

Morphological Image Compositing

Pierre Soille

Abstract—Image mosaicking can be defined as the registration of two or more images that are then combined into a single image. Once the images have been registered to a common coordinate system, the problem amounts to the definition of a selection rule to output a unique value for all those pixels that are present in more than one image. This process is known as image compositing. In this paper, we propose a compositing procedure based on mathematical morphology and its marker-controlled segmentation paradigm. Its scope is to position seams along salient image structures so as to diminish their visibility in the output mosaic even in the absence of radiometric corrections or blending procedures. We also show that it is suited to the seamless minimization of undesirable transient objects occurring in the regions where two or more images overlap. The proposed methodology and algorithms are illustrated for the composition of satellite images minimizing cloud cover.

Index Terms—Mathematical morphology, image compositing, seam, watershed transformation, segmentation, mosaicking, remote sensing.

1 INTRODUCTION

IMAGE mosaicking (also called *image montage* [13] or *image stitching* [11]) can be defined as the registration of two or more images that are then combined into a single image. It is of interest to numerous applications in various fields such as remote sensing, astronomy, microscopy, ophthalmology, and, more recently, digital photography and video recording. Registration deals with the geometric alignment of scenes taken from different viewpoints and/or by different sensors [43]. Once the images have been registered, they are assembled so as to create a scene having a field of view corresponding to the union of the field of views of the input scenes. The assembling of the individual registered images must address the problem of determining how the pixels belonging to regions visible from more than one scene (i.e., overlapping domains) should be represented. This problem is usually referred to as *image compositing* [30]. The composition of two or more images may generate geometric and radiometric discontinuities along the line where two separate images abut. This line is called a *seam* or *cut line*. Geometric discontinuities are usually caused by approximations of the registration method and, in the context of remote sensing, the orthorectification procedure to suppress parallax errors due to relief and tilt. Radiometric discrepancies are due to variations of one or more parameters of the image acquisition process such as sensor type, viewpoint, and acquisition time. For instance, in satellite remote sensing, even a slight variation of the latter parameter may lead to different atmospheric and illumination conditions. Furthermore, dynamic processes occurring at the level of the observed scene may also create radiometric and geometric discontinuities.

An ideal compositing algorithm should automatically output seamless mosaics both in the geometric and radiometric sense. This goal is not easy to achieve and, therefore, has motivated considerable research efforts since the very beginning of digital image processing. Historically, image compositing predates the birth of digital image processing since it was originally performed by pasting together photographic prints manually cut along carefully chosen seam lines [25], [42, chapter 12]. One of the first published algorithms for digital image compositing is credited to Milgram [27]. He defined the seam point along each line of the overlapping domain between two images as the point minimizing the sum of the gray-level differences between the left and right images, the sum being carried out over a fixed number of pixels on each side of the considered point. An enhanced approach is put forward by the same author in 1977. First, pixel-by-pixel differences are computed in the overlapping domain. The seam through the difference image is then defined as a least-cost path. This path is determined by analyzing the successive lines of the overlapping domain. Consequently, every line contains one and only one seam pixel reducing thereby the set of possible shapes of the seam line. Davis [14] suggests a similar approach where the least-cost path is computed with Dijkstra's algorithm [16]. This procedure is also adopted by Efros and Freeman [17] and referred to as the minimum-error boundary cut.

Peleg [29] proposed removing seams in a photomosaic by calculating a seam-eliminating function that imposes the gray-level differences on each side of the seam to be null. However, this assumption is too strong because it does not hold if the seam is located along an edge between two regions of different gray levels. Burt and Adelson [10] advocate the use of a multiresolution scheme (also called pyramid) whereby the images are first decomposed into a set of band-passed images by applying Gaussian filters of increasing size. Resulting component images are mosaicked using a weighted average with a transition zone that is proportional in size to the wavelengths represented in the band. Finally, these band-pass mosaics are summed to

• The author is with the Joint Research Center of the European Commission, via Enrico Fermi, T.P. 262, I-21020 Ispra, Italy. E-mail: Pierre.Soille@jrc.it.

Manuscript received 3 Mar. 2005; revised 13 July 2005; accepted 22 Aug. 2005; published online 13 Mar. 2006.

Recommended for acceptance by J. Goutsias.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-0115-0305.

generate the desired image mosaic. This multiresolution blending technique performs much better than direct blending of the original images within the regions of overlap. In this latter case, the width of the transition zone of the blending function is a critical parameter: If it is too small (compared to image features), a seam will be clearly visible, but, on the other hand, if it is too large, features from both images may appear as ghost artifacts within the transition zone, as in photographic double exposure. In situations where many images are overlapping at the same position (e.g., mosaicking from video sequences [20]), this effect may be removed by a selection filter such as the temporal median filter if an odd number of samples is available. Finally, Afek and Brand [1] suggest automating the mosaicking process in five steps:

1. Define a rough seam line in the rectangular overlapping domain as the diagonal linking the corners that do not match the corners of the input images.
2. Select tie points close to the rough seam line based on a matching algorithm.
3. Define the seam line by using a minimal path approach linking the tie points.
4. Apply geometric corrections.
5. Apply radiometric corrections.

Surprisingly, none of the previously described approaches forces the seam to adapt to the morphology of the image objects by following their boundaries. This is, however, a desirable property because object boundaries are natural transitions so that a seam following these boundaries becomes less visible in the output composed image even in the absence of radiometric corrections. Standard computer-assisted procedures require the assistance of an operator to produce seams following lines of definite tone demarcation, such as streams, roads, edges of fields, or any other image objects, while ensuring that spurious objects will not be visible in the output mosaic [19]. The need for a fully automatic algorithm fulfilling this property motivated us to search for a solution based on mathematical morphology [32], [18], [36]. A complex approach addressing the composition of a pair of gray-tone images by applying a series of morphological operators on a correlation image defined as the absolute difference of the pixels in the overlapping regions is developed in [3]. In contrast, we propose in this paper a method based on the marker-controlled segmentation of the gradient intensity and suited to the processing of an arbitrary number of input multispectral images. The method is directly applicable to three-dimensional images. In addition, we show that it can be extended for removing spurious effects or dynamic features, such as clouds and shadows, in the regions of overlap.

This paper is organized as follows: The proposed morphological compositing methodology is detailed in Section 2. Pseudocode for implementation of the algorithm is detailed in Section 3. A comparison with other methodologies is given in Section 4. Finally, conclusions and hints for further developments are presented in Section 5.

2 PROPOSED METHODOLOGY

When assembling two overlapping images, most systems require a human operator to carefully define a seam line. This line crosses the overlapping domain by following a series of points corresponding to the same set of features in each image. In this section, we assume that both geometric and radiometric discontinuities can be minimized by automatically selecting the most salient seam line occurring in the overlapping domain. We present the principle of the methodology in the case of two images and, then, extend it to any number of images while including the removal of transient data such as clouds and shadows. Notations and background definitions of morphological operators used in this paper are detailed in [36]. A shorter introduction to mathematical morphology as well as an overview of its applications to geoscience and remote sensing can be found in [38].

2.1 Principle

Let \mathcal{D}_1 and \mathcal{D}_2 denote the definition domain of two digital images f_1 and f_2 . We assume that the intersection of these definition domains is nonempty ($\mathcal{D}_1 \cap \mathcal{D}_2 \neq \emptyset$) and that they are not ordered in the sense that $\mathcal{D}_1 \not\subseteq \mathcal{D}_2$ and $\mathcal{D}_2 \not\subseteq \mathcal{D}_1$. We aim at creating an image f whose definition domain equals $\mathcal{D}_1 \cup \mathcal{D}_2$. This requires us to define a rule for determining the value of the pixels falling in the overlapping domain $\mathcal{D}_1 \cap \mathcal{D}_2$. We propose to automatically follow the most salient edges in this domain by applying the watershed transformation to the pointwise minimum (infimum) of the morphological gradient of each input image. Note that another approach proceeding in a gradient domain is put forward in [22]. It consists of minimizing a cost function defined by a dissimilarity measurement between the sum of the distance (L^p norm) between the derivatives of the initial image and the composed image. A similar gradient-domain fusion based on Poisson equations and combined with a graph-cut optimization scheme relying on a minimization procedure described in [8] is developed by Agarwala et al. [2]. The use of graph-cut techniques for image compositing was originally explored by Kwatra et al. [21] in the context of texture synthesis, where patch regions from a sample image or video are transformed and stitched together along optimal seams to generate a new output. In our approach, the pointwise minimum of the morphological gradient of each input image ensures that gray-scale variations need to occur at the same place for both images in order to be considered. Indeed, seams should not follow clouds or other dynamic features occurring in only one input image. (They will be treated separately; see Section 2.3.) The oversegmentation produced by the watershed transformation applied to raw gradient images (given the presence of many irrelevant minima) is avoided by considering the following marker sets: the internal boundary of \mathcal{D}_1 and the internal boundary of \mathcal{D}_2 belonging to the overlapping domain. Points belonging simultaneously to the internal border of \mathcal{D}_1 and \mathcal{D}_2 are excluded from the marker sets. This is illustrated in Fig. 1.

In mathematical terms, we denote by f_{marker} the marker image whose definition domain equals $\mathcal{D}_1 \cap \mathcal{D}_2$ and whose values are defined as follows:

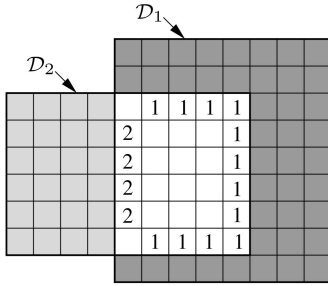


Fig. 1. Image compositing: A case of two overlapping images with rectangular definition domains. The overlapping domain is displayed in white together with the numerical values of the marker sets for Images 1 and 2, respectively.

$$[f_{\text{marker}}](\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in \rho_B^-(\mathcal{D}_2) \text{ and } \mathbf{x} \notin \rho_B^-(\mathcal{D}_1), \\ 2, & \text{if } \mathbf{x} \in \rho_B^-(\mathcal{D}_1) \text{ and } \mathbf{x} \notin \rho_B^-(\mathcal{D}_2), \\ 0, & \text{otherwise (no marker),} \end{cases} \quad (1)$$

where ρ_B^- denotes the internal gradient (or gradient by erosion) by the elementary symmetric structuring element (SE) B of the considered grid (3×3 square for the square grid). The mask image, denoted by f_{mask} , has the same definition domain as the marker image and is defined as follows:

$$f_{\text{mask}} = \rho_{nB}(f_1) \wedge \rho_{nB}(f_2), \quad (2)$$

where ρ denotes the morphological gradient, nB the isotropic SE of size n , and \wedge the pointwise minimum operator. Ideally, n equals 1 so that the elementary symmetric SE is considered. In practice, however, the value of n should be set proportionally to the average error of the registration process. By doing so, we ensure that edges produced by the same objects visible in both images overlap to some degree. We then propagate the nonnull labels of the marker image into the mask image using the marker-controlled watershed segmentation [26]. The resulting image corresponds to labeled catchment basins CB of the mask image where the labeled markers initiate the immersion simulation. We denote this image by $\text{CB}_{f_{\text{marker}}}(f_{\text{mask}})$. It provides us with a decision rule to select the input for each pixel of the overlapping domain:

$$f(\mathbf{x}) = f_i(\mathbf{x}), \quad (3)$$

where the subscript i is defined by the value of the labeled catchment basin image at position \mathbf{x} :

$$i = [\text{CB}_{f_{\text{marker}}}(f_{\text{mask}})](\mathbf{x}).$$

An example is depicted in Fig. 2. The input images (Figs. 2a and 2b) are subsets of Landsat images (band 5) of Ireland. Fig. 2c shows the pointwise minimum of the morphological gradients of the input images using a 3×3 SE. Note that the highest gradient values are located at the boundaries of the most salient image objects, here, the lake and river boundaries. The top horizontal line of the overlapping domain is used as a marker for the second input image, while the bottom diagonal line of this domain is used as a marker for the first input image. These two markers are used for the marker-controlled watershed segmentation of the gradient image. It produces the two catchment basins displayed in Fig. 2d. The

dark gray segment corresponds to the index value 2, indicating that the second image must be considered in this segment. Similarly, the light gray segment corresponds to the index value 1, indicating that the first image must be considered in this segment. This leads to the composed image displayed in Fig. 2e. (See also Fig. 2f with the seam line superimposed.) Even though the input images do not show identical contrasts, it is hard to locate the seam in the composed image displayed in Fig. 2e. Indeed, the watershed lines automatically follow the most salient edges occurring in the gradient image: the streams and lake boundaries, in this example. Note that histogram-matching techniques could have been applied to the input image to further reduce radiometric discrepancies. However, in this paper, the emphasis is on the automatic positioning of the seam line rather than the correction for radiometric variations. (This latter topic is briefly addressed in Section 5.)

Before computing the gradient, it may also be necessary to filter the input image so as to mitigate the local variability due to noise and suppress small objects so as to preserve only the most salient structures. This filtering stage can be achieved by the iterative sequential flat-zone filtering introduced in [35], [37] and extended to quasi-flat zones and multichannel images in [9].

The proposed morphological compositing methodology is directly suitable for the processing of multispectral images \mathbf{f} by substituting the morphological gradient of (2) with a multispectral scalar gradient such as that defined by the largest Euclidean distance (L^2 norm) separating a point from those falling in its neighborhood defined by an SE, the distance being measured in the feature space defined by the channel values of the input image:

$$[\rho_{nB}(\mathbf{f})](\mathbf{x}) = \max_{\mathbf{b} \in nB_{\mathbf{x}}} L^2[\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{b})],$$

where $nB_{\mathbf{x}}$ denotes the SE nB centered at position \mathbf{x} . See also [36, p. 304.]. Other multispectral gradients are detailed in [15].

2.2 General Framework

We aim at generalizing the method described for a pair of overlapping images to an arbitrary number n of overlapping images. An adaptation of the method is required because, in situations where three or more images overlap, nothing prevents a marker related to a given image from expanding beyond the definition domain of this image. We propose solving this problem by means of an iterative procedure. For conciseness, the procedure is described by generating the composed image using simultaneously all input images. By doing so, we avoid the previously described explicit definition of markers within the overlapping domains. Indeed, markers will be automatically defined along the external boundaries of the overlapping domains. This simplifies the algorithm description in the general case as well as its implementation.

We assume that none of the n input images has a definition domain included in the union of the definition domains of the other images: For all $i \in \{1, \dots, n\}$, $\mathcal{D}_i \not\subseteq \bigcup_{j \neq i} \mathcal{D}_j$. The definition domain of the desired composed image f is defined as the union of the input definition domains: $\mathcal{D}_f = \cup_i \mathcal{D}_i$. We also define an additional

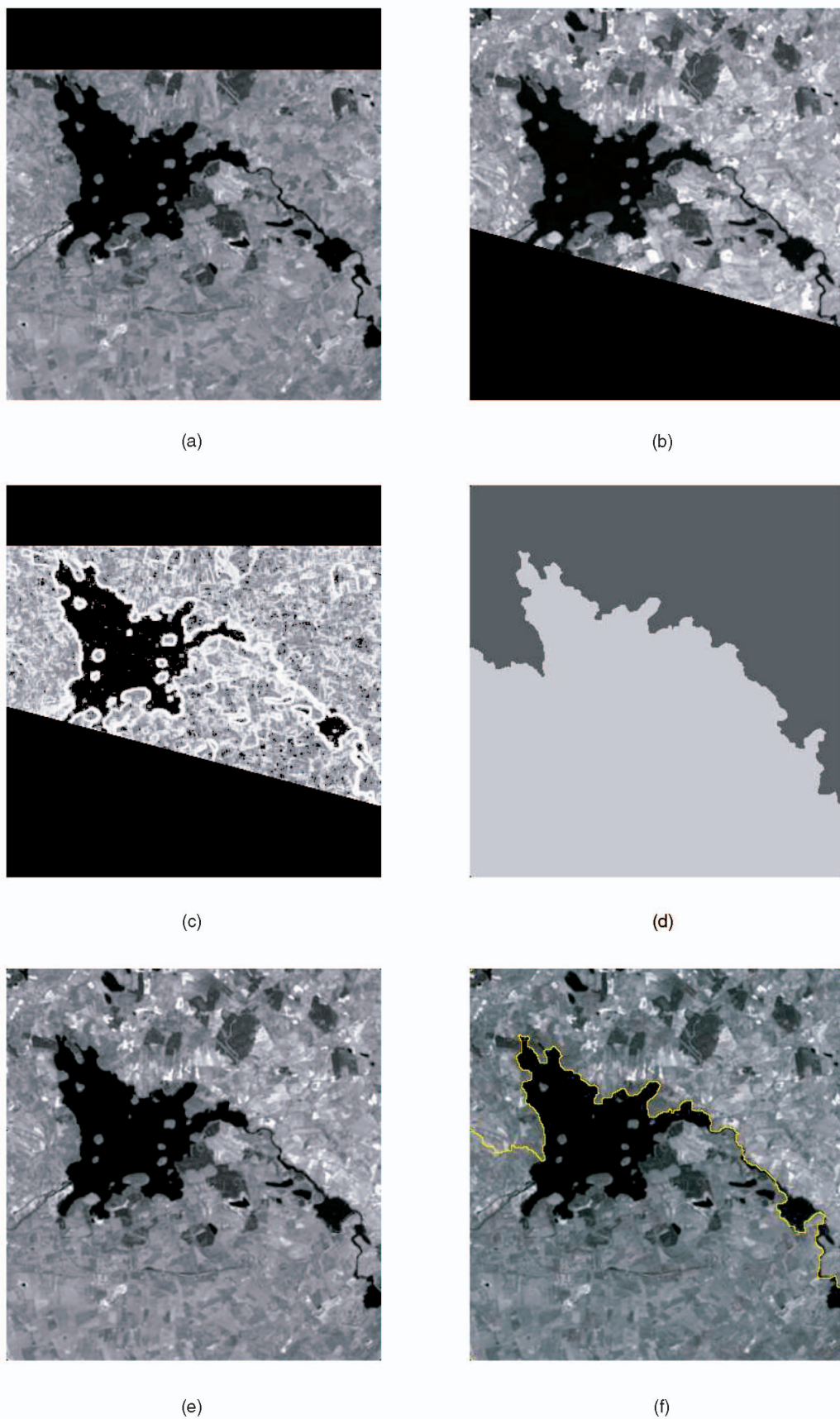


Fig. 2. Morphological compositing of two overlapping images: a real example. (a) First input image f_1 . (b) Second input image f_2 (c) Pointwise minimum of gradients in overlapping domain. (d) Marker-controlled segmentation of (c). (e) Composed image. (f) Composed image with superimposed seam line.

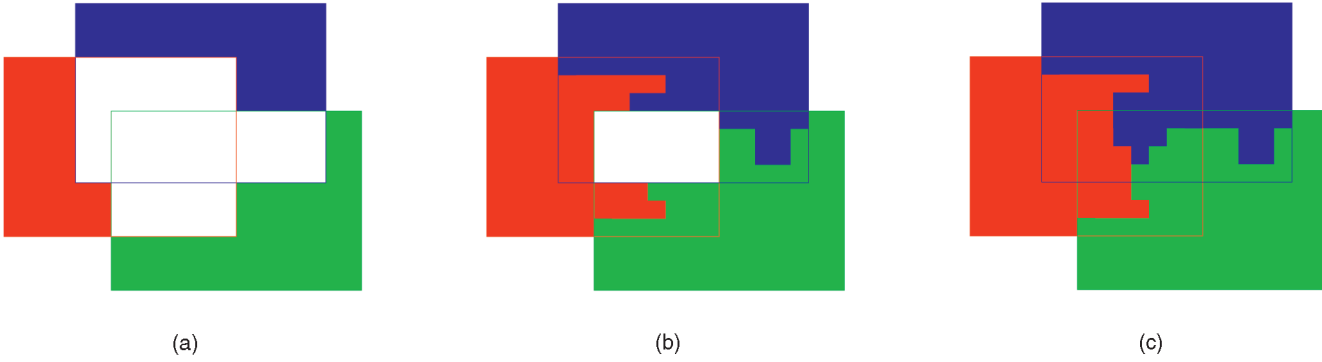


Fig. 3. Morphological image compositing in the case of three overlapping images. (a) Input images, with plain colors for nonoverlapping domains. (b) Watershed propagation of the nonoverlapping definition domains within the domains where two and only two images overlap. (c) Watershed propagation of the domains obtained in (b) and leading to a decision rule for all pixels located in the domain where all three input images overlap.

image g indicating, for each pixel of \mathcal{D}_f , how many input values are available:

$$g(\mathbf{x}) = \text{card}\{i \mid \mathbf{x} \in \mathcal{D}_i\}, \quad (4)$$

where card refers to the cardinal number (number of elements) of its argument set. The output image f is iteratively constructed, the successive iterations being denoted by the index k . Initially, k equals 1 and the values of f are uniquely defined for all those pixels \mathbf{x} of g such that $g(\mathbf{x}) = 1$. At the second iteration, our algorithm defines the values of all those pixels where two input values are available, i.e., for those pixels \mathbf{x} of g such that $g(\mathbf{x}) = 2$. We then proceed for all subsequent levels until the maximum number of overlap is reached, this value being given by the maximum value of g .

We now present a formal description of the proposed iterative procedure. We denote by $f^{(k)}$ the values of f which are defined at the end of iteration k . It follows that the definition domain of $f^{(k)}$ is defined by the set of pixels \mathbf{x} of g such that $g(\mathbf{x}) \leq k$:

$$\mathcal{D}_{f^{(k)}} = \{\mathbf{x} \mid g(\mathbf{x}) \leq k\}.$$

In addition, we use the notation $\mathcal{D}_i^{(k)}$ to refer to those pixels of $\mathcal{D}_{f^{(k)}}$ whose values originate from the input image f_i :

$$\mathcal{D}_i^{(k)} = \{\mathbf{x} \in \mathcal{D}_{f^{(k)}} \mid f^{(k)}(\mathbf{x}) \leftarrow f_i(\mathbf{x})\}.$$

Initially, k equals 1 and, therefore, the values of $f^{(1)}$ are defined for all pixels where only one image is available (no overlap). For each subsequent step, the selection rule allowing us to determine which input image should be selected relies on the marker-controlled segmentation paradigm. The marker image is defined as the set of pixels whose output values are already known at the previous step:

$$f_{\text{marker}}^{(k)}(\mathbf{x}) = \begin{cases} i, & \text{if } \mathbf{x} \in \mathcal{D}_i^{(k-1)}, \\ 0, & \text{otherwise (i.e., no marker)}. \end{cases}$$

The mask image (also called segmentation function) is defined by the pointwise minimum between the morphological gradient of all input images:

$$f_{\text{mask}} = \bigwedge_i \rho_{\lambda B}(f_i). \quad (5)$$

Note that this mask image does not depend on the index k . Hence, a fixed mask image is considered for all iteration steps. At a given iteration step, the marker-controlled watershed transformation of the mask image propagates the nonnull labels of the marker image and, therefore, provides us with the following decision rule since there is a one-to-one correspondence between the label values and the indices of the input images:

$$f^{(k)}(\mathbf{x}) = f_i(\mathbf{x}),$$

where the subscript i is defined by the value of \mathbf{x} in the image of labeled catchment basins CB:

$$i = \left[\text{CB}_{f_{\text{marker}}^{(k)}}(f_{\text{mask}}) \right](\mathbf{x}).$$

The propagation of the labels by the watershed transformation is restricted to those domains whose number of overlap equals k , i.e., pixels \mathbf{x} such that $g(\mathbf{x}) = k$. In addition, the propagation of a given label is confined to the definition domain of the input image corresponding to this label. The index k is then incremented by 1 and the whole procedure is repeated until a decision is reached for all pixels of the composed image. Fig. 3 shows a simple example in the case of three overlapping images. In this example, the colored regions represent the successive marker images. At the second step of the iterative procedure (i.e., when k equals 2), the markers are simply defined by those regions where only one image is available (see Fig. 3a). Then, these initial markers are propagated by the marker-controlled segmentation within the regions where two and only two images overlap (see Fig. 3b). The propagated markers are then used as markers for the third (final) step, where they are further propagated in the central region where all three images overlap (see Fig. 3c). The colored regions of this figure indicate, for each pixel, which image should be considered for generating the composed image.

2.3 Removal of Specific Structures

When composing a series of images, some objects occurring in the overlapping domains may be visible in one image but not in another. A rule is, therefore, needed to decide whether to include them or not in the composed image. For

example, when compositing satellite images, one should ideally produce a mosaic minimizing clouds and their shadows. The proposed morphological compositing method is generalized as follows to reach this goal: First, the image objects that should be suppressed must be extracted. In the region of overlap, pixelwise differences can help the detection of transient data. However, in remote sensing, pixel values vary a lot for scenes acquired at different times (see Section 1). Hence, we developed a specific methodology [4] to automatically extract clouds and their shadows from Landsat images using a combination of double threshold (also called the hysteresis threshold) and directional morphological operations [39], the direction of the SE being defined by the azimuthal angle of the sun. In any case, we presuppose in the following that a mask of structures to remove has been produced (either automatically or manually through user interaction) and denote by M_i the binary mask of objects present in the image f_i that should not appear in the composed image. Let us first assume here that only two input images are available. The marker image is then defined as follows in the overlap region (compare to (1)):

$$[f_{\text{marker}}](\mathbf{x}) = \begin{cases} 1, & \text{if } [\mathbf{x} \in \rho_B^-(\mathcal{D}_1) \text{ and } \mathbf{x} \notin \rho_B^-(\mathcal{D}_2)] \\ & \text{or } [\mathbf{x} \in M_2 \text{ and } \mathbf{x} \notin M_1], \\ 2, & \text{if } [\mathbf{x} \in \rho_B^-(\mathcal{D}_2) \text{ and } \mathbf{x} \notin \rho_B^-(\mathcal{D}_1)] \\ & \text{or } [\mathbf{x} \in M_1 \text{ and } \mathbf{x} \notin M_2], \\ 0, & \text{otherwise (i.e., no marker).} \end{cases}$$

That is, in the case of satellite images, clouds and their shadows present in an image are suppressed by considering markers of the other image at their position, provided that clouds and shadows do not occur at the same position. This is illustrated in Fig. 4 for the composition of two Landsat scenes of Ireland. By definition, only cloud-shadow complexes occurring in the overlap domain and that are present in only one scene are removed.

When three or more images overlap (i.e., general framework), the procedure becomes more complex. Indeed, assuming that a given pixel belongs to a feature detected in a single image and that should not appear in the composed image, there is no direct selection rule to choose among the remaining pixels' values. We propose to solve this problem by remapping the original image indices $i \in \{1, \dots, n\}$ to 2^{i-1} so as to produce unique values when summing two or more indices (the sum is then equivalent to a bitwise OR operation). If an undesired feature such as a cloud-shadow complex occurs in a region of overlap, we create a marker whose spatial extent corresponds to this complex and whose value equals the bitwise OR operation between the remapped indices of all images overlapping this complex, except the one containing this cloud-shadow complex. We call the resulting index a *composite* index since it represents the union of two or more image indices. For instance, in the example depicted in Fig. 3, in case a cloud-shadow complex of the second image would occur in the region where all three images overlap, a marker with a composite index value equal to 5 would be generated where this complex occurs (the sum of remapped index values

corresponding to the first and third images: $2^0 + 2^2 = 5$). In the sequel, the set of indices used for generating of a composite index is called its generator set. Indices referring to a unique image are called *plain* indices. By definition, plain indices are in the form 2^i . A formal definition of the marker image follows:

$$f_{\text{marker}}^{(k)}(\mathbf{x}) = \begin{cases} 2^i, & \text{if } \mathbf{x} \in \mathcal{D}_i^{(k-1)}, \\ \sum_i \{2^i \mid f_i(\mathbf{x}) \text{ is defined} \\ \text{and } \mathbf{x} \notin M_i\}, & \text{if } g(\mathbf{x}) = k \text{ and } \exists j \mid \mathbf{x} \in M_j, \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

This equation is suited to arbitrary combinations of objects to remove in the regions of overlap. Once the watershed-based propagation of the indices has been performed, we still need to give an actual image index value to each propagated composite index. By construction, this index must be chosen among its generator set. The selection must be performed so as to minimize the length of the seams. This can be achieved by using a majority rule. That is, the output index of a region with a composite index is defined as the most frequent noncomposite index value occurring among the external neighbors of the considered region and matching an index of the generator set of the considered composite index.

The main steps of the proposed morphological compositing procedure are shown in Fig. 5 for the composition of nine Landsat scenes covering Ireland. The bottom row shows the whole processed area while the top row displays a zoom within the rectangle area highlighted by a red frame. The first column corresponds to the marker image with a specific color (label) for each scene and composite labels for clouds occurring in areas where more than two images overlap (e.g., yellow color). The zoom is within a region where only two images overlap (the blue and the pink colors) so that the clouds are noncomposite markers. Their color (pink) indicates that they originate from the blue (upper right) image, so that markers of the pink image should be defined where they appear. The second column shows the mask image defined as a scalar gradient within a 3×3 neighborhood. The third column displays the output of the propagation of the markers. Once the composite labels (such as the yellow color) still present in this image are resolved by mapping them to actual image labels, the composition of the nine input images can be achieved as shown in the fourth and last column.

Note that if transient data, such as clouds and shadows are present at the same position on all available images, it will still be visible in the output composition. In applications fields such as remote sensing or biology where image analysis is performed on the output composition, it is usually not desirable to fill in these occluded regions. (It is sufficient to know where they occur to avoid biases in the subsequent image measurements.) However, in applications where the goal is to create a pleasing image for the observer in the sense that transient objects are suppressed, interpolation techniques such as those developed for removing occlusions or for image restoration might be

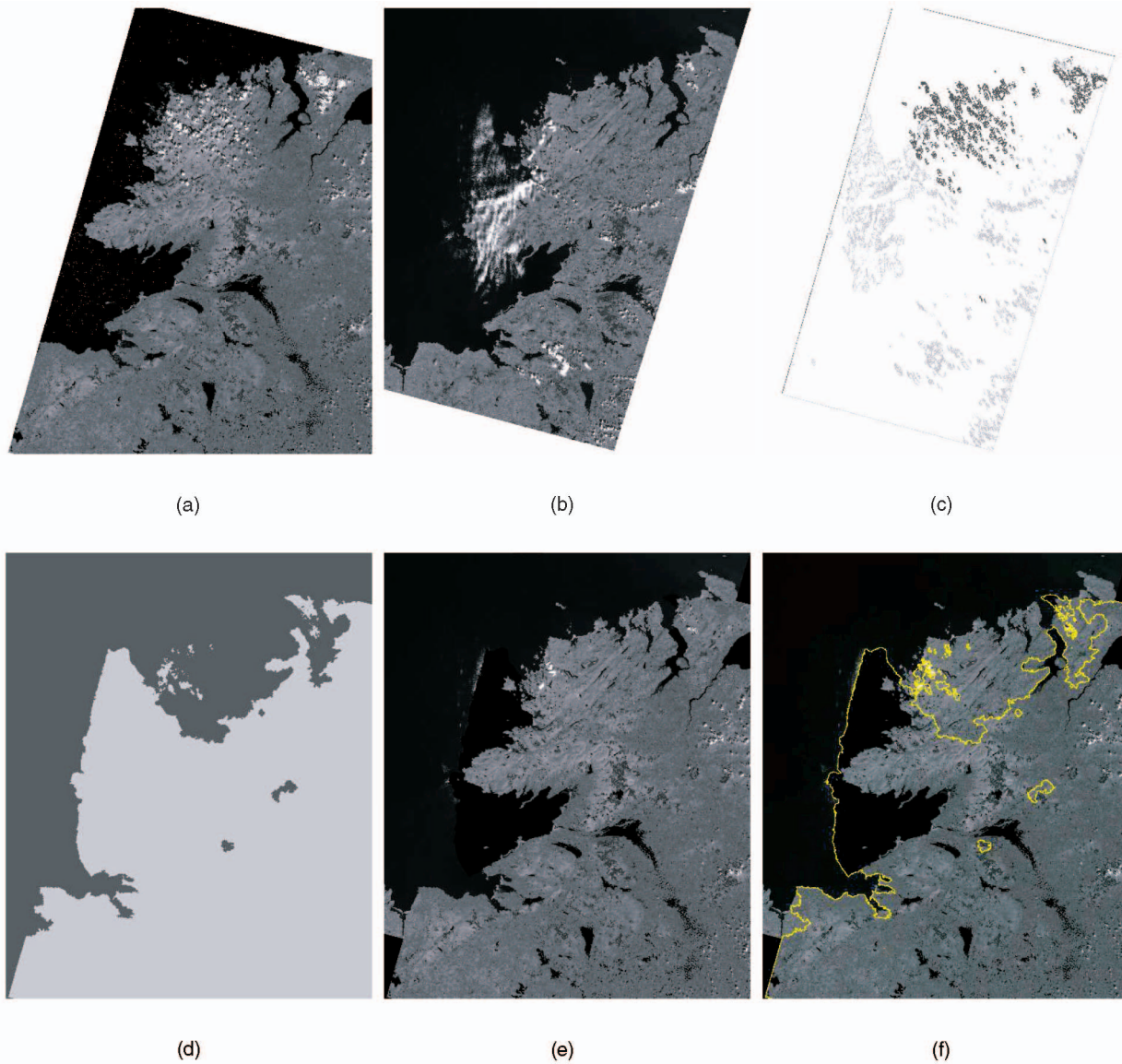


Fig. 4. Morphological image compositing including minimization of the presence of undesired objects (here, clouds and their shadows). The light gray markers visible in (c) are defined by the cloud-shadow complexes of the second scene plus the bottom and right borders of the region of overlap, while the dark gray markers are defined by the cloud-shadow complexes of the first scene plus the top and left borders of the region of overlap. (a) First scene. (b) Second scene. (c) Markers. (d) Catchment basins. (e) Resulting composition. (f) Seam lines superimposed on composition.

considered. See, for example, [24], [23], [6], [5] for methods based on a variational approach or [12] for a method combining texture synthesis and inpainting techniques.

3 IMPLEMENTATION

The following pseudocode summarizes the implementation of the proposed morphological compositing algorithm of n input images, denoted by f_i , and assuming that spurious objects have been detected for each image and stored in auxiliary binary images denoted by M_i , where $M_i(\mathbf{x}) = 1$ if \mathbf{x} belongs to a spurious object; otherwise, it equals 0:

```

1: Create images marker (6), mask (5),  $g$  (4), and  $f$ .
2: for all pixels  $\mathbf{x}$  of  $\mathcal{D}_f$  do // initialize marker, mask, and  $g$ .
3:   marker( $\mathbf{x}$ )  $\leftarrow$  0
4:   mask( $\mathbf{x}$ )  $\leftarrow$  0xFF // maximum value of pixel data
                               type
5:    $g(\mathbf{x}) \leftarrow 0$ 
6:   flag  $\leftarrow$  false
7:   for all indices  $i$  of the input images do
8:     if  $\mathbf{x} \in \mathcal{D}_i$  then
9:        $g(\mathbf{x}) \leftarrow g(\mathbf{x}) + 1$ 
10:      if  $M_i(\mathbf{x}) = 1$  then
11:        flag  $\leftarrow$  true
12:      else
13:        marker( $\mathbf{x}$ )  $\leftarrow$  marker( $\mathbf{x}$ ) +  $2^i$ 
14:        mask( $\mathbf{x}$ )  $\leftarrow$  min(mask( $\mathbf{x}$ ),  $[\rho_B(f_i)](\mathbf{x})$ )
15:      end if
16:    end if
17:  end for
18:  if flag = false and  $g(\mathbf{x}) > 1$  then
19:    marker( $\mathbf{x}$ )  $\leftarrow$  0
20:  end if
21: end for

```

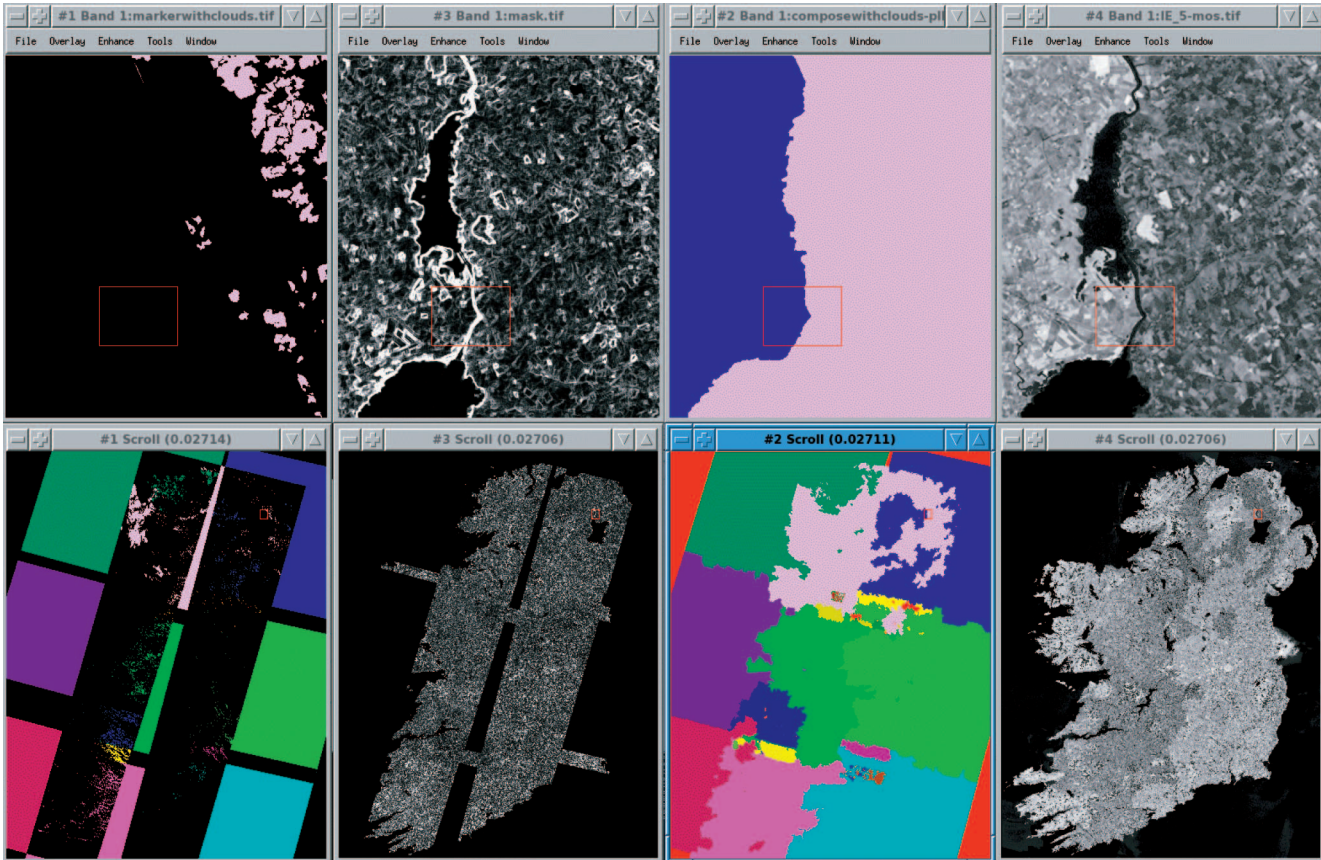


Fig. 5. Composition of nine Landsat scenes covering Ireland (bottom row: overview; top row: zoom within the red-framed rectangle). From left to right: marker image, mask image, propagated labels (including composite labels), and final composition of band 5. Note the positioning of the seam line and the presence of clouds in the marker image, allowing for their removal in the composed image. The size of the composed image is $14,001 \times 17,851$ pixels at a spatial resolution of 25 meters (i.e., 250 MBytes per composed band).

```

22:  $k \leftarrow 2$ 
23: while  $k \leq \max_{\mathbf{x} \in \mathcal{D}_f} g(\mathbf{x})$  do // iterative propagation of
      the markers
24:   marker  $\leftarrow \text{CB}_{\text{marker}}(\text{mask})$  // flood only pixels such
      that  $g(\mathbf{x}) = k$ 
25:    $k \leftarrow k + 1$ 
26: end while
27: Resolve composite labels of propagated markers using
      majority rule.
28: for all pixels  $\mathbf{x}$  of  $\mathcal{D}_f$  do // Actual composition
29:    $i \leftarrow \text{marker}(\mathbf{x})$ 
30:    $f(\mathbf{x}) \leftarrow f_i(\mathbf{x})$ 
31: end for

```

Note that the definition domain of the images marker, mask, g , and f (see line 1) is defined by the bounding box of the union of the definition domains of the input images. The complexity of the algorithm is linear in terms of the number of pixels thanks to the efficient watershed algorithm based on queue data structures [40] and adapted for an ordered propagation for increasing overlap values. In terms of processing time, about $7x/f$ seconds are required for generating the mosaic of all seven multispectral bands of a rectangular area of x MBytes (1 byte per pixel) on a CPU running at a frequency of f GHz. This time includes all steps, starting from the original bands, and includes the computation of the region with data points in each scene,

the automatic extraction of clouds and shadows, and the writing on disk of several intermediate results using the lossless compression technique described in [41]. For example, the generation of the mosaic of all seven multi-spectral bands of Ireland (250 MBytes per band, see Fig. 5) on a CPU running at a frequency of 2GHz takes less than 15 minutes. However, when dealing with larger input data, the mask, marker, and g images cannot be held together in random access memory. For instance, the composition of a country such as Austria, whose enclosing rectangle area requires 550 MBytes per 8-bit channel, pushes a 32-bit computer to its limit in terms of random access memory availability. This calls for an equivalent scheme allowing for the processing of the zones of overlap occurring in one image at a time. An efficient algorithm allowing for such a scheme will be detailed in a subsequent paper [7].

4 COMPARISON WITH OTHER METHODOLOGIES

In this section, we compare our approach with approaches referred to in the previous sections. The criteria used for comparison are listed and explained hereafter:

- **edge following:** indicates whether the seam follows the boundaries of image objects visible in all images covering the region of overlap;

TABLE 1
Merit Table for Image Compositing Algorithms Referred to in This Paper

	path cost [28, 14, 17]	pyramid [10]	tie points [1]	correlation watersheds [3]	dissimilarity of derivatives [22]	graph cuts [21] [2]	proposed
edge following	no	no	no	no	yes	no yes	yes
multisource	no	no	no	no	no	no yes ^a	yes
transient	no	no	no	no	no	no yes	yes
crisp	yes	no	yes	yes	yes	yes yes	yes
automatic	yes	yes	no	yes	yes	yes yes	yes
interactive	no	no	yes	no	no	no yes	yes ^b
3-D	no	yes	no	yes	no	yes yes	yes
multichannel	no	no	no	no	no ^c	yes yes	yes

^aAll source images must cover the same domain. ^bMarkers can be defined manually if necessary. ^cColor images are handled by applying the algorithm to each channel separately.

- **multisource:** ability to *simultaneously* assemble more than two images covering the same domain of overlap;
- **transient:** yes if the method allows for the removal of transient data;
- **crisp:** yes if the method leads to the selection of one and only one value among the available inputs;
- **automatic:** yes if the method can be used without user interaction;
- **interactive:** yes if the method allows for user interaction;
- **3D:** yes if the method is applicable to 3D images;
- **multichannel:** yes if the method treats all channels *simultaneously*.

The algorithm complexity (linear for the proposed method) has not been incorporated in the list because most discussed methods do not provide us with this information. Yet, according to [21], finding the best cut for a graph with n nodes appears to be $O(n \log(n))$, while the worst case is known to be $O(n^2)$, [31].

Table 1 summarizes the results of our comparison. The method proposed by Milgram in 1975 [27] has not been considered since it has been superseded by the method he proposed in 1977 [28]. The method by Peleg [29] concerns the elimination of existing seams rather than the actual composition of overlapping images, so it has not been included in the comparison. Finally, to reduce the size of the table, methods based on the computation of paths of minimal cost [28], [14], [17] have been grouped as well as those based on graph cuts [21], [2]. The table is binary in the sense that its scope is to indicate whether a given method satisfies a given property or not. For the edge following criterion, it shows that, in contrast to methods based on pixel differences (also called correlation or error), all methods proceeding in the gradient domain are edge following. However, the evaluation does not indicate to what degree relevant edges are followed because the relevance of an edge is usually application dependent. Consequently, the best method with respect to the edge following criterion may vary according to the considered

application. As for the multisource criterion, except for the interactive digital photomontage [2], none of the previous methodologies proposes a formal scheme for simultaneously assembling more than two images covering the same domain of overlap. However, contrary to the proposed approach, the interactive digital photomontage [2], requires all source images to cover the same domain (e.g., macro photographs taken at different focal length or time series of a scene with moving people). Beyond the proposed method, the built-in ability to remove transient objects is only detailed in [2], using a maximum likelihood objective function. All methods are crisp with the exception of the pyramid method, where multiresolution blending occurs. Except for the tie points method, all methods can be used in an automatic mode, although the selection of parameters may require some user interaction. The digital photomontage [2] and the proposed method allow for user interaction if required by the user. The multiresolution pyramid blending, the correlation watersheds, the graph cuts, and the proposed method are all suitable for the processing of 3D images. Finally, only the graph cuts and the proposed method allow for the simultaneous processing of all channels of multichannel images.

Fig. 6 shows the results obtained by the correlation watersheds [3] and the interactive digital photomontage [2] methods, starting with the same pair of input images as for Fig. 2. This example illustrates that the seam produced by the first method (automatic) fails to follow natural boundaries while the seam generated by the second method (interactive) follows local edges but misses most of the river and lake boundaries, although the markers were set close to their boundaries.

5 CONCLUSION AND PERSPECTIVES

The methodology described in this paper allows for the automatic positioning of the mosaic seam lines along salient image edges. It is generic in the sense that it can be applied to any series of overlapping images provided that they have been registered beforehand. Although it has been described

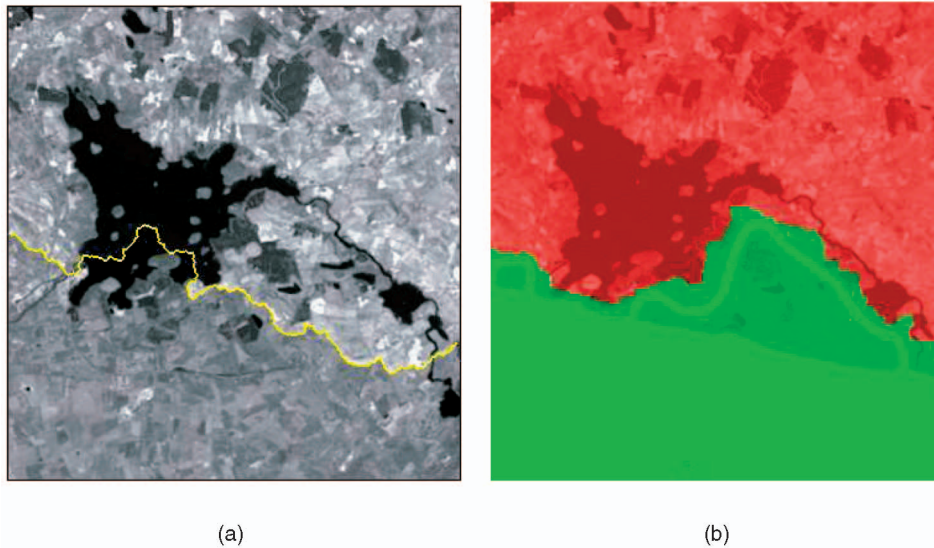


Fig. 6. Composition of the two input images shown in Fig. 2 using (a) correlation watersheds [3] and (b) interactive digital photomontage [2]. In (a), the final seam line is superimposed on the composed image. For (b), the image shown in Fig. 2b was selected as the source image. Markers were then manually drawn onto the image shown in Fig. 2a and appear in dark green in the resulting composite image.

for two-dimensional (2D) images, it is directly applicable to 3D images owing to the extension of the watershed algorithm to the processing of 3D images [40]. If required by the application, the versatility of the marker-controlled segmentation paradigm enables user interaction by adding to the marker set a series of manually drawn markers marking the core of regions that should be taken from a specific image.

Currently, we apply our methodology for automatically mosaicking Landsat scenes at continental scale. It allows for the minimization of cloud and shadow covers while reducing the visibility of seams by automatically following image structures where abrupt radiometric variations occur in all overlapping domains. The procedure is also utilized for composing images falling within a given hydrographic basin so as to automatically extract river networks without having to deal with complex border problems breaking the connectivity of these networks.

In situations where large radiometric differences occur between adjacent scenes, the seams may be noticeable even though they follow the boundaries of salient image structures. Beyond preprocessing steps reducing the effect of radiometric distortions (e.g., histogram matching, calibration, or atmospheric corrections in remote sensing), existing techniques for blending the input images along the detected seams and within a fixed distance can be applied. For example, an alpha channel [30] can be used to achieve an even mixture along the seam line and a weighting function driven by geodesic time [34] computations on each side of the seam (rather than a Euclidean distance map). The geodesic time computation is restricted to a buffer defined by the generalized geodesic dilation of the seam with respect to the morphological gradient of the image selected on each side of the seam. This leads to an adaptive spatial range where blending occurs since the spatial extent of this generalized geodesic dilation is inversely proportional to the image edge content. However, blending approaches require perfect geometric alignment and the absence of transient objects or dynamic effects which would

otherwise generate ghost artifacts. To address this issue, we are currently investigating the use of morphological interpolation techniques [33] to replace the blending procedure in the above-mentioned buffer by a gradual modification of the image values, starting from the borders of the buffer and proceeding toward its center until a medial line, i.e., the seam, is reached.

To conclude this paper, let us go back in time and quote Wolf [42, p. 240] from his 1974 book, a year before the first paper [27] reporting computer methods for creating photomosaics: "Good-quality mosaic preparation requires a person not squeamish about getting his hands wet and covered with sticky adhesive. Because it is an art, a great deal of personal satisfaction can be derived from a successfully prepared mosaic."

In a sense, this is still very true and we thus hope that the methods presented in this paper will provide "artists" with an additional type of "razor blade" and "sticky adhesive" useful for the successful preparation of mosaics on a computer.

REFERENCES

- [1] Y. Afek and A. Brand, "Mosaicking of Orthorectified Aerial Images," *Photogrammetric Eng. and Remote Sensing*, vol. 64, no. 2, pp. 115-125, Feb. 1998.
- [2] A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, and M. Cohen, "Interactive Digital Photomontage," *ACM Trans. Graphics*, vol. 23, no. 3, pp. 294-302, Aug. 2004, <http://doi.acm.org/10.1145/1015706.1015718>, source code and executable: <http://grail.cs.washington.edu/projects/photomontage/release>, 2005.
- [3] F. Araújo Jr. and N. Leite, "A Morphological Algorithm for Photomosaicking," *Proc. Eighth European Signal Processing Conf. (EURASIP)*, pp. 181-186, 1996.
- [4] S. Bagli and P. Soille, "Automatic Delineation of Shoreline and Lake Boundaries from Landsat Satellite Images," *Proc. Eco-Imagine Conf. GI and GIS for Integrated Coastal Management*, 2004, <http://www.gisig.it/eco-imaginer/abstract/Borsisti/Posters/Bagli.htm>.
- [5] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera, "Filling-in by Joint Interpolation of Vector Fields and Gray Levels," *IEEE Trans. Image Processing*, vol. 10, no. 8, pp. 1200-1211, Aug. 2001, <http://dx.doi.org/10.1109/83.935036>.

- [6] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image Inpainting," *SIGGRAPH'00: Proc. 27th Ann. Conf. Computer Graphics and Interactive Techniques (SIGGRAPH '00)*, pp. 417-424, 2000. <http://doi.acm.org/10.1145/344779.344972>.
- [7] C. Bielski and P. Soille, "Order Independent Image Compositing," *Lecture Notes in Computer Science*, vol. 3617, pp. 1076-1083, Sept. 2005. http://dx.doi.org/10.1007/11553595_132.
- [8] Y. Boykov, O. Veksler, and R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222-1239, Nov. 2001. <http://dx.doi.org/10.1109/34.969114>.
- [9] D. Brunner and P. Soille, "Iterative Area Seeded Region Growing for Multichannel Image Simplification," *Math. Morphology: 40 Years On*, C. Ronse, L. Najman, and E. Decencière, eds., *Computational Imaging and Vision*, vol. 30, pp. 397-406, Dordrecht, The Netherlands: Springer-Verlag, 2005.
- [10] P. Burt and E. Adelson, "A Multiresolution Spline with Application to Image Mosaics," *ACM Trans. Graphics*, vol. 2, no. 4, pp. 217-236. Oct. 1983. <http://doi.acm.org/10.1145/245.247>.
- [11] C.-Y. Chen and R. Klette, "Image Stitching: Comparisons and New Techniques," *Lecture Notes in Computer Science*, vol. 1689, pp. 615-622, 1999.
- [12] A. Criminisi, P. Pérez, and K. Toyama, "Region Filling and Object Removal by Exemplar-Based Image Inpainting," *IEEE Trans. Image Processing*, vol. 13, no. 9, pp. 1200-1212, Sept. 2004.
- [13] P. Dani and S. Chaudhuri, "Automated Assembling of Images: Image Montage Preparation," *Pattern Recognition*, vol. 28, no. 3, pp. 431-445, Mar. 1995. [http://dx.doi.org/10.1016/0031-3203\(94\)00106-V](http://dx.doi.org/10.1016/0031-3203(94)00106-V).
- [14] J. Davis, "Mosaics of Scenes with Moving Objects," *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR '98)*, pp. 354-360, 1998.
- [15] S. Di Zenzo, "A Note on the Gradient of a Multi-Image," *Computer Vision, Graphics, and Image Processing*, vol. 33, pp. 116-125, 1986.
- [16] E. Dijkstra, "A Note on Two Problems in Connexion with Graphs," *Numerische Mathematik*, vol. 1, pp. 269-271, 1959.
- [17] A. Efros and W. Freeman, "Image Quilting for Texture Synthesis and Transfer," *Proc. Int'l Conf. Computer Graphics and Interactive Techniques (SIGGRAPH '01)*, pp. 341-346, Aug. 2001. <http://doi.acm.org/10.1145/383259.383296>.
- [18] H. Heijmans, *Morphological Image Operators*. Boston: Academic Press, 1994.
- [19] S. Hummer-Miller, "A Digital Mosaicking Algorithm Allowing for an Irregular Join 'Line,'" *Photogrammetric Eng. and Remote Sensing*, vol. 55, no. 1, Jan. 1989.
- [20] M. Irani, P. Anandan, J. Bergen, R. Kumar, and S. Hsu, "Efficient Representations of Video Sequences and Their Applications," *Signal Processing: Image Comm.*, vol. 8, no. 4, pp. 327-351, Apr. 1996.
- [21] V. Kwatra, A. Schödl, I. Essa, G. Turk, and A. Bobick, "Graphcut Textures: Image and Video Synthesis Using Graph Cuts," *ACM Trans. Graphics*, vol. 22, no. 3, pp. 277-286, 2003. <http://doi.acm.org/10.1145/882262.882264>.
- [22] A. Levin, A. Zomet, S. Peleg, and Y. Weiss, "Seamless Image Stitching in the Gradient Domain," *Lecture Notes in Computer Science*, vol. 3024, pp. 377-389, 2004.
- [23] S. Masnou, "Disocclusion: A Variational Approach Using Level Lines," *IEEE Trans. Image Processing*, vol. 11, no. 2, pp. 68-76, Feb. 2002. <http://dx.doi.org/10.1109/83.982815>.
- [24] S. Masnou and J. Morel, "Level Lines Based Disocclusion," *Proc. Int. Conf. Image Processing*, vol. 3, pp. 259-263, 1998. <http://dx.doi.org/10.1109/ICIP.1998.999016>.
- [25] D. McNeil, "The Wet Process of Laying Mosaics," *Photogrammetric Eng.*, vol. 15, no. 2, p. 315, 1949.
- [26] F. Meyer and S. Beucher, "Morphological Segmentation," *J. Visual Comm. and Image Representation*, vol. 1, no. 1, pp. 21-46, Sept. 1990.
- [27] D. Milgram, "Computer Methods for Creating Photomosaics," *IEEE Trans. Computers*, vol. 24, no. 11, pp. 1113-1119, Nov. 1975.
- [28] D. Milgram, "Adaptive Techniques for Photomosaicking," *IEEE Trans. Computers*, vol. 26, no. 11, pp. 1175-1180, Nov. 1977.
- [29] S. Peleg, "Elimination of Seams from Photomosaics," *Computer Graphics and Image Processing*, vol. 16, pp. 90-94, 1981.
- [30] T. Porter and T. Duff, "Compositing Digital Images," *Computer Graphics*, vol. 18, no. 3, pp. 253-259, July 1984.
- [31] R. Sedgewick, *Algorithms in C: Fundamentals, Data Structures, Sorting, Searching, and Graph Algorithms. Part 5: Graph Algorithms*, third ed. Addison-Wesley, 2001.
- [32] J. Serra, *Image Analysis and Mathematical Morphology*. London: Academic Press, 1982.
- [33] P. Soille, "Spatial Distributions from Contour Lines: An Efficient Methodology Based on Distance Transformations," *J. Visual Comm. and Image Representation*, vol. 2, no. 2, pp. 138-150, June 1991.
- [34] P. Soille, "Generalized Geodesy via Geodesic Time," *Pattern Recognition Letters*, vol. 15, no. 12, pp. 1235-1240, Dec. 1994. [http://dx.doi.org/10.1016/0167-8655\(94\)90113-9](http://dx.doi.org/10.1016/0167-8655(94)90113-9), download PDF preprint: <http://ams.jrc.it/soille/soille94.pdf>.
- [35] P. Soille, "On the Morphological Processing of Objects with Varying Local Contrast," *Lecture Notes in Computer Science*, vol. 2886, pp. 52-61, Nov. 2003. <http://dx.doi.org/10.1007/b94107>.
- [36] P. Soille, *Morphological Image Analysis: Principles and Applications*. Berlin and New York: Springer-Verlag, corrected second printing of second ed., 2004.
- [37] P. Soille, "Beyond Self-Duality in Morphological Image Analysis," *Image and Vision Computing*, vol. 23, no. 2, pp. 249-257, Feb. 2005. <http://dx.doi.org/10.1016/j.imavis.2004.06.002>.
- [38] P. Soille and M. Pesaresi, "Advances in Mathematical Morphology Applied to Geoscience and Remote Sensing," *IEEE Trans. Geoscience and Remote Sensing*, vol. 40, no. 9, pp. 2042-2055, Sept. 2002.
- [39] P. Soille and H. Talbot, "Directional Morphological Filtering," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1313-1329, Nov. 2001. <http://dx.doi.org/10.1109/34.969120>.
- [40] L. Vincent and P. Soille, "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, no. 6, pp. 583-598, June 1991. <http://dx.doi.org/10.1109/10.1109/34.87344>.
- [41] T. Welch, "A Technique for High Performance Data Compression," *Computer*, vol. 17, no. 6, pp. 8-19, June 1984.
- [42] P. Wolf, *Elements of Photogrammetry (with Air Photo Interpretation and Remote Sensing)*. New York: McGraw-Hill, 1974.
- [43] B. Zitová and J. Flusser, "Image Registration Methods: A Survey," *Image and Vision Computing*, vol. 21, pp. 997-1000, 2003.



Pierre Soille received the engineering degree in 1988 from the Université catholique de Louvain, Belgium. Research on mathematical morphology for the processing of digital elevation models and satellite images provided the material for the doctorate degree he received in 1992 from the same university and in collaboration with the Centre de Morphologie Mathématique of the Ecole des Mines de Paris. He then pursued

research on image analysis and mathematical morphology at the CSIRO Mathematical and Information Sciences, Sydney, Australia; the Centre de Morphologie Mathématique, Fontainebleau, France; and the Pattern Recognition Department of the Fraunhofer-Institut IPK, Berlin. In 1994, he coorganized with Jean Serra the second International Symposium on Mathematical Morphology. During the period 1995-1998, he has been an assistant professor at the Ecole des Mines d'Alès and EERIE, Nîmes, France. He received the Habilitation à Diriger des Recherches degree in 1997 from the Université Montpellier II, France. In 1999, he led a project on the application of novel image analysis techniques to agri-food industries at the Silsoe Research Institute of the Biotechnology and Biological Sciences Research Council, United Kingdom. Since December 1999, he has been with the Joint Research Centre of the European Commission, Ispra, Italy, focusing on the research of advanced methodologies and algorithms for the generation of enhanced pan-European geospatial data sets. He is the author of the reference monograph *Morphological Image Analysis: Principles and Applications*, published in English and German by Springer-Verlag, and editor of *Morphology Digest*, an electronic newsletter devoted to morphology (<http://ams.jrc.it/mdigest>).

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.