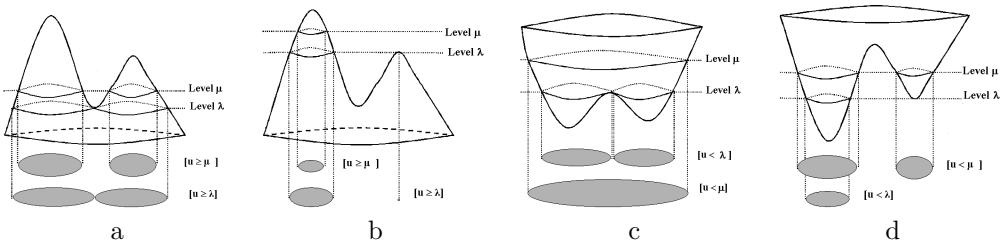

FIG. 3.4. *a) For each level μ we compute the connected components of $[u \geq \mu]$. At level $\mu > \lambda$ there are two components of $[u \geq \mu]$ which increase and merge as μ decreases to λ. b) While decreasing the level μ, we see that at level λ there is a new local maximum of u and a new connected component of $[u \geq \lambda]$ appears. c) For each level μ we compute the connected components of $[u < \mu]$. At level λ there are two components of $[u < \lambda]$ which merge for any level $\mu \geq \lambda$. d) Observe that at level λ there is only a connected component of $[u < \lambda]$ but at any level $\mu > \lambda$ a new connected component of $[u < \mu]$ appeared.*

Let us explain the phenomena reflected by a change of signature. Let $\lambda < \mu$. If $sig([u \geq \lambda], [u < \lambda]) \neq sig([u \geq \mu], [u < \mu])$, then either (a) $sig([u \geq \lambda]) \neq sig([u \geq \mu])$, or (b) $sig([u < \lambda]) \neq sig([u < \mu])$. Suppose that (a) holds. Since $[u \geq \mu] \subseteq [u \geq \lambda]$ each connected component of $[u \geq \mu]$ is contained in a component of $[u \geq \lambda]$. Two critical phenomena may happen: (a_1) there are two different components of $[u \geq \mu]$ which are contained in the same component of $[u \geq \lambda]$, i.e., two connected components of $[u \geq \mu]$ merged at level λ, or (a_2) there is a component of $[u \geq \lambda]$ with no component of $[u \geq \mu]$ contained in it, i.e., a new connected component of the upper level sets appeared at level $[u \geq \lambda]$. Suppose that (b) holds. Since $[u < \lambda] \subseteq [u < \mu]$, each component of $[u < \lambda]$ is contained in a component of $[u < \mu]$. Again, the same two critical phenomena may happen: (b_1) there are two different connected components of $[u < \lambda]$ which are contained in the same component of $[u < \mu]$, i.e., a connected component of $[u < \mu]$ has splitted at level λ, (b_2) there is a component of $[u < \mu]$ with no component of $[u < \lambda]$ contained in it, i.e., a connected component of the lower level sets is present at level μ while it was absent at level λ. (see Fig. 3.4). We call the first type of criticality of upper type, while the second is called of lower type.

Let us state the fundamental theoretical result of this paper which justifies our algorithm in next section. The proof will be given in Appendix 11.

THEOREM 3.9. *Assume that* (**H**) *holds. Let $\lambda \in R$. Then λ is a critical level of u if and only if λ is a singular level of the topographic map of u.*

4. The computational algorithm. For the sake of simplicity we only describe the algorithm to compute the critical values of upper type. The critical values of lower type can be computed using the same algorithm applied to the inverted image $\max(u) - u$. Note that it is possible to compute both types of criticalities at the same time. Anyway, the computational cost will be similar either if we compute both type criticalities separately or at the same time. As we will see in Algorithm 1, the computational cost derives mainly from the computation of the connected components at each level. Since we perform this computation by means of a region growing strategy, it makes no difference to do both computations (upper and lower connected components) simultaneously or not.

We have observed that, due to low oscillations in the image, a huge number of criticalities may appear. In order to select the most relevant ones we filter the image with extrema filters, which simplify but do not distort the topographic map structure

[7, 12]. We stress here the fact that this pre-filtering step is only done to compute the relevant criticalities, but the compression algorithm (see Section 8) is applied to the original (not to the filtered) image. After the above pre-filtering step (which could also not be used), our strategy consists on discarding some critical sections which do no satisfy a minimum contrast criterion. The contrast criterion is specified by means of a parameter $MinContrast$.

We assume that our image u ranges from $m = min(u)$ to $M = max(u)$ (for example from 0 to 255). Let us give some additional explanations before making explicit the algorithm taking into account the contrast criterion. Our algorithm computes critical levels of upper type, choosing only a subset of them in such a way that the regions comprised between two consecutive critical levels have sufficient contrast. The contrast of u on a set X is the oscillation of the gray level there, i.e., the difference $\max_{x\in X} u(x) - \min_{x\in X} u(x)$. We shall refer to a contrasted M-section as a union of monotone sections with contrast $\geq MinContrast$. Our algorithm starts from $\lambda = M$ and decreases the values of λ computing the connected components of $[u \geq \lambda]$ and its markers (let us denote them by X^{λ,p_i}). When a monotone section is born the contrast is set to 0. Then we decrease λ until we find a critical level. If λ is such a critical level and X^{λ,p_i} is a maximal monotone section ending at this level whose contrast is lower than $MinConstrast$, we do not declare it as critical and we continue growing it by decreasing λ. If at some lower value of λ we find again a critical level, we shall declare it as critical only in the case that the contrast of the growing section is higher than $MinContrast$. In the case we declare λ as a critical level, a new monotone section starts and we have to store this level to compute the contrast when we arrive at the next critical level. This will permit us to compute the critical levels, and to choose a subset of them giving us contrasted M-sections. If the parameter $MinContrast = 0$, the algorithm gives us all critical levels of upper type. Applying the algorithm to u and to $\max(u) - u$ we obtain all critical levels and all monotone sections.

To follow the above strategy, we shall need two quantities: β^{λ,p_i} which is the starting level of the contrasted M-section being computed, and the contrast of X^{λ,p_i}, defined by $C(X^{\lambda,p_i}) = \beta^{\lambda,p_i} - \lambda$. To store the computed connected components we use a dynamical data structure L consisting on a vector, ranging from m to M, of lists of couples $(X^{\lambda,p_i}, \beta^{\lambda,p_i})$. The specific algorithm is:

Algorithm 1

1 Set $\lambda = M$, and compute the connected components of $[u \geq \lambda]$ and its signature points p_i which are chosen as points of the connected component where u attains its maximum. We denote them by X^{λ,p_i}. Set $\beta^{\lambda,p_i} = M$, $C(X^{\lambda,p_i}) = 0$, $\forall i$.

2 Store the couples $(X^{\lambda,p_i}, \beta^{\lambda,p_i})$ in $L[\lambda]$.

3 If $\lambda - 1 < m$, we go to step **5**. Otherwise, set $\lambda = \lambda - 1$, and compute the connected components of $[u \geq \lambda]$. Denote them by $X^{\lambda,i}$, $i = 1, \ldots, N_\lambda$. Three cases are possible

a) If $X^{\lambda,i}$ contains no connected component of $[u \geq \lambda+1]$, then we associate to it a new signature point p_i, denote X^{λ,p_i} instead of $X^{\lambda,i}$, and define $\beta^{\lambda,p_i} = \lambda$, $C(X^{\lambda,p_i}) = 0$.

b) If $X^{\lambda,i}$ was already present at level $\lambda+1$, and contains only one component $X^{\lambda+1,p_i}$ of $[u \geq \lambda+1]$, then we associate to it the same marker p_i. We update $\beta^{\lambda,p_i} = \beta^{\lambda+1,p_i}$, $C(X^{\lambda,p_i}) = \beta^{\lambda,p_i} - \lambda$.

c) If $X^{\lambda,i}$ was already present at level $\lambda + 1$, and contains more than one component of $[u \geq \lambda+1]$, say $X^{\lambda+1,p_{i_1}}, \ldots, X^{\lambda+1,p_{i_k}}$, then we associate to it the marker $p_i := p_{i_j}$ with highest value of $\beta^{\lambda+1,p_{i_j}}$. We denote X^{λ,p_i} instead of $X^{\lambda,i}$, and update $\beta^{\lambda,p_i} = \beta^{\lambda+1,p_i}$, $C(X^{\lambda,p_i}) = \beta^{\lambda,p_i} - \lambda$. If $C(X^{\lambda,p_i}) \geq MinConstrast$, then we mark X^{λ,p_i} as critical and redefine $\beta^{\lambda,p_i} = \lambda$, $C(X^{\lambda,p_i}) = 0$. In this case, store the set of markers

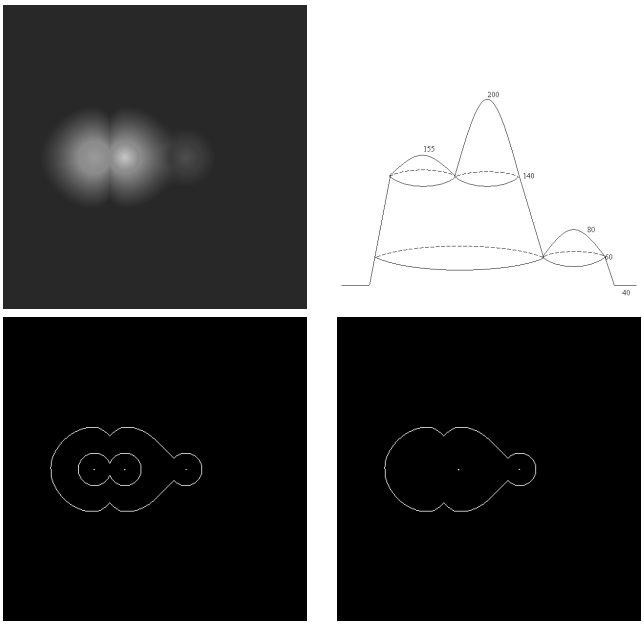

FIG. 4.1. *a) The original gray level image representing three peaks. b) A sectional view of Figure a. We display the values at the maximum points, the saddles and the minimum value. The curves delimit the maximal monotone sections c) The critical levels of the Figure represented in a). It has been computed with our algorithm taking the parameter* $MinContrast = 0$. *After crossing each saddle the parameters* β^{λ,p_i} *and* $C(X^{\lambda,p_i})$ *are reset according to Step 3.c in our Algorithm. Note that at the saddle we have taken the level curve immediately below, thus the two connected components merging there are inside this curve. d) We display the critical levels obtained with our algorithm with the parameter* $MinContrast = 80$ *(any value between* 61 *and* 140 *would produce the same result). In this case the first saddle is not a critical level since the contrast of the regions merging there is* 60*, strictly less than* 80. *After crossing level* 140 *the first two maxima merged there and have a single representative (the maximum point at level* 200*). When the merging at the second saddle happens the contrast is already sufficient and equal to* 140. *The markers of the two merging regions are depicted here. Note that the marker of the minimum region is also depicted, though not clearly visible, at the point of coordinates* $(0, 0)$.

$\{p_{i_1}, \ldots, p_{i_k}\}$ which are represented by p_i.[1]

4 Return to step **2**

5 Select from L the list of couples $(X^{\lambda,p_i}, \beta^{\lambda,p_i})$ marked as critical.

6 The output of the algorithm will be the set of curves Γ_M corresponding to the boundaries of the selected critical sections X^{λ,p_i}, and the markers represented by p_i.

It is important to remark that the parameter $MinContrast$ selects the critical levels in a *causal* way. That is, if we denote as $C(u, \alpha)$ the set critical levels of u computed with our Algorithm with the parameter $MinContrast = \alpha$, then $C(u, \alpha) \supset C(u, \beta)$ when $\alpha < \beta$. Fig. 4.1 displays an example where we have computed the critical levels of a synthetic image.

Let us finally say some words concerning the discrete versus continuous formulation. The discrete case could be subsumed under the continuous case, by considering a discrete function on a grid as a function on a continuous domain by identifying each

[1] Case c) corresponds to the case where two or more connected components of $[u \geq \lambda + 1]$ merge at level λ. This merging is only labeled as critical if the contrast parameter is larger than $MinConstrast$.

pixel with a square. Moreover, we may do this in such a way that our function become upper semi-continuous (i.e., such that the upper level sets $[u \geq \lambda]$ are closed), as it is done in [8]. Then the discrete case would be covered if we had developed our results for upper semi-continuous functions. Indeed, parts of the theory can be generalized to this context (essentially the results in [7]), but at the moment being we have not yet addressed Proposition 11.2 in this context. We believe that it is true, but it is not yet done. From the practical point of view, the same algorithms would hold. On the other hand, we would like to mention that our algorithm does not select edges in the sampling process. As a result, we would have large errors at the edge position. The inclusion of this curves requires the use of some edge detection along level lines of the image in the spirit of the work [16] (this remains to be explored in the future).

5. Related work. In computer graphics and geographic information systems, topographic maps represent a high level description of the data. Topographic maps are represented by the contour maps, i.e., the isocontours of the given scalar data. The description of the varying isocontours requires the introduction of data structures, like the *topographic change tree* or *contour tree* which can represent the nesting of contour lines on a contour map (or a continuous topographic structure) [39, 27, 52]. In all cases, the proposed description can be considered as an implementation of Morse theory. Given the scalar data u defined in a domain Ω of R^N ($u : \Omega \to R$), the contour map is defined in the literature as the family of isocontours $[u = \lambda]$, $\lambda \in R$, or in terms of the boundaries of upper (or lower) level sets $[u \geq \lambda]$ ($[u \leq \lambda]$). The first description is more adapted to the case of smooth data while the second description can be adapted to more general continuous data where there are plateaus of constant elevation or discontinuous data. The second description has been addressed in [15, 27], while the first description has been used in [6, 5, 52], where an a priori interpolation of the discrete data is required to permit the isocontour description.

The contour map is organized in a data structure, either the contour tree [27, 52], or the Reeb graph [51, 36]. The *contour tree* represents the nesting of contour lines of the contour map. The contour tree encodes the topological changes of the level curves of the data. Critical values and its associated features, peaks (maxima), pits (minima), or passes (saddles), can be extracted from the contour trees [27]. Contour trees can also be used as a tool to compute other terrain features such as ridges and ravines [27]. For practical applications, the data structure has to be implemented with a fast algorithm and with minimal storage requirements. In [52] this is accomplished with a variant of the contour tree where the criticalities (maxima, minima, saddles, computed in a local way) are computed first. In [6] several attributes have been added to the contour data which can be used to select a subsampled family of contours which are representative of the data. As examples of such attributes the authors choose the length or area of the isocontours, the ratio length of the isocontour/area of the enclosed set, or the integral of the gradient along the isocontour. The *Reeb graph*, which represents the splitting and merging of the isocontours, was proposed in [51] as a data structure for encoding topographic maps.

In the context of computer graphics, Morse theory has also been used to encode surfaces in $3D$ space [46]. In [46], the authors also use a tree structure like the Reeb graph complemented with information about the Morse indexes of the singularities and including enough (information about) intermediate contours to be able to reconstruct by interpolation the precise way in which the surface is embedded in $3D$ space.

In image processing, the topographic description was advocated as a local and contrast invariant description of images (i.e., invariant under illumination changes),

and has lead to an underlying notion of shapes of an image as the family of connected components of upper or lower level sets of the image [11, 42]. An efficient description of the family of shapes in terms of a tree was proposed in [33, 32] and further developed in [28]. The tree of shapes as proposed in [33, 32] fuses the information of both the trees of upper $[u \geq \lambda]$ and lower $[u \leq \lambda]$ level sets of the scalar image u. The key idea for this fusion is the notion of shape as a connected component of an upper $[u \geq \lambda]$ or lower $[u < \lambda]$ level set in which the holes are filled-in. This topographic structure has been further studied in [7, 8, 32], where a Morse description of this topographic structure was developed. In [28], after a bilinear interpolation of the discrete data, the image could be treated as a continuous function and a tree of bilinear level lines $[u = \lambda]$ was computed. The tree of bilinear level lines is more related to the contour tree computed with the isocontours of the interpolated image. The work in [25] can be considered as a mathematical description of the (iso) contour tree in the case of two-dimensional functions.

In [15], Morse theory has also been used as a basic model to describe the geometric structures of $2D$ and $3D$ images, and in general, of multidimensional data. Applications have been given in different domains, in particular, to visualize structures in $3D$ medical images.

Let us finally mention that a morphological approach to image compression has been proposed by several authors, for instance [34], [41], [37], [18]. In [41] the authors propose to use binary partition trees to select the level curves which have to be encoded. The trees take into account the cost in bits to encode the selected level boundaries and the approximation error (measured with an L^2 norm). In [18], the author selected the level lines taking into account its perceptual significance which was measured in terms of the number of T and X junctions contained in it. Because of our application to the encoding of DEM data, we use a description more adapted to the topographic features of the data.

6. Computing the drainage structures. Besides the Morse description, there are other structures which are of special interest due to their topographic significance in DEM data. These mainly correspond to the drainage structures (e.g., rivers and ravines). There exists many different algorithms accurately computing such structures, see [29] and references therein. We will present an approach which is related to the one in [45]. Strictly speaking, we do not compute the drainage structures but a simplified version of them which is adapted to our purposes. In a simplistic way we can think of the drainage structures as the set of points for which there exists at least one direction in which the flow of water is accumulated or repealed. We can write down this definition mathematically by considering the set of points $\vec{x}$ such that the image u has an extremum (i.e, a maximum or a minimum) at $\vec{x}$ in some direction $\vec{v}$, hence it contains, in particular, the maxima and minima of u. Intuitively, this set contains the drainage structures (ridges and valleys), and it gives also information about boundaries of plateaus for example.

In the discrete case we shall consider only 4 different directions (values of v) corresponding to 4 different profiles in the image. Concretely, we search maxima and minima of the image u in the vertical, horizontal and diagonal profiles (in [45] only two directions were used, namely the horizontal and vertical ones). Thus we are lead to compute the local extrema of a 1D function. Let us recall that the coordinate i_0 is a local maximum (resp. minimum) of w if $w(i_0 + j) \leq w(i_0)$ (resp. $w(i_0 + j) \geq w(i_0)$) when $j \in \{+1, -1\}$. The problem is that a large number of extrema can appear due to low oscillations, mainly created by noise. In order to solve this problem we choose

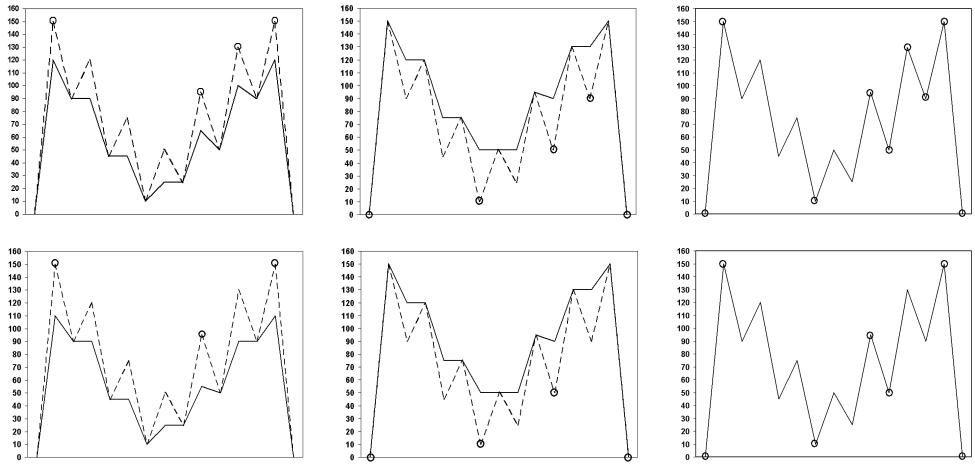

FIG. 6.1. *a) Top left: The pointed curve is an example of a 1D signal which takes values only at the points which are extrema. For the purposes of visualization, we have displayed it as a dashed curve. We display as a continuous curve the corresponding reconstruction by dilation with contrast parameter thr* $= 30$*. With a circle we display the maxima computed by Algorithm 2. b) Top middle: The corresponding reconstruction by erosion of the original signal with the superimposed minima (thr* $= 30$*) c) Top Right: The original signal with the superimposed maxima and minima (thr* $= 30$*). d) Bottom left: The same example of a 1D signal with the corresponding reconstruction by dilation with contrast parameter thr* $= 40$*. With a circle we display the maxima computed by Algorithm 2. e) Bottom middle: The corresponding reconstruction by erosion of the original signal with the superimposed minima (thr* $= 40$*) f) Bottom Right: The original signal with the superimposed maxima and minima (thr* $= 40$*).*

the most significant extrema, in our case those which have a large *contrast*. Following the standard tools in mathematical morphology, to compute the local maxima of the function w with contrast larger o equal than thr, we compute the local maxima of the reconstruction by dilation of $w - thr$ by w. Let us briefly recall this notion (see [47]). If $\delta^{(1)}(f)(i) = \max_{j\in\{-1,0,1\}} f(i+j)$ denotes the dilation of f in a neighborhood of radius 1 of i, the geodesic dilation of f by g is $\delta_g^{(1)}(f)(i) = min(\delta^{(1)}(f)(i), g(i))$. Let $\delta_g^{(n)}(f) = \delta_g^{(1)}[\delta_g^{(n-1)}(f)]$, $n \geq 2$. Then the reconstruction by dilation of a function f with respect to a function g, denoted by $R_g^\delta(f)$, is the iteration until stability of $\delta_g^{(n)}(f)$, i.e., $R_g^\delta(f) = \delta_g^{(n)}(f)$, where n is such that $\delta_g^{(n+1)}(f) = \delta_g^{(n)}(f)$ [47]. Thus, the algorithm to compute the maxima of w with contrast larger than thr is

Algorithm 2

1 *Compute* $R_w^\delta(w - thr)$.

2 *If* i_0 *is a strict local maximum of* $R_w^\delta(w - thr)$*, we store* i_0 *as a local maximum of* w *with contrast* $\geq thr$*. For any regional maxima of* $R_w^\delta(w - thr)$ *we compute the maxima of* w *there and we store them. In case that the maxima of* w *form an interval we store the end points of the interval. In case* $R_w^\delta(w - thr)$ *is a constant image, the contrast of* w *is less than* thr *and we store the values of* w *at the end-points of the*

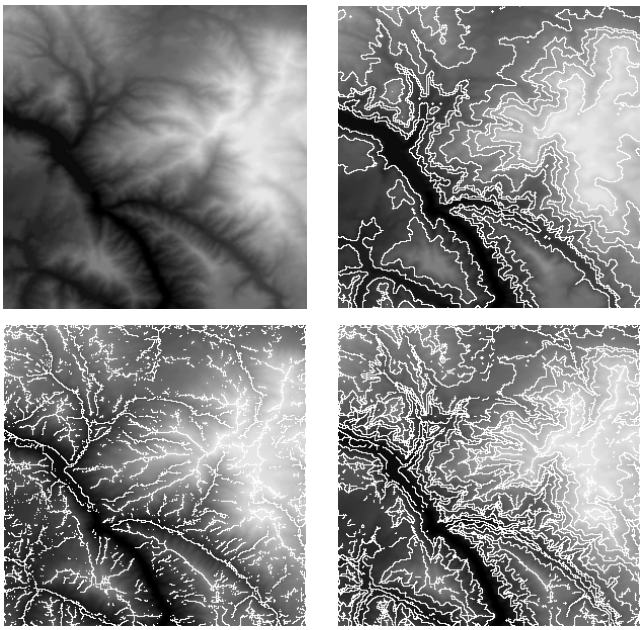

FIG. 6.2. *a) The original DEM image, b) Computed level lines (white), $MinContrast = 30$, c) Computed ridge/valley structure (white), $thr = 15$, d) The whole image sampling (white).*

interval of definition.

The minima of w with contrast larger than thr can be computed by using the reconstruction by erosion $R^{\epsilon}_{w}(w + thr)$. Since, for any two functions f, g, we have the formula $R^{\epsilon}_{g}(f) = k - R^{\delta}_{k-g}(k - f)$, k being an upper bound for f and g, we may adapt Algorithm 2 to compute the minima of w with contrast larger or equal than thr.

Figure 6.1 displays the result of applying this algorithm to a test example.

To apply Algorithm 2 to DEM data we do it along the horizontal, vertical and the two diagonal directions. The four sets of selected extrema, one in each direction, constitute our version of the simplified drainage structure, and we shall denote it by Γ_D. Figure 6.2 illustrates the whole geometric sampling process. From left to right and top to bottom we show the original DEM image (with 8 bits/sample), the level lines Γ_M corresponding to its Morse structure (computed after filtering Figure 6.2.a with the extrema filter with threshold area = 10, and using $MinContrast = 30$), the curves Γ_D corresponding to the extrema of the profiles ($thr = 15$), and the final sampling $\Gamma_{xy} = \Gamma_M \cup \Gamma_D$. We notice that the filtering applied here is used only to select the level curves corresponding to the Morse and drainage structures. A thinning step has been performed in order to obtain one pixel width curves.

7. Interpolation. Our purpose in this Section is to describe some interpolation algorithms based on the solution of a partial differential equation given the boundary data on a set of curves and/or points. Given the elevation data $u(x, y)$, we have explained in previous sections how to compute the level curves Γ_M at the critical values of u, and a simplified description of its drainage structures, Γ_D. These are stored together with the values (or an approximation to them) of u on them. Then we interpolate those values to reconstruct an approximation of the image u.

In [13] we studied and classified the interpolation algorithms which satisfy a reasonable series of axioms in terms of the solution of a partial differential equation. Of particular interest for our purposes here are two of them: the Laplacian interpolation and the Absolutely Minimizing Lipschitz Extension, denoted as AMLE in the sequel. We shall recall both of them and we shall precise in which sense they are shape preserving interpolators.

For simplicity, we shall assume that Q is a bounded connected domain in R^2. The Laplacian interpolation is based on solving the PDE

$$-\Delta u = 0 \tag{7.1}$$

with specified boundary data on ∂Q

$$u = \varphi \quad \text{on } \partial Q. \tag{7.2}$$

To guarantee the existence and uniqueness of solutions we need to assume that Q is a domain with a Lipschitz boundary, and φ is a continuous function on ∂Q. This excludes the possibility to use this model to interpolate values on irregular domains or data given on points, as it will happen in our case. In Figure 7.1.c we display the result obtained interpolating the data given on a curve and a point. We see that the vertex of the cone becomes a cusp, a well known behavior. In spite of this, since the isolated points in our data are scarce we have used (7.1) with satisfactory results.

The AMLE interpolation ([3, 4]) is based on solving the PDE

$$D^2u\,(Du, Du) = 0 \quad \text{in } Q. \tag{7.3}$$

with boundary data (7.2) (here Du and D^2u denote the gradient and the Hessian matrix of u, respectively, so that in coordinates, $D^2u\,(Du, Du) = \sum_{i,j=1}^{N} \frac{\partial^2 u}{\partial x_i \partial x_j} \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j}$). In this case we may consider more general domains and boundary data, in particular the data can be given in a finite number of Jordan curves and a finite number of points. Indeed, we may assume that Q is a bounded connected domain in R^2, and the boundary data $\varphi \in Lip_\partial(Q)$ where

$$Lip_\partial(\Omega) = \{g \in C(\partial Q) : |||g||| = \sup_{x,y \in \partial Q} \frac{|g(x) - g(y)|}{d_{\partial Q}(x, y)} < \infty\},$$

and $d_Q(x, y)$ is the geodesic distance between x and y in Q, i.e., the minimal length of all possible paths joining x and y and contained in Q [22]. If ∂Q is smooth, then $Lip_\partial(Q) = W^{1,\infty}(\partial Q)$ but $Lip_\partial(Q)$ is defined for more general domains. Let us recall that, if X is an open set or a smooth manifold in R^N, $W^{1,\infty}(X)$ denotes the space of functions $u \in L^\infty(X)$ such that $\nabla u \in L^\infty(X)$. By $W_0^{1,\infty}(X)$ we denote the closure in $W^{1,\infty}(X)$ of the smooth functions with compact support in X.

Existence and uniqueness of viscosity solutions for the AMLE model (7.3) with boundary data $\varphi \in Lip_{\partial Q}(Q)$ was proved by Jensen [22]. Moreover, he proved that the viscosity solution of (7.3) is an absolutely minimizing Lipschitz extension of φ, i.e., $u \in W^{1,\infty}(Q) \cap C(\overline{Q})$ and satisfies

$$\|Du\|_{L^\infty(Q';R^N)} \leq \|Dw\|_{L^\infty(Q';R^N)} \tag{7.4}$$

for all $Q' \subseteq Q$ and w such that $u - w \in W_0^{1,\infty}(Q')$. Let us say that the AMLE model was introduced by Aronsson in [3, 4] as the Euler-Lagrange equation of the

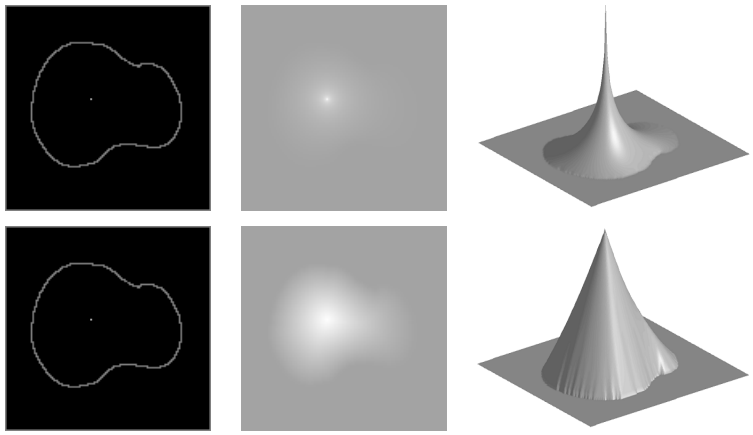

FIG. 7.1. *Top: From left to right the initial data, the interpolated data using Laplacian interpolation and its 3D representation. Bottom: From left to right the initial data, the interpolated data using AMLE interpolation and its 3D representation.*

variational problem (7.4). The above results were extended in [22] to the case of continuous boundary data $\varphi \in C(\partial Q)$, and Jensen proved that in that case, the $AMLE$ is locally Lipschitz continuous in Q [22].

In Figure 7.1, bottom, we see an example of the ability of $AMLE$ to interpolate data given on curves and points.

An important result proved by Aronsson in [4] was that smooth (C^2) $AMLE$ do not have critical points inside Q (see [38] for further results and references). In next Theorem we remark that a weaker but related result holds. We shall write it in a more general context.

Following [13], we shall consider an interpolation operator as a transformation E which associates to each open bounded set Q and each function $\varphi \in C(\partial Q)$ a function $E(\varphi, Q) \in C(\overline{Q})$ such that $E(\varphi, Q)|_{\partial Q} = \varphi$. We shall say that the interpolation operator satisfies the stability principle if

$$E(E(\varphi, Q)|_{\partial Q'}, Q') = E(\varphi, Q)|_{Q'}$$

for any open bounded set Q, any $\varphi \in \partial Q$, and any open bounded set $Q' \subseteq Q$. Suppose that the interpolation operator E satisfies the stability property, we say that E satisfies the maximum principle if

$$\inf_{\partial Q'} \varphi \leq \inf_{\overline{Q'}} E(\varphi, Q') \leq \sup_{\overline{Q'}} E(\varphi, Q') \leq \sup_{\partial Q'} \varphi \tag{7.5}$$

for any open bounded set Q, any open bounded set $Q' \subseteq Q$, and any $\varphi \in C(\partial Q')$. If E satisfies the maximum principle and $\varphi = \alpha$ in ∂Q, where $\alpha \in R$, then $E(\varphi, Q) = \alpha$ in $\overline{Q}$, and the same is true in any open bounded set $Q' \subseteq Q$.

THEOREM 7.1. *Let Q_1, Q_2 be two open simply connected sets in R^2 such that $\overline{Q_1} \subseteq Q_2$. Let $Q = Q_2 \setminus \overline{Q_1}$. Assume that $u|_{\partial Q_1} = \lambda$, $u|_{\partial Q_2} = \mu$ with $\lambda < \mu$ or $\lambda > \mu$. Let E be an interpolation operator satisfying the stability and the maximum principle. Then $E(u|_{\partial Q}, Q)$ contains only a monotone section in the sense that $\overline{Q}$ is a monotone section of $E(u|_{\partial Q}, Q)$. In particular, if $E(u|_{\partial Q}, Q)$ is the $AMLE$ extension of the boundary data inside Q, then $E(u|_{\partial Q}, Q)$ contains only a monotone section.*

Since we shall need the notion of saturation, which will be introduced in Appendix 11, the proof of Theorem 7.1 will be given in Appendix 12. We note that the Laplace interpolation requires more regularity of the boundary of the domain, but the above result can be adapted to include it. We shall not give the details here.

We can see in Figure 7.1 that, since both interpolators, Laplacian and $AMLE$, satisfy the stability and the maximum principle, we obtain no new maxima or minima. Theorem 7.1 explains in which sense they can be considered as shape preserving interpolators.

The AMLE interpolator can be applied to the interpolation of DEM data using the nonlinear sampling described in the above sections. We have shown that it is an excellent cone interpolator. The Laplace equation can also be used if we consider that isolated points are small circles, but strictly speaking its use is not theoretically justified when the data is given on scattered points. Finally, let us mention that it is possible to prove stability of the interpolations with respect to errors committed in the position of the boundary curves and the boundary data, but we shall not include the proof here (see [49]).

8. The coding step. We have described two algorithms that compute important points and curves from an image, thereby providing the basic geometric description of DEM data. These algorithms can be considered as a non uniform geometric sampling of the image. The next step is to interpolate the missing data from our sampling. There exist several algorithms to interpolate data from curves and/or points. We have in particular tested two of them: the Laplacian and the AMLE model. We remark that, rigorously speaking, only the AMLE model can be used to interpolate values specified on points [3, 13, 22] (see Fig. 7.1). In spite of this, we shall also use the Laplacian since there are many curves in the data and we may think of points as small regions. In order to evaluate these interpolation schemes we have chosen as a measure of goodness the entropy of the residual between the original image and the interpolated one. This is a natural choice since we want to minimize the number of bits used to encode the errors between the interpolated and the original images. Both models, AMLE and Laplacian, were tested using the whole image sampling (Γ_{xy} and $u|_{\Gamma_{xy}}$), and using only the level lines corresponding to critical values (Γ_M and $u|_{\Gamma_M}$). In the second case the AMLE model performed better than the Laplacian (the entropy of the residual, the maximum, and $RMSE$ errors were lower than in the Laplacian interpolation). Surprisingly, in the case of the whole sampling structure the winner was the Laplacian, although there is not much difference, and the interpolation for the AMLE model looked visually better. After these tests we have decided to use the Laplacian interpolation to obtain the first estimation of the image from the selected curves and points. In order to control the maximum (sup) error we simply store/encode the quantized error information (that is why the entropy of the residual was a natural measure of goodness for the interpolant).

At this point we need to consider how to encode both the geometrically sampled data and the quantized residual errors once a sup error e is specified. We proceed to address this now. The geometry of the sampled curves (Γ_{xy}) and their gray levels ($u|_{\Gamma_{xy}}$) are encoded separately. To encode the geometry we use a differential chain coding strategy, see [24, 30, 14]. In the future we plan to explore an encoding based on rate-distortion theory, as in [44]. For the gray levels, if we accept losses, we may use an ENO (Essentially Non Oscillatory) based encoding scheme [2] which also controls the sup error, a fundamental requirement of the application as stated before. Finally we compress both the geometry and the gray values of the curves using an arithmetic

coder. Having these curves and the data on them, we can interpolate them by means of the Laplace equation to obtain the first estimate of the image.

Finally, to control the maximum error, we need to store the residuals r. Encoding the residuals r can be simply done by quantizing them using

$$r^q = sign(r)\left\lfloor \frac{|r|}{e} \right\rfloor e, \tag{8.1}$$

and then coding the resulting r^q with an arithmetic coder.

Algorithm 3

1	*From the original image, eventually filtered to reduce its bandwidth, u compute a subsampled image $u_l(i,j) = u(li, lj)$ where l is the reducing factor.*
2	*Compute Γ_{xy} and $u_l|_{\Gamma_{xy}}$. Encode Γ_{xy}. Compute $\tilde{u_l}|_{\Gamma_{xy}}$ by applying an ENO encoding scheme, with maximum error e, to $u_l|_{\Gamma_{xy}}$.*
3	*Compute the interpolated image $\tilde{u_l}$ by solving Laplace equation with initial data $(\Gamma_{xy}, \tilde{u_l}|_{\Gamma_{xy}})$.*
4	*Compute and quantize the residual r_l^q between u_l and $\tilde{u}_l$. Let $\hat{u}_l = \tilde{u_l} + r_l^q$ be the approximation of u_l satisfying $\sup\{|u_l - \hat{u_l}|\} = e$.*
5	*Zoom out $\hat{u_l}$ and compute and quantize the new residual r^q between u and $\tilde{u}$ in order to satisfy $\sup\{|u - (\tilde{u} + r^q)|\} = e$.*
6	*Finally compress Γ_{xy}, $\tilde{u_l}|_{\Gamma_{xy}}$, r_l^q and r^q using an arithmetic coder.*

TABLE 8.1
Algorithm for compressing DEM data.

The compression ratios using this approach were already satisfactory. We observed that the encoding of the geometry represented the main cost in bits. This is due mainly to the irregularity of the curves and the inefficiency of the differential chain coding approach (3 bits/pixel with 8-connected curves or 2 bits/pixel with 4-connected curves). To further improve the encoding of the geometry we have adopted a simple multiscale approach. We compute and encode the curves and the residuals in a subsampled image and then zoom out the result and recompute new residuals. If required, before sampling, we may filter the given image with a low pass filter or with an anisotropic filter like motion by mean curvature, or affine invariant smoothing [19], [43]. In our experiments below, we only filtered the image with the extrema filter with area threshold = 20. In case of applying anisotropic filters it would be reasonable, in order to preserve some features of the topographic structure, to apply these filters while keeping some points fixed, namely, the points corresponding to the extrema values and the saddle points, or simply we could fix the multiple points of the sampled curves Γ_{xy}. Anisotropic filtering fixing points was studied in [10].

The zoom out process can be done by using a bicubic spline interpolation, although this can create new maxima and minima due to the well known oscillation

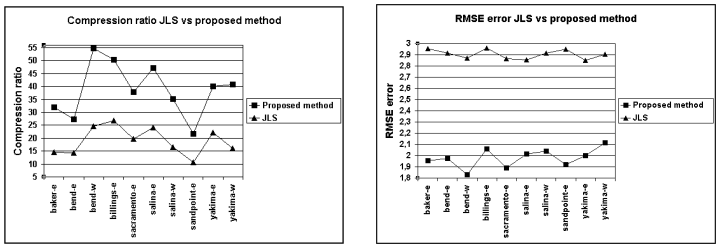

FIG. 9.1. *Left: Compression ratio of JLS vs the proposed method for a set of ten images (quantized with 8 bits/sample) compressed with a maximum error of 5. Right: Corresponding root mean square error (RMSE) for the two methods.*

problem of splines. In order to avoid this kind of errors we have used a *shape preserving spline*, which avoids the oscillation problem of classical splines, that is, respects the monotonicity of the original data (no new maxima or minima are created). Concretely, we have implemented the algorithm proposed in [21].

The complete algorithm for compressing DEM data is summarized in table 8.1.

9. Compression results. We now briefly present some experimental results. Additional examples can be found in [48]. In order to compare our results, we use the JPEG-LS standard for lossless and controlled lossy image compression, being this the only standard that permits a control on the maximal per pixel error [55]. We also used JPEG-2000 [1] in which, by reintroducing the errors, we are able to control the maximum error. In the comparisons below, JPEG-LS is denoted by JLS, while ours is denoted by ME, standing for morphological encoding. We report results on a set of 10 DEM images of size 1200 by 1200 pixels quantized with 8 or 16 bits/sample.

Figure 9.1 compares the performances of ME versus JLS for a set of 10 different DEM images when we allow a maximum error $e = 5$ (comparisons with smaller errors can be found in [48]). In the left plot we can see the compression ratio (denoted in the sequel by CR) of both methods for the set of 10 images, the right plot corresponds to the $RMSE$ for both methods. The average values for the CR and $RMSE$ of JLS and ME were computed (for these values per each individual image, see [48]). Average CR for the case of JLS and ME are **18.9872** and **38.7217** respectively. That is, our proposed scheme ME reaches twice the compression ratio of JLS. In addition, the average $RMSE$ for the cases of JLS and ME is **2.9038** and **1.9804**, respectively.

Table 9.1 compares the results obtained using ME and $JPEG-2000$ for the set of 10 images in Column 0 which are quantized with 16 bits/sample. Columns 1 and 2 contain the compression ratio and the $RMSE$ in meters, respectively, when compressed with ME (with $e = 5$ meters). The parameters used where the following: we compute the critical level curves of the data after filtering them with extrema filters with area threshold = 20 (but we checked that any area parameter between 20 and 200 gives similar results) and $MinContrast = 30\frac{\max(u)-\min(u)}{255}$ (i.e. $MinContrast = 30$ if the range of the image is $[0, 255]$); the drainage structures were computed using the threshold $thr = 10\frac{\max(u)-\min(u)}{255}$ (i.e., $thr = 10$ if the range of the image is $[0, 255]$). Most of the cost of compression is spend in encoding the quantized errors. The cost of encoding the whole set of sampled curves Γ_{xy} and the gray levels on it $u|_{\Gamma_{xy}}$ represents, in percentage, $18.17, 22.77, 20.34, 15.80, 13.46, 14.50, 16.89, 24.46, 19.16, 16.16$, respectively, for the set of images in Column 0. In Columns 3,4,5 of Table 9.1 we display the compression ratio (CR), the L^∞ error, and the $RMSE$, obtained by using

	CR (ME) (16 bits)	RMSE (meters)	CR ($JPEG-2000$) (adapted n. bits)	L^∞ error	$RMSE$
baker-e	11.2084	2.5602	11.2264	15	1.5906
bend-e	11.6345	2.5251	11.6594	12	1.6371
bend-w	15.8574	2.5728	16.3574	14	1.4220
billings-e	19.2926	2.4949	19.4837	14	1.2877
sacramento-e	9.0442	2.6483	9.1911	15	1.8514
salina-e	13.7823	2.6217	16.0282	13	1.7411
salina-w	11.5732	2.5993	11.9706	14	1.6899
sandpoint-e	9.5504	2.5189	9.5813	11	1.6459
yakima-e	17.3113	2.4832	17.4141	15	1.3741
yakima-w	7.3946	2.7296	8.3368	15	2.1757
AVERAGE	**12.6649**	**2.5754**	**13.1249**	**13,8**	**1.6415**

TABLE 9.1

Column 0 contains the name of the images. Columns 1 shows the compression ratio (CR) of the ME algorithm on the set of images of Column 0 quantized with 16 bits/sample. Column 2 shows the corresponding $RMSE$ error (given in meters). In Columns 3,4,5 we display the compression ratio, L^∞, and $RMSE$ errors obtained using $JPEG-2000$ [1]. We have used $JPEG-2000$ to compress the images in column 0; for that, we run the JasPer software but providing to it the actual range of the image (which amounts to the actual number b of bits/sample, $b = 11$, or 12, for the set of images used) so that the quantizer could be adapted to it; and asking for the same compression ratio than the one obtained in Column 1 rescaled by the factor $f = b/16$ so that we obtain a compression ratio similar to Column 1. The compression ratio is displayed in Column 3 and it is computed with respect to the original bit length. Note that the compression ratio may be slightly different to the one we asked for, due to the properties of the quantizer. Columns 4 contains the corresponding L^∞ error of the compressed image. Column 5 contains the corresponding $RMSE$ error. The last row shows the average values for the set of 10 images. We can also compute the points where the L^∞ error was greater than 5 and correct the corresponding values at those points. The corresponding compression ratio (compressing also the quantized errors with an arithmetic coder) and $RMSE$ are given in the text.

$JPEG-2000$ for the same set of images. Since the quantizer adapts to the actual number of bits/sample of the image, we run Jasper software ([1]) giving to it the actual range of the image (which amounts to the actual number b of bits/sample) and the same rate of compression obtained with ME rescaled by the factor $f = \frac{b}{16}$. In the set of images we used, the value of b is either 11, or 12. This re-scaling makes that the rate of compression, when computed using the 16 bits/sample, gives the rate of compression obtained with ME (which is computed for the images quantized with 16 bits/sample). Note that the rates of compression are only approximately equal, as obtained in the output of $JPEG-2000$. We also see that the $RMSE$ is lower than in the ME case. If we want to control the L^∞ error of the compressed image, we may quantize the errors between the original and the compressed images, compress them with an arithmetic coder, and recompute the compression ratio. After this correction the average compression ratios and $RMSE$ are 12.6152, and 1.5793, respectively. For a comparison in the case of 8 bits/sample we refer to [48].

We should also note that we have compared the visual quality both of the compressed gray scale images and their topographic maps, and observed that ME and $JPEG-2000$ are significantly better than JPEG-LS [48].

10. Conclusions. In this work we have presented some techniques to compute a basic geometric representation of images. This representation is given by the Morse

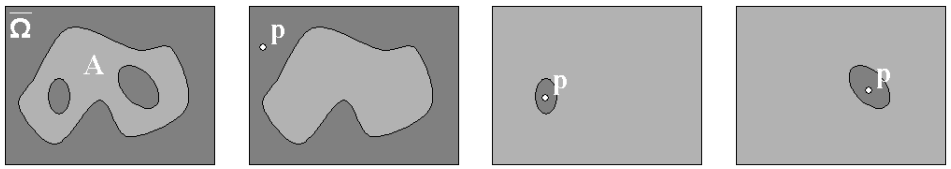

FIG. 11.1. *From left to right a set A (in light gray) and its saturations (also in light gray) with respect to different points p.*

and drainage structures of the image. We have given simple algorithms to compute these structures and have proved them to be well founded. Finally, we have proved that any interpolation process satisfying the stability and maximum principle is also shape preserving, in the sense that no new maxima or minima is created. As an application, this geometric image representation was used to derive a non uniform sampling strategy that when combined with interpolation and coding techniques, provided a novel DEM compression algorithm. This algorithm produces compression ratios far superior than JPEG-LS and similar to $JPEG-2000$, while permits to control the maximal error in the decoded image, a property which is fundamental for most DEM applications.

We are currently investigating the use of geometric sampling techniques for the compression of natural images. The question then is what is a good geometric representation of natural images that will lead to compression results as the ones obtained for DEM with the techniques here introduced.

11. Appendix: the proof of Theorem 3.9. Let us recall that we are assuming that (**H**) holds. Let us recall the definition of saturation, which consists essentially in filling the holes of a set. A detailed study of this operation and its properties can be found in [7], [8], [32].

DEFINITION 11.1. *Let $A \subseteq \overline{\Omega}$. We call holes of A in $\overline{\Omega}$ the components of $\overline{\Omega} \setminus A$. Let $p_\infty \in \overline{\Omega} \setminus A$ be a reference point, and let T be the hole of A in $\overline{\Omega}$ containing p_∞. We define the saturation of A with respect to p_∞ as the set $\overline{\Omega} \setminus T$ and we denote it by $sat(A, p_\infty)$. We shall refer to T as the external hole of A and to the other holes of A as its internal holes. By extension, if $p_\infty \in A$, by convention we define $sat(A, p_\infty) = \overline{\Omega}$. Note that $sat(A, p_\infty)$ is the union of A and its internal holes.*

The reference point p_∞ acts as a point at infinity. In all what follows, we assume that the point $p_\infty \in \overline{\Omega}$ on which the saturations are based is fixed, i.e., all saturations will be computed with respect to p_∞. To simplify our notation, we shall write $sat(A)$ instead of $sat(A, p_\infty)$, unless we change the point at infinity and we need to specify it. We shall also speak of holes of A instead of holes of A in $\overline{\Omega}$. We refer to Figure 11.1 for an example.

PROPOSITION 11.2. *If $\lambda \in R$ is a critical value, then λ is also a singular value.*

To prove Proposition 11.2, we shall need the following two Lemmas. Similar results, though in a different context, were proved in [7].

Let A, B be two closed connected sets such that $A \cap B = \emptyset$. Then, by Lemma 3.2 in [8], we have that either *a)* $sat(A) \cap sat(B) = \emptyset$, or *b)* $sat(A) \subseteq sat(B)$, or $sat(B) \subseteq sat(A)$. If *a)* holds, then we define $Q(A) = \partial sat(A)$, $Q(B) = \partial sat(B)$. If *b)* holds, then $sat(A)$ is contained in a hole H of B. Take a point $p \in H \setminus sat(A)$. Then $sat(A, p) \cap sat(B, p) = \emptyset$ and we define $Q(A) = \partial sat(A, p)$, $Q(B) = \partial sat(B, p)$. In a similar way, we define $Q(A)$ and $Q(B)$ when $sat(B) \subseteq sat(A)$.

LEMMA 11.3. *Let $u \in C(\overline{\Omega})$. Let $\lambda \in R$. Suppose that X is a connected*

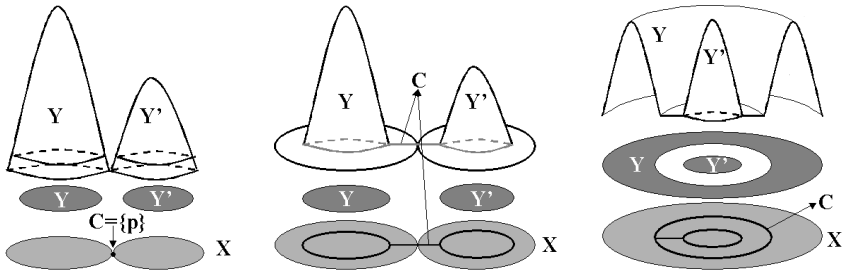

FIG. 11.2. *From left to right we present three different cases that may happen when considering the hypothesis of Lemmas 11.3 and 11.4.*

component of $[u \geq \lambda]$ *such that* $[u > \lambda] \cap X$ *is not connected. Then there exist* $Y, Y' \in \mathcal{CC}([u > \lambda] \cap X)$, $Y \neq Y'$, *and a continuum* $C \subseteq [u = \lambda] \cap X$ *such that* $\overline{Y} \cup \overline{Y'} \cup C$ *is connected. If* $\overline{Y} \cap \overline{Y'} = \emptyset$, *then we may assume that* $C \supseteq Q(\overline{Y}) \cup Q(\overline{Y'})$.

Figure 11.2 illustrates the statement of Lemma 11.3 by displaying some examples. At the top of the figure we display the considered case (different functions u), in the middle the associated sets Y and Y', and finally the set X containing the continuum C for each case.

Proof. 1: If A, B are two connected components of $[u > \lambda] \cap X$ such that $\overline{A} \cap \overline{B} \neq \emptyset$, then we may take $C = \{p\}$ for some point $p \in \overline{A} \cap \overline{B}$. Thus, we may assume that $\overline{A} \cap \overline{B} = \emptyset$, for any two connected components A, B of $[u > \lambda] \cap X$.

Let us denote by $Y_1, ..., Y_n$ the connected components of $[u > \lambda] \cap X$. Since we assume (**H**), using [7] Proposition 13, we know that there is only a finite number of connected components of $[u = \lambda] \cap X$, let them be $\Lambda_1, ..., \Lambda_m$. By last paragraph, we may assume that $\overline{Y}_1, ..., \overline{Y}_n$ are two by two disjoint. Obviously $X = \cup_{i=1}^n \overline{Y}_i \cup \cup_{j=1}^m \Lambda_j$. Suppose that each Λ_i intersects at most one of the $\overline{Y}_i$, $i = 1, ..., n$. Since $\partial Y_i \subseteq [u = \lambda]$, given $\overline{Y}_i$, $i = 1, ..., n$, there is some Λ_{k_i}, $k_i \in \{1, .., m\}$, such that $\overline{Y}_i \cap \Lambda_{k_i} \neq \emptyset$. Suppose that there is some $k \in \{1, ..., m\} \setminus \{k_1, .., k_n\}$. Since Λ_k are closed sets and X is connected, it cannot happen that Λ_k does not intersect $\cup_{i=1}^n \overline{Y}_i \cup \Lambda_{k_i}$. On the other hand, Λ_k does not intersect $\cup_{i=1}^n \overline{Y}_i \cup \Lambda_{k_i}$. We conclude that $m = n$ and $\{k_1, ..., k_n\} = \{1, ..., n\}$. Since the sets $\overline{Y}_i \cup \Lambda_{k_i}$ are closed, two by two disjoint and $X = \cup_{i=1}^n \overline{Y}_i \cup \Lambda_{k_i}$, we have a contradiction. Thus, we may assume that some Λ_k intersects at least two sets, say $\overline{Y}_j$ and $\overline{Y}_r$. Since Λ_k is a continuum and $\overline{Y}_j$ and $\overline{Y}_r$ do not intersect, there is a point $p \in \Lambda_k \setminus (\overline{Y}_j \cup \overline{Y}_r)$. We claim that

$$\partial sat(\overline{Y_j}, p) \cap \Lambda_k \neq \emptyset, \quad \partial sat(\overline{Y_r}, p) \cap \Lambda_k \neq \emptyset. \tag{11.1}$$

Assume, by contradiction, that $\partial sat(\overline{Y_j}, p) \cap \Lambda_k = \emptyset$. Since $\overline{Y_j} \cap \Lambda_k \neq \emptyset$, we have that $sat(\overline{Y_j}, p) \cap \Lambda_k \neq \emptyset$. Since Λ_k is a continuum, this implies that $\Lambda_k \subseteq int(sat(\overline{Y_j}, p))$. Since $p \in \Lambda_k$, we have that $p \in sat(\overline{Y_j}, p)$, a contradiction. This implies that $\partial sat(\overline{Y_j}, p) \cap \Lambda_k \neq \emptyset$. Similarly we prove the other assertion of (11.1). Since $\partial sat(\overline{Y_j}, p)$, and $\partial sat(\overline{Y_r}, p)$ are continuum contained in $[u = \lambda]$, we conclude that both sets are contained in Λ_k. This proves the Lemma.□

LEMMA 11.4. *Let* $u \in C(\overline{\Omega})$. *Let* $\lambda \in R$. *Suppose that* X *is a connected component of* $[u \geq \lambda]$ *such that* $[u > \lambda] \cap X$ *is not connected. Let* $Y, Y' \in \mathcal{CC}([u > \lambda] \cap X)$, $Y \neq Y'$, *and* $C \subseteq [u = \lambda] \cap X$ *be a continuum such that* $\overline{Y} \cup \overline{Y'} \cup C$ *is connected. If* $\overline{Y} \cap \overline{Y'} \neq \emptyset$, *we assume that* $C = \{x'\}$ *where* $x' \in \overline{Y} \cap \overline{Y'}$. *If* $\overline{Y} \cap \overline{Y'} = \emptyset$ *we assume that* $C \supseteq Q(\overline{Y}) \cup Q(\overline{Y'})$. *Let* $x \in C$. *Then* $\eta_+(u, x, \lambda) = \lambda$.

Proof. Let $\alpha > \lambda$. Let Y_α, Y'_α be connected components of $\overline{Y} \cap [u \geq \alpha], \overline{Y'} \cap [u \geq \alpha]$. Note that $Y_\alpha \subseteq \overline{Y}$, $Y'_\alpha \subseteq \overline{Y'}$, and $Y_\alpha = \mathcal{CC}([u \geq \alpha], Y_\alpha)$, $Y'_\alpha = \mathcal{CC}([u \geq \alpha], Y'_\alpha)$. Let $\lambda < \mu \leq \alpha$. Let $Y_\mu = \mathcal{CC}([u \geq \mu], Y_\alpha)$, $Y'_\mu = \mathcal{CC}([u \geq \mu], Y'_\alpha)$. Observe that $Y_\mu \subseteq Y$, $Y'_\mu \subseteq Y'$.

In case $x' \in \overline{Y} \cap \overline{Y'}$ and $C = \{x'\}$ observe that $x' \in \partial\overline{Y} \cap \partial\overline{Y'}$. In any case, let $x \in C$. Suppose that $\eta_+(u, x, \lambda) > \lambda$ where $\eta_+(u, ., .)$ denotes the η_+, defined in Section 3, corresponding to u. Let $\lambda < \mu < \hat{\mu} < \eta < \eta_+(u, x, \lambda)$. Let $X_{\lambda,x} = \mathcal{CC}(\{z \in \overline{\Omega} : u(z) \in [\lambda, \eta_+(u, x, \lambda)] \cap I((x, \lambda)\}, x)$ (see the end of Section 3 for the definition of $I(x, \lambda)$). Let $X_{\lambda,\eta} = \{z \in X_{\lambda,x} : \lambda \leq u(z) \leq \eta\}$ (a connected set). Note that $\{z \in X_{\lambda,x} : \mu \leq u(z) \leq \hat{\mu}\}$ is a nonempty set.

Let $p_0 \in \partial\overline{Y}$, with $p_0 = x'$ in case $\overline{Y} \cap \overline{Y'} \neq \emptyset$ and $p_0 \in Q(\overline{Y})$ in case $\overline{Y} \cap \overline{Y'} = \emptyset$. Observe that $p_0 \in C \subseteq X_{\lambda,\eta}$. Let $p_1 \in Y_\mu$, and let K be a continuum contained in $\overline{Y}$ joining p_0 and p_1.

Now, we claim that there is a point $p \in \overline{Y} \cap X_{\lambda,\eta}$ such that $\mu \leq u(p) \leq \hat{\mu}$. Let $L_0 = \{y \in K : u(y) < \mu\}$, $L_1 = \{y \in K : u(y) > \hat{\mu}\}$. Observe that both are open sets in K and $L_0 \subseteq \overline{Y}$. Since $u(p_0) = \lambda < \mu$, and $u(p_1) \geq \mu$, then $p_0 \in L_0$, $p_1 \notin L_0$. Then L_0 is a neighborhood of p_0 in K. We observe that $L_0 \subseteq [\lambda \leq u \leq \eta]$. Indeed, since $L_0 \subseteq K \subseteq \overline{Y}$, then $u(y) \geq \lambda$ for all $y \in L_0$, and, on the other hand, $u(y) < \mu < \eta$, for all $y \in L_0$. Given $k \geq 1$, there is a finite sequence of points $p_0^k, p_1^k, ..., p_{N_k}^k$ in K with $p_0^k = p_0$, $p_{N_k}^k = p_1$, and $d(p_i^k, p_{i+1}^k) < \frac{1}{k}$. Let j_k be the first index i such that $p_i^k \in L_0$ and $p_{i+1}^k \notin L_0$. Observe that $j_k \leq N_k - 1$. Since K is a compact set, we may assume that $p_{j_k}^k \to p$ as $k \to \infty$. Then, also $p_{j_k+1}^k \to p$. Since $u(p_{j_k}^k) < \mu$ and $u(p_{j_k+1}^k) \geq \mu$, we have that $u(p) = \mu$. On the other hand, $p \in \overline{L_0} \subseteq [\lambda \leq u \leq \eta]$, and, being limit of points not in L_0, then $p \neq p_0$. Let $m \geq 1$, and let $k_0 \geq m$ be such that $|p_{j_k}^k - p| < \frac{1}{m}$ for all $k \geq k_0$. Recall that $p_i^k \in L_0$ for all $i \leq j_k$ and $d(p_i^k, p_{i+1}^k) \leq \frac{1}{k} \leq \frac{1}{m}$ for all i. Let γ_k be the polygonal joining p_i^k to p_{i+1}^k for all $0 \leq i \leq j_k$. Then $\sup_{p' \in \gamma_k} d(p', \overline{L_0}) \leq \frac{1}{m}$. Since $p_0 \in \gamma_k$ for all k, $p_0 \in \liminf_k \gamma_k$. Then $\limsup \gamma_k$ is a continuum ([26], vol. II, p. 111) joining p_0 to p such that $\limsup_k \gamma_k \subseteq \overline{L_0} \subseteq \overline{Y} \cap [\lambda \leq u \leq \eta]$. Since $p_0 \in X_{\lambda,\eta}$ we conclude that $p \in \overline{Y} \cap X_{\lambda,\eta}$. In a similar way we prove that there is some $q \in \overline{Y'} \cap X_{\lambda,\eta}$ such that $\mu \leq u(q) \leq \hat{\mu}$.

Summarizing, we have shown that the sets $Y_{\mu,\hat{\mu}} := \{z \in \overline{Y} \cap X_{\lambda,\eta} : \mu \leq u(z) \leq \hat{\mu}\}$, $Y'_{\mu,\hat{\mu}} := \{z \in \overline{Y'} \cap X_{\lambda,\eta} : \mu \leq u(z) \leq \hat{\mu}\}$ are non empty. Since $X_{\lambda,x}$ is a monotone section, we have that the set $X_{\lambda,x} \cap [\mu \leq u \leq \hat{\mu}]$ is connected, contained in $[u > \lambda]$, and it contains $Y_{\mu,\hat{\mu}}$ and $Y'_{\mu,\hat{\mu}}$. Notice that, being connected, contained in $[u \geq \lambda]$, and intersecting X, we have $X_{\lambda,x} \cap [\mu \leq u \leq \hat{\mu}] \subseteq X$. This contradicts the fact that Y, Y' are two different connected components of $[u > \lambda] \cap X$. This contradiction proves that $\eta_+(u, x, \lambda) = \lambda$.□

Proof of Proposition 11.2. Let $\mu_n > \lambda$ be such that $\mu_n \downarrow \lambda$ and $sig([u \geq \mu_n], [u < \mu_n]) \neq sig([u \geq \lambda], [u < \lambda])$ for all n. By taking a subsequence, if necessary, we may assume that either $sig([u \geq \mu_n]) \neq sig([u \geq \lambda])$ for all n, or $sig([u < \mu_n]) \neq sig([u < \lambda])$ for all n. As described after Definition 3.6, again, modulo a subsequence, we may assume that one of the following situations happens (see Fig. 3.4):

(i) for each μ_n, there is a connected component of $[u \geq \lambda]$ which does not contain a connected component of $[u \geq \mu_n]$
(ii) for each μ_n there are two connected components of $[u \geq \mu_n]$ contained in the same connected component of $[u \geq \lambda]$
(iii) for each μ_n, there is a connected component of $[u < \mu_n]$ which contains no connected component of $[u < \lambda]$

(iv) for each μ_n there are two connected components of $[u < \lambda]$ which are contained in the same connected component of $[u < \mu_n]$.

Assume that we are in case (i). For each μ_n there is $X_n \in \mathcal{CC}([u \geq \lambda])$ such that X_n does not contain a connected component of $[u \geq \mu_n]$. Since the number of connected components of $[u \geq \lambda]$ is finite ([7], Lemma 5), by taking a subsequence, if necessary, we may assume that $X_n = X \in \mathcal{CC}[u \geq \lambda])$ is independent of n. It follows that the level λ contains a zonal maximum of u, thus, it is a singular value of u. Assume that we are in case (iii). For each μ_n there is $X_n \in \mathcal{CC}([u < \mu_n])$ such that X_n contains no connected component of $[u < \lambda]$. Since for each n we have $\mu_{n+1} < \mu_n$ then $[u < \mu_{n+1}] \subseteq [u < \mu_n]$. Assume that for each $n \in N$, there is an $m \geq n$ such that $X_n \cap [u < \mu_m] = \emptyset$. Then we find a sequence m_i such that $m_{i+1} < m_i$ for all i and X_{m_i} are two by two disjoint. Since each X_{m_i} has measure $\geq \delta$ this would imply an infinite measure for $\overline{\Omega}$. Hence we may assume that there is an $n_0 \in N$ such that for each $n \geq n_0$ we have that $X_{n_0} \cap [u < \mu_n] \neq \emptyset$. Repeating the above argument inside X_{n_0} we find $n_1 > n_0$ and $X_{n_1} \in \mathcal{CC}([u < \mu_{n_1}])$ with $X_{n_1} \subseteq X_{n_0}$ such that for all $n \geq n_1$ we have that $X_{n_1} \cap [u < \mu_n] \neq \emptyset$. In this way we construct a sequence X_{n_i} such that $X_{n_{i+1}} \subseteq X_{n_i}$ and each X_{n_i} does not contain a connected component of $[u < \lambda]$. This implies that there is a zonal (local) minimum at level λ contained in all X_{n_i}. We deduce that λ is a singular value of u.

Suppose that we are in case (ii), i.e., for each μ_n the set $[u \geq \mu_n] \cap X$ contains two different connected components, which are components of $[u \geq \mu_n]$. Let us inductively choose these connected components. Let Y_1, Y_1' be two different connected components of $[u \geq \mu_1]$. Suppose that we have already chosen Y_i, Y_i' for all $i \leq n$. Since $[u \geq \mu_n] \subseteq [u \geq \mu_{n+1}]$, if $\mathcal{CC}([u \geq \mu_{n+1}], Y_n) \cap \mathcal{CC}([u \geq \mu_{n+1}], Y_n') = \emptyset$, we take $Y_{n+1} = \mathcal{CC}([u \geq \mu_{n+1}], Y_n)$, $Y_{n+1}' = \mathcal{CC}([u \geq \mu_{n+1}], Y_n')$. If $\mathcal{CC}([u \geq \mu_{n+1}], Y_n) = \mathcal{CC}([u \geq \mu_{n+1}], Y_n')$, then we take $Y_{n+1} = \mathcal{CC}([u \geq \mu_{n+1}], Y_n)$, and Y_{n+1}' a different connected component of $[u \geq \mu_{n+1}]$. We call this a bifurcation. We note that this bifurcation cannot happen an infinite number of times since this would amount to an infinite area contained in $[u \geq \lambda]$ because this set would contain an infinite number of connected components, two by two disjoint, of the sets $[u \geq \mu_n]$. Thus, we may assume that the families of sets Y_n and Y_n' are increasing. Let

$$Y = \cup_n Y_n \quad Y' = \cup_n Y_n'.$$

Then Y and Y' are different connected components of $[u > \lambda] \cap X$. Indeed, if p_1, p_2 are such that $u(p_1) = \max_Y u$, $u(p_2) = \max_{Y'} u$, then

$$Y = \mathcal{CC}([u > \lambda], p_1), \quad Y' = \mathcal{CC}([u > \lambda], p_2).$$

Thus, the set $[u > \lambda] \cap X$ is not connected. By Lemma 11.3, there exist two connected components Z, Z' of $[u > \lambda] \cap X$ satisfying the properties stated in that Lemma. Now, by Lemma 11.4, there is a point $x \in X$ such that $\eta_+(u, x, \lambda) = \lambda$, i.e., λ is a singular value of u.

Finally, assume that we are in case (iv). Let X_n be the connected component of $[u < \mu_n]$ which contains two different connected components A_n, B_n of $[u < \lambda]$. Since the number of connected components of $[u < \lambda]$ is finite, also is finite the number of pairs of them. Thus, by extracting a subsequence, if necessary, we may assume that X_n contains two different connected components A, B of $[u < \lambda]$ which do not depend on n. Since $\mu_{n+1} < \mu_n$ we have that $X_{n+1} \subseteq X_n$. Let $\mu_{n+1} < \mu_n' < \mu_n$.

Then, if $X'_n = \mathcal{CC}([u \leq \mu'_n], A)$, we have that $X_{n+1} \subseteq X'_n \subseteq X_n$, and therefore $X'_n = \mathcal{CC}([u \leq \mu'_n], A \cup B)$ and

$$\cap_n X_n = \cap_n X'_n.$$

Since X'_n are compact and connected, also is $\cap_n X'_n$, and we have that

$$\cap_n X'_n = \mathcal{CC}([u \leq \lambda], A \cup B) =: X.$$

Now we observe that $[u < \lambda]$ has two components in X. By defining $v = -u$, and applying Lemmas 11.3 and 11.4, we obtain a point $x \in X$ such that $\eta_-(u, x, \lambda) = \lambda$. We conclude that λ is a singular value of u.□

One of the consequences of Proposition 11.2 is that the number of critical values of u is finite. Thus the signature of $[u \geq \lambda]$ is locally constant at each side of a critical value, i.e., if λ is a critical value, then there is $\epsilon > 0$ such that

$$sig([u \geq \mu]) = sig([u \geq \lambda]) \neq sig([u \geq \mu']) \quad \text{and } sig([u \geq \mu']) \text{ is constant}$$

for each $\mu < \lambda < \mu'$, $\mu \in (\lambda - \epsilon, \lambda)$, $\mu' \in (\lambda, \lambda + \epsilon)$.

In the proof of next Proposition, we shall need the following definition.

DEFINITION 11.5. *A sequence $A_1, \ldots, A_p$ of subsets of $\overline{\Omega}$ is called a chain if each A_i is contained in an internal hole of A_{i-1}, $i = 2, \ldots, p$.*

PROPOSITION 11.6. *Let $\lambda \in R$. If λ is a singular value of u, then λ is a critical value of u.*

Proof. Suppose that λ is a singular value which corresponds to a maximum value. Then there is a connected component X of $[u \geq \lambda]$ which does not intersect any connected component of $[u \geq \mu]$ for all $\mu > \lambda$. Let $p \in X$ be its marker. Then $p \in sig([u \geq \lambda])$ and $p \notin sig([u \geq \mu])$ for any $\mu > \lambda$. Thus λ is a critical value of u.

If λ is a minimum value, there is $q \in [u = \lambda]$ such that, if $\mu > \lambda$, then $\mathcal{CC}([u < \mu], q) \neq \emptyset$ and $\mathcal{CC}([u < \lambda], q) = \emptyset$. Then $q \in sig([u < \mu])$, $q \notin sig([u < \lambda])$. Thus λ is a critical value of u.

Now, suppose that $\eta_+(x, \lambda) = \lambda$. Then there is $X \in \mathcal{CC}([u = \lambda])$ such that $X^\epsilon := \mathcal{CC}([\lambda \leq u \leq \lambda + \epsilon], X)$ is not a monotone section for any $\epsilon > 0$. Since $\cap_{\epsilon > 0} X^\epsilon = X$, we have that, for $\epsilon > 0$ small enough, the only connected component of $[u = \lambda]$ contained in X^ϵ is X. Take such an ϵ. By Proposition 3.2 we find a sequence $\mu_n \downarrow \lambda$ such that $X^\epsilon \cap [u = \mu_n]$ are not connected. Let Y_n^1, Y_n^2 be two connected components of $X^\epsilon \cap [u = \mu_n]$. Let $Z = \mathcal{CC}([u \geq \lambda], X)$, $Z_n^1 = \mathcal{CC}([u \geq \mu_n], Y_n^1)$, $Z_n^2 = \mathcal{CC}([u \geq \mu_n], Y_n^2)$. Observe that, since $Y_n^1, Y_n^2 \subseteq X^\epsilon \subseteq Z$, we have $Z_n^1, Z_n^2 \subseteq Z$. If we prove that $Z_n^1 \neq Z_n^2$ for n large enough, we conclude that $sig([u \geq \lambda]) \neq sig([u \geq \mu_n])$, and, therefore, λ is a critical value of u. By Lemma 8 in [7], we may assume that there is a chain formed by the sets Y_n^1 and a chain formed by the sets Y_n^2 (first, take a subsequence of Y_n^1 with indexes n_i such that $Y_{n_i}^1$ is a chain, then take a subsequence n_{i_j} of n_i such that $Y_{n_{i_j}}^2$ is a chain). By extracting a subsequence we may assume that $\liminf Y_n^i \neq \emptyset$. Then $Y^i := \limsup Y_n^i$ is a continuum ([26], vol. II, p. 111). Observe that $Y^i \subseteq X_\epsilon \cap [u = \lambda] \subseteq X$, $i = 1, 2$. Let us also observe that if Y_n^1 is contained in a hole of Y_n^2 then, by Lemma 7 in [7], there is a connected component of $[u < \mu_n]$, or a connected component of $[u > \mu_n]$, in between both sets, thus, the set in between has area $\geq \delta$. By extracting subsequences, if necessary, we may assume that one of the following cases happens (see Fig. 11.3):

(i) the two chains have disjoint saturations

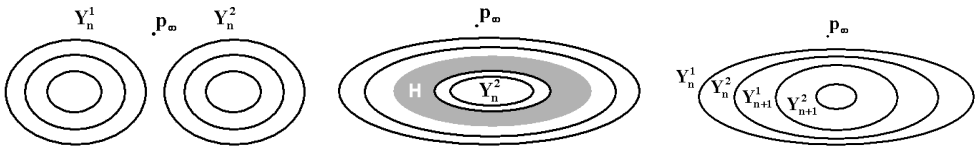

FIG. 11.3. *From left to right the three cases (i) to (iii) considered in the proof of Proposition 11.6.*

(ii) there is a hole in all sets of the first chain and the second chain is contained in it

(iii) the two chains are intertwined, i.e., all sets Y^1_n, Y^2_n form part of the same chain and no subsequence satisfies (ii).

We shall say that a chain Y^i_n is increasing (resp. decreasing) $Y^i_n \subseteq sat(Y^i_{n+1})$ for all n (resp., if $Y^i_{n+1} \subseteq sat(Y^i_n)$ for all n). Assume first that one of the chains, say Y^1_n is increasing while the other, say Y^2_n is decreasing. We know that, by extracting a subsequence, if necessary, we have either (a) $sat(Y^1_n) \subseteq sat(Y^2_n)$ for all n, or (b) $sat(Y^2_n) \subseteq sat(Y^1_n)$ for all n, or (c) $sat(Y^1_n) \cap sat(Y^2_n) = \emptyset$ for all n. If (a) or (b) hold, we are in case (ii). If (c) holds, we are in case (i).

Let us prove that one of the alternatives $(i) - (ii) - (iii)$ hold in case that both chains Y^1_n and Y^2_n are decreasing. If, for some n the saturations of Y^1_n and Y^2_n are disjoint, then we are in case (i). Thus we may assume that for each n, the saturations of Y^1_n and Y^2_n are not disjoint. Then either Y^1_n is contained in a hole of Y^2_n, or Y^2_n is contained in a hole of Y^1_n. By extracting a subsequence, if necessary, we may assume that Y^1_n is contained in a hole of Y^2_n for all n, or Y^2_n is contained in a hole of Y^1_n for all n. To fix ideas, let us assume that Y^2_n is in a hole of Y^1_n for all n. Take $n = n_1 = 1$. If all Y^1_n, $n \geq 2$, contain Y^2_1 in one of their holes then we are in case (ii). Thus we may assume there is some $n_2 > n_1$ such that $Y^1_{n_2}$ is contained in a hole of $Y^2_{n_1}$. Observe that $Y^2_{n_2}$ is contained in a hole of $Y^1_{n_2}$. If all Y^1_n with $n > n_2$ contain $Y^2_{n_2}$ in one of their holes, we are again in case (ii). Otherwise there is some $n_3 > n_3$ such that $Y^1_{n_3}$ is contained in a hole of $Y^2_{n_2}$. Proceeding in this way we shall find a subsequence of Y^1_n and of Y^2_n such that either (ii) or (iii) hold.

The case where both chains are increasing can be analyzed with the same arguments as above.

Suppose that case (i) happens. Then there are two disjoint saturated sets A, B such that $Y^1_n \subseteq A$, $Y^2_n \subseteq B$. If the chains Y^1_n and Y^2_n are decreasing or one is increasing while the other is decreasing, then we conclude that the sets Y^1 and Y^2 would be separated by one of the sets Y^i_n and then could not be connected inside X (i.e., there is no continuum $Z \subseteq X$ such that $Y^1 \cup Y^2 \cup X$ is connected), a contradiction. Thus, both chains Y^1_n, Y^2_n are increasing. In this case, since Y^1_n is contained in a hole of Y^1_{n+1}, we cannot connect Y^1_n to Y^2_n without crossing Y^1_{n+1} which is at level μ_{n+1}. Hence, Z^1_n and Z^2_n cannot be connected without crossing level μ_{n+1}. We have that $Z^1_n \neq Z^2_n$ and our conclusion follows.

Suppose that case (ii) happens. Without loss of generality we may assume that the Y^1_n are inside holes of the Y^2_n, i.e., there is a saturated set H such that the chain $Y^1_n \subseteq H$ and $H \subseteq sat(Y^2_n)$ for all n. If Y^1_n is decreasing, then Y^1 and Y^2 would be separated by one of the sets Y^1_n, and could not be connected inside X, a contradiction. If Y^1_n is increasing, then to connect Y^1_n to Y^2_n we cross Y^1_{n+1} which is at level μ_{n+1}. Then we have that $Z^1_n \neq Z^2_n$, since to connect both sets we would need to cross level

μ_{n+1}.

Suppose that (iii) happens. We may assume that Y_n^1 and Y_n^2 are monotone, either increasing or decreasing. If one of them is increasing and the other is decreasing we would be in case (ii). Thus we may assume that both families are increasing or decreasing. Suppose that both are increasing. For simplicity we shall say that two sets are ordered if one of them is contained in a hole of the other. We know that Y_1^1 and Y_1^2 are ordered. Set $n_1 = 1$. Choosing n_2 sufficiently large, we may assume that $Y_{n_2}^1, Y_{n_2}^2$ are ordered and contain $Y_{n_1}^1, Y_{n_1}^2$ in one of their holes. In this way we construct a subsequence n_j such that $Y_{n_j}^1, Y_{n_j}^2$ are ordered and contained in a hole of each of the ordered pair of sets $Y_{n_{j+1}}^1, Y_{n_{j+1}}^2$ for all j. By extracting a subsequence, if necessary, we may assume that $Y_{n_j}^1$ is contained in a hole of $Y_{n_j}^2$ and both are contained in a hole of $Y_{n_{j+1}}^1$ for all j (or the same relations with 1 and 2 interchanged). Since, as we already mentioned above, each set between $Y_{n_j}^1$ and $Y_{n_j}^2$ has area $\geq \delta$ ([7], Lemma 7), this would represent an infinite area in the intertwined chain. The same conclusion would follow in case that both chains are decreasing.

In any case, we conclude that there is a sequence of n's such that $Z_n^1 \neq Z_n^2$, hence $sig([u \geq \lambda]) \neq sig([u \geq \mu_n])$, and, therefore, λ is a critical value of u.

Finally, we assume that $\eta_-(x, \lambda) = \lambda$. Then there is $X \in \mathcal{CC}([u = \lambda])$ such that $X_\epsilon := \mathcal{CC}([\lambda - \epsilon \leq u \leq \lambda], X)$ is not a monotone section for any $\epsilon > 0$. Since $\cap_{\epsilon>0} X_\epsilon = X$, for $\epsilon > 0$ small enough we have that the only connected component of $[u = \lambda]$ contained in X_ϵ is X. Take such an ϵ. By Proposition 3.2, we find a sequence $\mu_n \uparrow \lambda$ such that $X_\epsilon \cap [u = \mu_n]$ are not connected. Let Y_n^1, Y_n^2 be two connected components of $X_\epsilon \cap [u = \mu_n]$. Observe that $Y_n^1, Y_n^2 \subseteq X_\epsilon \cap [u = \mu_n] \subseteq X$.

By changing $u \to -u$ and repeating the argument above we conclude that the sets $Z_n^1 = \mathcal{CC}([u \leq \mu_n], Y_n^1)$ and $Z_n^2 = \mathcal{CC}([u \leq \mu_n], Y_n^2)$ are different connected components of $[u \leq \mu_n]$. Since

$$\mathcal{CC}([u \leq \mu_n], Y_n^1) = \cap_{\mu>\mu_n} \mathcal{CC}([u < \mu], Y_n^1)$$

$$\mathcal{CC}([u \leq \mu_n], Y_n^2) = \cap_{\mu>\mu_n} \mathcal{CC}([u < \mu], Y_n^2)$$

if there is a sequence $\mu_n^k \downarrow \mu_n$ such that $\mathcal{CC}([u < \mu_n^k], Y_n^1) = \mathcal{CC}([u < \mu_n^k], Y_n^2)$, we would obtain that $Z_n^1 = \mathcal{CC}([u \leq \mu_n], Y_n^1) = \mathcal{CC}([u \leq \mu_n], Y_n^2) = Z_n^2$, a contradiction. Thus, for each n, there is an $\epsilon_n > 0$ such that $\mathcal{CC}([u < \mu], Y_n^1) \neq \mathcal{CC}([u < \mu], Y_n^2)$, for all $\mu \in (\mu_n, \mu_n + \epsilon_n)$. Hence we find a sequence $\mu_n' \uparrow \lambda$ such that

$$\mathcal{CC}([u < \mu_n'], Y_n^1) \neq \mathcal{CC}([u < \mu_n'], Y_n^2), \tag{11.2}$$

for all n. Let $\lambda' > \lambda$. Let $Z = \mathcal{CC}([u < \lambda'], X)$. Observe that $Y_n^1, Y_n^2 \subseteq X \subseteq Z$, hence $Z_n^1, Z_n^2 \subseteq Z$. Then (11.2) implies that $sig([u \geq \mu_n'], [u < \mu_n']) \neq sig([u \geq \lambda'], [u < \lambda'])$. Since the signature is locally constant at the left of λ, we conclude that $sig([u \geq \lambda], [u < \lambda]) = sig([u \geq \mu_n'], [u < \mu_n']) \neq sig([u \geq \lambda'], [u < \lambda'])$, and this holds for all $\lambda' > \lambda$, i.e., λ is a critical value of u.□

Remark. From the Remark after Definition 3.5 and the above results it follows that λ is a critical value of u if and only if $-\lambda$ is a critical value of $-u$.

12. Appendix: the proof of Theorem 7.1. To prove Theorem 7.1, we freely use the properties of saturations [7, 8].

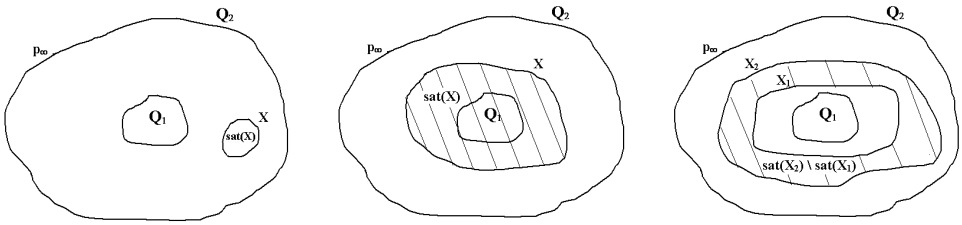

FIG. 12.1. *a) Left: the case where $sat(X) \subseteq Q$, case a) in the proof of Theorem 7.1. b) Middle: the case where $sat(X) \supseteq \overline{Q_1}$, corresponding to case b) in the text. c) Right: The curves X_1 and X_2 are two connected components of $[u = \alpha]$, and we displayed in bars the region $sat(X_2) \setminus sat(X_1)$ (see the proof of Theorem 7.1.)*

Proof of Theorem 7.1. To fix ideas, assume that $\lambda > \mu$. Let $\alpha \in (\mu, \lambda)$. Let $X \in \mathcal{CC}([u = \alpha])$. Obviously $X \cap \partial Q_i = \emptyset$, $i = 1, 2$. Let us choose a point $p_\infty \in \partial Q_2$, and let $sat(X)$ be the saturation of X in Q_2 with respect to p_∞, i.e., $sat(X)$ is X union the connected components of $\overline{Q_2} \setminus X$ with do not contain p_∞. By Lemma 3.2 in [8], either *a)* $sat(X) \subseteq Q$, or *b)* $sat(X) \supseteq \overline{Q_1}$. Assume that we are in case *a)* (Figure 12.1.a). In this case, since $u|_{\partial sat(X)} = \alpha$, we deduce that $u|_{sat(X)} = \alpha$. Then there is $\epsilon > 0$ and $Z \in \mathcal{CC}([\alpha - \epsilon \leq u \leq \alpha + \epsilon])$ such that $sat(X) \subseteq sat(Z) \subseteq Q$. Since $\partial sat(Z)$ is connected [8], then either $\partial sat(Z) \subseteq [u = \alpha - \epsilon]$, in which case $u|_{sat(Z)} = \alpha - \epsilon$, or $\partial sat(Z) \subseteq [u = \alpha + \epsilon]$, in which case $u|_{sat(Z)} = \alpha + \epsilon$. In any case, we obtain a contradiction. Thus *b)* must hold (Figure 12.1.b). Now, if there are two connected components $X_1, X_2 \in \mathcal{CC}([u = \alpha])$, then without loss of generality, we may assume that $\overline{Q_1} \subseteq sat(X_1) \subseteq sat(X_2)$ (Figure 12.1.c). Since $\partial sat(X_1), \partial sat(X_2) \subseteq [u = \alpha]$, we obtain that $u|_{sat(X_2) \setminus sat(X_1)} = \alpha$, and X_1, X_2 cannot be different. This proves that $[u = \alpha]$ consists of a single connected component. Now, we consider the case $\alpha = \mu$. If there is a connected component of $[u = \mu]$ disjoint to ∂Q_2, then arguing as above we obtain that $sat(X)$ must contain $\overline{Q_1}$. Since $u = \mu$ on $\partial Q_2 \cup \partial sat(X)$, then by interpolating we deduce that $u = \mu$ in $\overline{Q_2} \setminus sat(X)$. Hence X cannot be disjoint to ∂Q_2. This proves that $[u = \mu]$ is connected. Similarly we prove that $[u = \lambda]$ is also connected.□

REFERENCES

[1] M. D. ADAMS, *The jpeg-2000 still image compression standard*, (2001), available at http://www.ece.ubc.ca/ mdadams.

[2] F. ARÀNDIGA AND R. DONAT, *Nonlinear multiscale decompositions: The approach of A. Harten*, Numerical Algorithms, 23 (2000), pp. 175-216.

[3] G. ARONSSSON, *Extension of functions satisfying lipschitz conditions*, Ark. for Math., 6 (1967), pp. 551-561.

[4] G. ARONSSSON, *On the partial differential equation* $u_x^2 u_{xx} + 2u_x u_y u_{xy} + u_y^2 u_{yy} = 0$, Ark. for Math., 7 (1968), pp. 395-425.

[5] C. BAJAJ, V. PASCUCCI, AND D.R. SCHIKORE, *Fast isocontouring for improved interactivity*, Proc. IEEE Symp. on Volume Visualization, San Francisco, Oct. 7-8, 1996, pp. 39-46.

[6] C. BAJAJ, V. PASCUCCI, AND D.R. SCHIKORE, *The contour spectrum*, Proceedings Visualization, Phoenix, AZ, IEEE Comp. Soc. and ACM SIGGRAPH, 1997, pp. 167-173.

[7] C. BALLESTER AND V. CASELLES, *The m-components of level sets of continuous functions in wbv*, Publicacions Matemàtiques, 45 (2001), pp. 477-527.

[8] C. BALLESTER, V. CASELLES, AND P. MONASSE, *The tree of shapes of an image*, ESAIM: Control, Opt. and Calc. of Variations 9 (2003), pp. 1-18.

[9] P.J. Besl and R.C. Jain, *Segmentation through symbolic surface descriptions*, Proc. II Conf. Comp. Vision and Pattern Recognition, Miami, 1986, pp. 77-85.
[10] V. Caselles, B. Coll, and J.M. Morel, *A Kanisza programme*, Progress in Nonlinear Diferential Equations and their Applications, 25, 1996, pp. 35-55.
[11] V. Caselles, B. Coll, and J.M. Morel, *Topographic maps and local contrast changes in natural images*, Int. J. Comp. Vision, 33 (1999), pp. 5-27.
[12] V. Caselles and P. Monasse, *Grain filters*, Journal of Mathematical Imaging and Vision, 17 (2002), pp. 249-270.
[13] V. Caselles, J.M. Morel, and C. Sbert, *An axiomatic approach to image interpolation*, IEEE Transactions on Image Processing, 7 (1998), pp. 376-386.
[14] B. B. Chaudhuri and S. Chandrashekhar, *Neighboring direction runlength coding: An efficient contour coding scheme*, IEEE Transactions on Systems, Man and Cybernetics, 20 (1990), pp. 916-921.
[15] J.L. Cox, D.B. Karron, and N. Ferdous, *Topological Zone Organization of Scalar Volume Data*, Journal of Mathematical Imaging and Vision, 18 (2003), pp. 95-117.
[16] A. Desolneux, L. Moisan, and J.M. Morel, *Edge detection by helmholtz principle*, Journal of Mathematical Imaging and Vision, 14 (2001), pp. 271-284.
[17] W.R. Franklin and A. Said, *Lossy compression of elevation data*, Proc. 7th Int. Symposium on Spatial Data Handling, 1996, pp. 29-41.
[18] J. Froment, *A functional analysis model for natural images permitting structured compression*, ESAIM: COCV, 4 (1999), 473-495.
[19] F. Guichard and J.M. Morel, *Partial differential equations and image iterative filtering*, Ceremade, 9535, Université Paris-Dauphine, France, 1995.
[20] R. Haralick, L. Watson, and T. Laffey, *The topographic primal sketch*, Int. J. Rob. Research, 2 (1983), pp. 50-72.
[21] J. M. Hyman, *Accurate Monotonicity preserving cubic interpolation*, SIAM J. Sci. Stat. Comp., 4 (1983), pp. 645-654.
[22] R. Jensen, *Uniqueness of lipschitz extensions: Minimizing the sup norm of the gradient*, Arch. Rat. Mech. Anal., 123 (1993), pp. 51-74.
[23] E.G. Johnston and A. Rosenfeld, *Digital detection of pits, peaks, ridges and ravines*, IEEE Trans. Systems Man Cybernetics, 5 (1975), pp. 472-480..
[24] R. R. Estes Jr. and V. Ralph Algazi, *Efficient error free chain coding of binary documents*, Proc. of the Data Compression Conference, Snowbird, Utah, April 1995, pp. 122-132.
[25] A.S. Kronrod, *On functions of two variables*, Uspehi Math. Sciences (NS), 35 (1950), 24-134.
[26] C. Kuratowski, *Topologie I, II*, Editions J. Gabay, 1992.
[27] I.S. Kweon and T. Kanade, *Extracting topograpic terrain features from elevation maps*, CVGIP: Image Understanding, 59 (1994), 171-182.
[28] J.L. Lisani, *Comparaison automatique d'images par leurs formes*, Ph.D Thesis, Université de Paris-Dauphine, July, 2001.
[29] A. López, F. Lumbreras, J. Serrat, and J. Villanueva, *Evaluation of methods for ridge and valley detection*, IEEE Trans. Pattern Anal. and Machine Intell., 21 (1999), pp. 327-335.
[30] C. Lu and J. Dunham, *Highly efficient coding schemes for contour lines based on chain code representation*, IEEE Trans. Commun., 39 (1991), pp. 1511-1514.
[31] J. Milnor, *Morse theory*, Annals of Math. Studies 51, Princeton University Press, 1963.
[32] P. Monasse, *Représentation morphologique d'images numériques et application au recalage*, PhD thesis, Université de Paris-Dauphine, 2000.
[33] P. Monasse and F. Guichard, *Fast computation of a contrast invariant image representation*, IEEE Trans. on Image Proc., 9 (2000), pp. 860-872.
[34] F. Meyer P. Salembier, L. Torres and C. Gu, *Region-based video coding using mathematical morphology*, Proceedings of IEEE (Invited paper), 83 (1995), pp. 843-857.
[35] S.D. Rane and G. Sapiro, *Evaluation of JPEG-LS, the new lossless and controlled-lossy still image compression standard, for compression of high-resolution elevation data*, IEEE Transactions on Geoscience and Remote Sensing, 39 (2001), pp. 2298-2306.
[36] G. Reeb, *Sur les poits singuliers d'une forme de Pfaff completement integrable ou d'une fonction numérique*, Comptes Rendus Acad. Sciences Paris, 222 (1946), pp. 847-849.
[37] E. Reusens, *Joint optimization of representation model and frame segmentation for generic video compression*, Signal Processing Image Commun., 46 (1995), pp. 105-117.
[38] E. Rosset, *A lower bound for the gradient of ∞-harmonic functions*, Electronic J. of Diff. Equations, 2 (1996), pp. 1-7.
[39] J. Roubal and T.K. Poiker, *Automated contour labelling and the contour tree*, Proc. Auto-Carto, 7 (1985), pp. 472-481.
[40] A. Said and W. A. Pearlman, *An image multiresolution representation for lossless and lossy*

comression, IEEE Transaction on Image Processing, 5 (1996), pp. 1303-1310.

[41] P. Salembier and F. Marqués, *Region-based representations of image and video: Segmentation tools for multimedia services* IEEE Transactions on Circuits and Systems for Video Technology, 9 (1999), pp. 1147-1169.

[42] P. Salembier and J. Serra, *Flat zones filtering, connected operators and filters by reconstruction* IEEE Trans. Image Processing, 4 (1995), pp. 1153-1160.

[43] G. Sapiro and A. Tannenbaum, *Affine invariant scale-space*, International Journal of Computer Vision, 11 (1993), pp. 25-44.

[44] G.M. Schuster and A.K. Katsaggelos, *Rate-distortion based video compression*, Kluwer Academic Publishers, 1997.

[45] W.W. Seemuller, *The extraction of ordered vector drainage networks from elevation data*, Comp. Vision Graphics and Image Processing, 47 (1989), pp. 45-58.

[46] Y. Shinagawa, T.L. Kunii, and Y.L. Kergosien, *Surface coding based on morse theory*, IEEE Computer Graphics and Appl., 11 (1991), pp. 66-78.

[47] P. Soille, *Morphological Image Analysis*, Springer Verlag, 2003.

[48] A. Solé, V. Caselles, G. Sapiro, and F. Aràndiga, *Morse Description and Geometric Encoding of Digital Elevation Maps*, Preprint, 2003.

[49] A. Solé, *Geometric Image Coding, Filtering and Restoration*, PhD thesis, Universitat Pompeu Fabra, 2002.

[50] S.M. Srivastava, *A Course on Borel Sets*, Springer Verlag, San Diego, 1998.

[51] S. Takahashi, T. Ikeda, Y. Shinagawa, T. Kunii, and M. Ueda, *Algorithms for extracting correct critical points and constructing topological graphs from discrete geographic elevation data*, Eurographics '95, 14 (1995), pp. 181-192.

[52] M. van Kreveld, R. va Oostrum, C. Bajaj, V. Pascucci, and D. Schikore, *Contour trees and small seed sets for isosurface traversal*, 13th ACM Symp. on Comput. Geometry 1997, pp. 212-220.

[53] L. Vincent, *Gray scale area openings and closings, their efficient implementation and applications*, Proc. Workshop Mathematical Morphology and Applications to Signal Processing, Barcelona, Spain, May 1993, pp. 22-27.

[54] L. Vincent, *Morphological area openings and closings for gray-scale images*, Proc. of the Workshop "Shape in Picture",1992, Driebergen, The Netherlands, Springer-Berlin, 1994, pp. 197-208.

[55] M.J. Weinberger, G. Seroussi, and G. Sapiro, *From LOCO-I to the JPEG-LS standard*, Proceedings of the International Conference on Image Processing, 4 (1999), pp. 68-72.

[56] M.J. Weinberger, G. Seroussi, and G. Sapiro, *LOCO-I: A low complexity, context-based, lossless image compression algorithm*, Proc. IEEE Data Compression Conf., Snowbird, Utah, April, 1996, pp. 140-149.