

Perceptual filtering with connected operators and image inpainting

MARIELLA DIMICCOLI, AND PHILIPPE SALEMBIER^a

^a *Technical University of Catalonia (UPC), Barcelona, SPAIN*
mariella,philippe@gps.tsc.upc.edu

Abstract This paper focuses on a class of morphological filtering tools called connected operators. These operators act by merging flat zones that do not fulfill a given simplification criterion. This filtering approach offers the advantage of simplifying the image, because some flat zones are removed, as well as preserving contour information, because the flat zones that are not removed are perfectly preserved. However, for some applications, connected operators present a drawback in the way they restore the areas where flat zones have been merged. These areas may be perceptible in the filtered image appearing as single flat zones inside smoothed regions or across edges. To overcome such drawback, this paper proposes a solution based on the use of image inpainting.

Keywords: connected operators, image inpainting.

1. Introduction

In the context of image segmentation or region-based image analysis, the purpose of a filter is often to remove some image details that do not fulfill a certain simplification criterion. Many classical image filtering strategies are based on the use of a specific signal $h(x)$, such as an impulse response or a structuring element, which modifies the pixel values in a local window. However, these filters introduce distortions in the output because the signal $h(x)$ is not related at all with the input signal. A completely different approach is taken by morphological connected operators [1]. These operators act directly on the partition of flat zones of the image by removing and merging those flat zones that do not fulfill the simplification criterion. Thanks to this filtering approach, connected operators cannot introduce any contour distortion related to a specific signal and, as a consequence, they are attractive in a large number of applications where the image has to be simplified without losing information about contours [2, 4, 5, 11, 14, 16, ?Vicent93].

However, connected operators present a limitation in the way they restore the areas where flat zones have been merged. These areas are always restored as single flat zones that may be perceptible in the filtered image. The basic idea presented in this paper is of removing the perceptual presence of the areas where flat zones have been merged using image inpainting [15, 18]. Image inpainting is to restore missing or inaccessible image pixels in a plausible way based upon the available information. The use of this technique would allow to estimate the areas where flat zones have been merged as natural extension of their surrounding.

The organization of the paper is as follows. The next section provides an introduction to connected operators and analyzes the perceptual effect of the restoring strategy. Section 3 is devoted to image inpainting and a practical algorithm is described. In section 4, the proposed technique is presented and its structure is discussed. In section 5, several examples are reported and performances are evaluated. Finally, section 6 is devoted to the conclusions.

2. Connected Operators

2.1 Construction strategies

Connected operators are morphological filtering tools that can simplify part of the image content, while preserving the contours of the remaining parts of the image. This property is a direct consequence of their definition: gray level connected operators filter the image by removing and merging those flat zones that do not fulfill a given simplification criterion. The most successful strategies to construct connected operators rely on a reconstruction process or on a region-tree pruning. The reconstruction process [1, 2, 11, ?Vicent93] is called anti-extensive, extensive or self-dual, depending on the image components it allows to simplify. Anti-extensive and extensive reconstruction processes deal with either bright or dark image components respectively, whereas self-dual reconstruction deals with all components in a symmetrical way. The self-dual reconstruction [11], also called *leveling*, is defined as follows: if u and v are two images (respectively called the *reference* and the *marker* image), the self-dual reconstruction ρ of v with reference u is defined by: $\rho(v|u) = \lim_{n \rightarrow \infty} v_n$ and $v_n = \epsilon_0(v_{n-1}) \vee [\delta_0(v_{n-1}) \wedge u]$, where ϵ_0 and δ_0 represent respectively an erosion and a dilation with square or a cross of 3x3, \vee and \wedge the infimum and supremum and $v_0 = v$.

An example of self-dual reconstruction is shown in Fig.1. In this example, the marker image is constant everywhere except for two points that mark a maximum and a minimum of the reference image. After reconstruction, the output has only one maximum and one minimum and their contours coincide with those of the reference signal.

In practice, useful connected operators are obtained by considering that the marker v is a transformation $\Phi(u)$ of the input image. The trans-

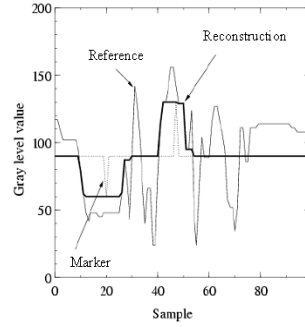


FIGURE 1. Example of self-dual reconstruction

formation Φ defines the simplification effect and the reconstruction process restores the contour of the flat zones that have not been completely removed by the simplification. As a result, most connected operators obtained by reconstruction can be written as: $\psi(u) = \rho(\Phi(u) \mid u)$, where ρ represents a generic reconstruction process. Examples of filtering effect include size oriented (resp. a contrast-oriented) simplification if Φ is an erosion or a dilation with a structuring element (resp. a subtraction or an addition of a constant gray level value).

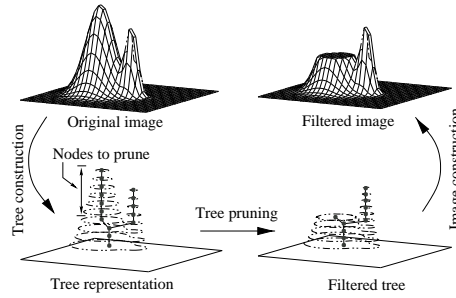


FIGURE 2. Max-tree processing of images

Region-tree pruning [4, 5] strategies offer an alternative way to create connected operators. The general approach involves three steps (Fig. 2). The first step consists in creating the tree representation of the image: the tree root represents the entire support of the image, the tree nodes represent image regions and the tree branches represent the inclusion relationships among the nodes. The second step is the filtering itself, which analyzes each node and takes a decision on which node has to be preserved and which has to be removed. Finally, the last step restores the filtered image by transforming the pruned tree into a gray level image. Classical trees suitable for connected operators are the Max-tree/Min-tree [4], the Binary Partition Tree (BPT) [5] and the Component Tree [6]. The nodes of a Max-tree (Min-tree) represent the connected components of level sets an image is made of,

and are obtained iteratively by thresholding the image at all possible gray levels. The nodes of a BPT represent the regions that can be obtained from an initial partition by merging neighboring regions following an homogeneity criterion until the tree root is obtained. The nodes of a component tree represent connected components of upper- or lower level sets with their holes filled, also called "shapes", which are obtained putting together bounded connected components of both upper and lower level sets. The interest of the tree representation is in that the tree structure is fixed and represented by the tree branches. As a consequence, sophisticated pruning strategies can be designed allowing to deal in particular with non-increasing criteria. Formally, a criterion \mathbb{C} assessed on a region R is said to be increasing if the following property holds: $\forall R_1 \subseteq R_2, \mathbb{C}(R_1) \subseteq \mathbb{C}(R_2)$. In the increasing case, there is a relation between the criterion value of a node and that of its descendants in the sense that if a node has to be removed all its descendants have also to be removed. In the case of non-increasing criteria, this relation does not hold and this fact implies a lack of robustness of the operator. Some practical rules have been reported in the literature to deal with the non-increasing case. One strategy is to formulate the problem as a dynamic programming issue and to solve it with the Viterbi algorithm [4, 5].

Summarizing, the region-tree pruning approach is conceptually more complex than the reconstruction approach. However, it leads to a very efficient implementation of connected operators. It allows dealing with non-increasing criteria and provides more flexibility in the choice of the simplification criterion.

2.2 Perceptual analysis

Even if connected operators allow image simplification without introducing distortion on the remaining contours, they may not perceptually remove the presence of areas where flat zones have been merged. Fig.3 illustrates this effect with a size-oriented connected operator: $\psi(u) = \rho(\epsilon(u) \mid u)$, where ϵ is an erosion with a square structuring element of size 5x5 (over a 512x512 image). This operator is known as the opening by reconstruction of erosion [Vicent93]. Its goal is to remove small maxima.

In the sequel, the use of inpainting to estimate the pixel values of areas where flat zones have been merged is proposed and discussed.

3. Image Inpainting

3.1 A variational approach

In the field of image processing, the term inpainting refers to the task of restoring the pixel values for a missing or consciously masked sub-region of the image domain. The basic idea of the algorithms proposed in the literature [8–10, 15, 17–19] is to restore the missing regions with the information

available from their surrounding. Depending on their goal and on the surrounding information they use, available methods can be broadly classified in *structural* or *textural* inpainting. Structural inpainting [15, 17–19] restores the geometric structures of the image using contour information. Textural inpainting [8–10] restores the texture of the image using patterns or texture exemplars. Since the use of a connected operator guarantees that at least the contour information is preserved, a structural inpainting algorithm has been used in this work [17]. However, in principle, any state of art inpainting algorithm can be used. The authors formulate the inpainting as the problem of restoring, in the regions of missing data, both the geometric and the photometric information represented respectively by the level lines and the gray level values. Let Ω be the region to be inpainted, and B a narrow band of pixels surrounding Ω . In order to extrapolate the shape information independently from the contrast, the image in $\tilde{\Omega} = (\Omega \cup B)$ is decomposed into level sets U_λ , which are inpainted individually. For each level set U_λ the solution on Ω is obtained minimizing the following functional over $\tilde{\Omega}$:

$$E_\lambda = \int_{\tilde{\Omega}} |\operatorname{div}(\theta)|(a + b|\nabla U_\lambda|)dx + \alpha \int_{\tilde{\Omega}} (|\nabla U_\lambda| - \theta \cdot \nabla U_\lambda)dx \quad (1.1)$$

where a , b and α are positive constants and ideally θ represents the normal vector field to the level lines of U_λ , that is: $\theta = \frac{\nabla U_\lambda}{|\nabla U_\lambda|}$. This constraint is expressed with the second integral term. The level set extrapolation is a function of the geometric quantities that appear in the functional over $\tilde{\Omega}$: the curvature of level lines and the perimeter of the discontinuities represented respectively by the quantities $\operatorname{div}(\theta)$ and $(a + b|\nabla U_\lambda|)$ in the first integral term. Finally, the solution is obtained by stacking the extrapolated level sets.

3.2 Optimization algorithm

Numerically, the minimization of the functional (1.1) is computed solving the variational problem via an iterative algorithm based on the gradient

Flat zones corresponding to the writing "MPEG4 WORLD" have been removed and merged into a single one whose gray level value depends on the surrounding flat zones. However, the region of support, that is the text, is still visible in the filtered image. In some applications as segmentation or editing, it is desirable to remove regions from the original image without leaving any perceptual information about them.



FIGURE 3. Drawback of connected operators

descent flow. Since the iteration number depends on the size of Ω and on the homogeneity of the level lines in B , a termination criterion has to be fixed. However, this aspect has not been addressed by the authors and is often neglected in the literature. The solution used in this work has been of considering the variation of the energy as a function of the iteration number and of fixing the termination criterion by a threshold on its slope. In practice, when the local slope of the functional variation is sufficiently close to 0, it is assumed that the algorithm has converged.

4. Proposed Approach

In this section, a new filtering strategy allowing to overcome the drawback of connected operators is proposed. The filtering process (Fig. 4) involves two major steps: simplification by a connected operator and restitution of the perceptually most important removed flat zones by inpainting. The intermediate step consists in computing the mask marking the regions to be inpainted. The question that may arise now is to know which flat zones, among the new ones created by the filtering process, need to be estimated by inpainting since their are perceptible. The method used in this work is as follows. First, the residue I_3 is binarized with threshold 1. Second, the mask obtained in this way is simplified by removing the connected components for which the residue remains smaller than a fixed threshold. The value of this second threshold guarantees the visibility of the simplification effect. In fact, the intuition behind this solution is that there exists a relationship between the perceptibility of the new flat zones and the visibility of the action of the connected operator. Over the mask obtained after the binarization at two levels (M_1), a closing is applied to merge regions which are very closed to each other (M_2). This operation guarantees that the inpainting algorithm could dispose of a band of at least tree pixels in order to compute geometric features. When the boundary of a flat zone to be inpainted corresponds to an object boundary in the original image, it is often surrounded by a transition zone of at least one or two pixels. This transition zone may disturb the process of capturing the geometry. To avoid this problem, before the inpainting, the regions of the filtered image marked by the complementary of the mask M_2^C are simplified using an opening by reconstruction of erosion (structuring element of size 3x3 or 5x5), followed by its dual, the closing by reconstruction of dilation.

After inpainting, both the simplified version of the band and the inpainted regions marked by the mask are inserted in the image I_2 . In order to copy also the simplified band, before of applying the toggle mapping, the mask M_2 is dilated of a size equal to the size of the band used by inpainting. This strategy guarantees that inpainted areas will be perceived as a smooth extension of the visual information contained in the band.

Fig.5 shows an example of the result obtained using the proposed ap-

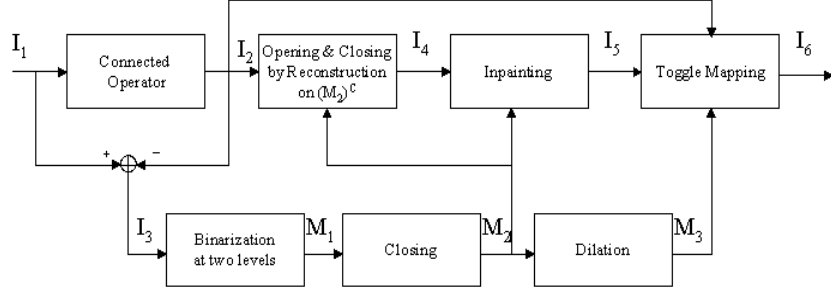


FIGURE 4. Proposed filtering approach. I refers to an image.; M refers to a mask; M^C refers to the complement of M

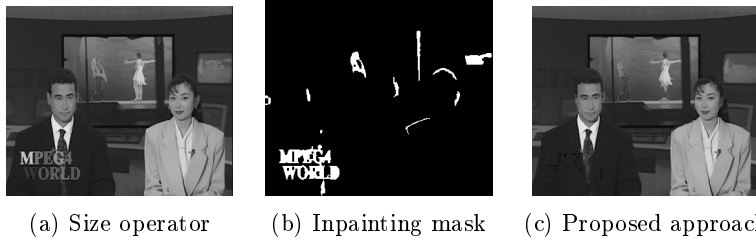


FIGURE 5. Example of filtering with the size criterion

proach. First, the original image I_1 is filtered using a size-oriented connected operator (Fig. 5 (a)). Second, a mask, defining the perceptually most important regions that have been removed by the connected operator, is computed and the regions that are very close to each other are merged by a closing. The resulting mask M_2 is shown in Fig. 5 (b). Third, the filtered image I_2 is smoothed in the areas surrounding the regions marked by the mask using an opening by reconstruction followed by its dual, the closing by reconstruction. Fourth, the regions of this smoothed image I_4 marked by the mask M_2 are inpainted. Finally, a toggle mapping copies the regions estimated by inpainting, as well as their smooth bands in the image I_2 . As can be noticed, the connected components marked by the mask (Fig. 5 (b)), as for instance the writing and the legs of the dancer, visible in Fig. 5 (a), are no longer perceived in Fig. 5 (c), since they have been completely removed by inpainting.

5. Experimental Results

5.1 Perceptual filtering examples

This section presents some results obtained using the proposed filtering approach. A first example considering size simplification has been shown in the section 4. The second example involves a contrast simplification.

The contrast simplification is obtained using a λ -max operator: $\psi(u) =$

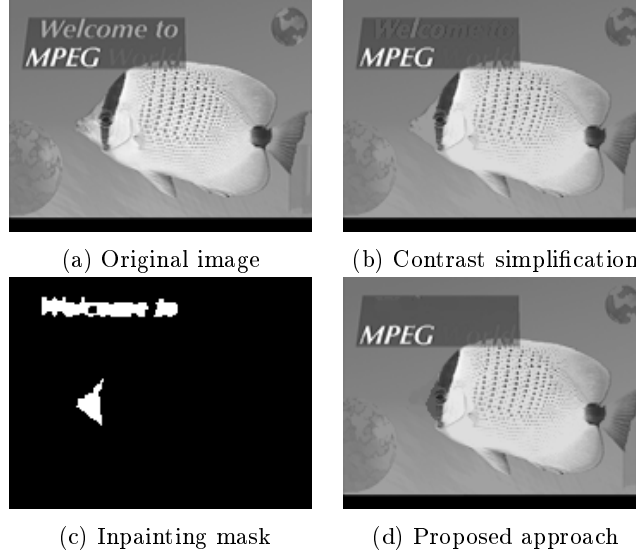


FIGURE 6. Example of filtering with contrast criterion

$\rho(u - c \mid u)$. This operator allows to remove all maxima (or minima) with contrast inferior to λ . In this example, the effect of the contrast simplification (with $\lambda = 100$) is specially visible in areas such as the writing "Welcome to" and the muzzle of the fish (Fig. 6 (b)). However, these regions are still perceptible. Instead, applying inpainting to the regions marked by the mask (Fig. 6 (c)), the perceptual presence of these areas is removed (Fig. 6 (d)).

The third example involves a motion criterion in image sequences [13]. Motion operator allows to remove from an image all objects that do not undergo a given motion. In this example, the operator has been applied to remove all moving objects from the image shown in Fig. 7 (a). In the considered sequence, the perceptually most important moving object is the dancer behind the two speaker. As can be observed in Fig. 7 (b), although flat zones corresponding to the dancer have been removed, they appear in the filtered image as a large flat zone inside a smoothed region. Using inpainting to estimate pixel values in this area (Fig. 7 (c)), the perception of this flat zone is completely removed (Fig. 7 (d)).

The last example involves an image obtained superimposing white lines to the original image. The goal of this example is to use the original image as reference to perform an objective quality assessment. Assume that the goal is to remove the superimposed lines from the image shown in Fig. 9 (a). Due to their shape, the white lines are suited to be removed using a complexity criterion [12]. This criterion is based on a measure of the ratio between the perimeter P and the area A of the connected component. Intuitively, if a connected component has a small area but a very long perimeter, it corresponds to a complex object. However, a complexity criterion alone

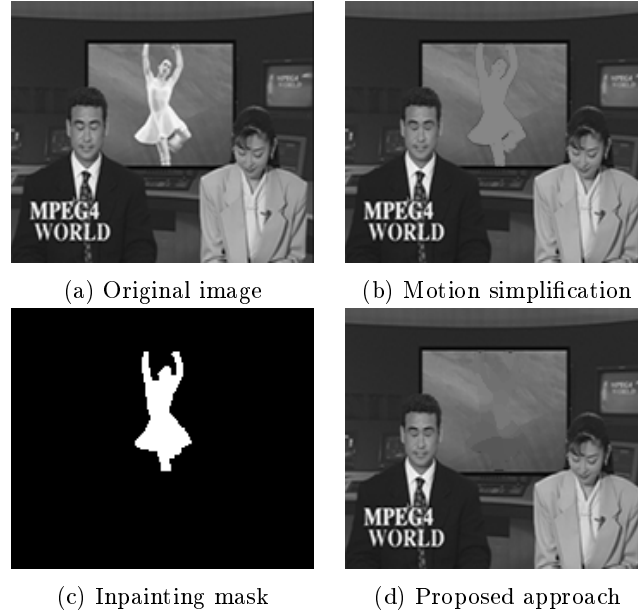


FIGURE 7. Example of filtering with a motion criterion

would remove also a large number of complex flat zones.

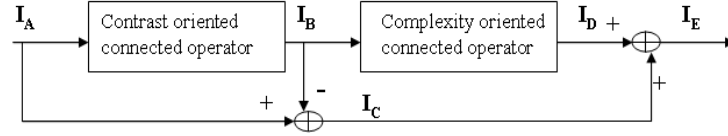


FIGURE 8. Connected operator used to extract the white lines

In order to extract the white lines removing the smallest number of other flat zones, the connected operator shown in Fig.8 has been used. This operator is based on the observation that the white lines are complex and highly contrasted, whereas other complex flat zones correspond to texture which is less visible because of low contrast. First, the image shown in Fig. 9 (a) is filtered using a contrast-oriented connected operator, which removes all maxima having contrast inferior to 170. This is the maximum value of the contrast allowing to preserve the white lines. Second, the resulting image is filtered by a complexity criterion, with complexity 42, that is the minimum value allowing to remove the white lines. Note that the complexity criterion is not increasing because there is not a relationship of complexity between two connected components R_1, R_2 such that $R_1 \subset R_2$. In order to deal with this non increasing criterion the "Max" decision has been employed. The

"Max" decision is defined as follows: a node C for which the evaluation of the criterion is lower than a given threshold is not removed if at least one of its descendants has to be preserved. Finally, the difference I_C between the original image I_A and the image filtered using a contrast criterion I_B is added to the filtered image using a complexity criterion I_D to restore the texture.

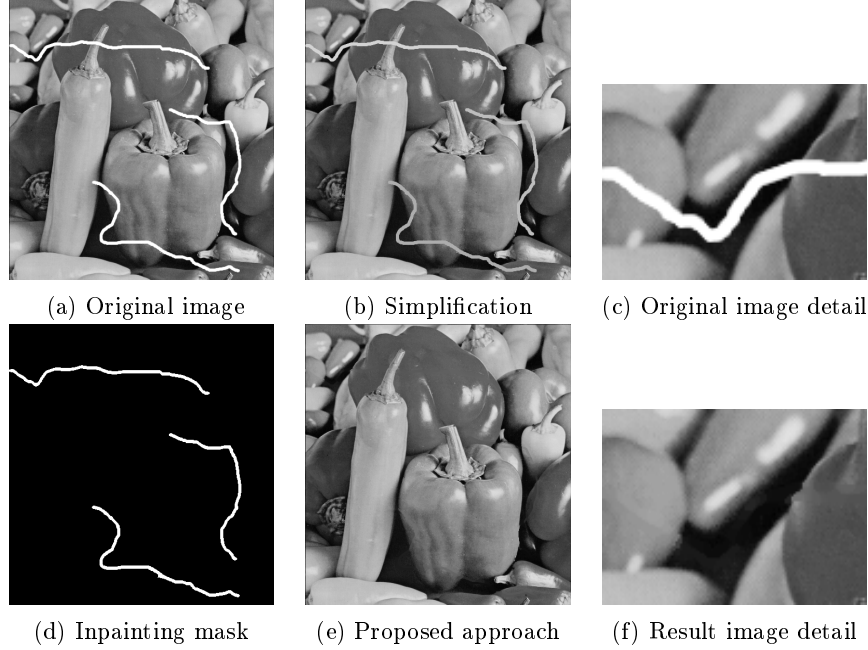


FIGURE 9. Example of filtering with the connected operator of Fig. 8

As can be observed in Fig. 9 (b), flat zones corresponding to the white lines have been removed and most of the texture preserved. However, the perceptual presence of the superimposed lines is still strongly visible. Instead, applying inpainting to the perceptual more important removed regions (Fig. 9 (c)) a much better result is obtained (Fig. 9 (d)).

5.2 Performances evaluation

In this section the quality performances of the proposed approach are compared objectively to classical filtering strategies using two well defined criteria: the peak signal-to-noise ratio (PSNR) and the Structural Similarity Image Measure (SSIM) [7, 20].

The PSNR relies on the mean square error (MSE) which is the square of the Euclidean distance between a reference and a distorted image. This definition does not include any perceptual features and therefore MSE and

TABLE 1. PSNR and SSIM values obtained using different filtering approaches.

<i>Type of filter</i>	<i>PSNR value</i>	<i>SSIM value</i>
Low pass filter (11x11 average) (5x5 average)	19.8	0.06
Median filter (15x15) (5x5)	23.2	0.87
Connected operator	21.9	0.91
Proposed approach	39.2	0.99

PSNR do not really assess the perceived image quality. To evaluate the perceived image quality, the SSIM has been used. The SSIM relies on the assumption that human vision is highly sensitive to structural information and, as consequence, a measure of the structural information change should provide a good approximation of perceived image distortion. In practice, the SSIM quantifies the differences between a distorted image and a reference image independently from average luminance and contrast. The SSIM is equal to 1 for two identical images. Performances have been evaluated for the last example described in section 5, for which the reference image (image without the superimposed lines) is available. Table 1 shows the PSNR values as well as the SSIM values obtained using different filtering techniques. In the case of median and low pass filter, the table reports the best value obtained using different sizes of window ranging from 5x5 to 25x25. As can be observed, the proposed scheme drastically increases both the PSNR and the SSIM, meaning that the approach proposed in this paper reduces the error visibility and gives results consistent with the qualitative visual appearance.

The proposed filtering strategy is computationally more complex than connected operators, depending its complexity on the inpainting technique used. However, it permits to achieve a good trade-off between quality result and computational cost.

6. Conclusions

In this paper, a new filtering technique improving connected operator performances has been presented and discussed. The proposed technique involves two broad steps: first, the image is simplified using connected operators. Second, the perceptually most important filtered regions are estimated using inpainting. The mask marking the regions to be inpainted is automatically computed and no user interaction is required. Comparative experiments have shown that the proposed technique outperforms classical filtering strategies in terms of both visibility of error (PSNR) and structural perceptual quality (SSIM). The presented approach is general in the sense that any connected operator can be used. As a result, it is suitable for a large set of advanced filtering applications such as objects, writing or defects removal. The future work will be devoted to improve the current method

of selecting the regions to be inpainted.

References

- [1] P.Salembier and J.Serra, *Flat zones filtering, connected operators and filters by reconstruction*, IEEE Trans. on Image Processing **3(8)** (1995), 1153-1160.
- [2] J.Crespo and J.Serra and R.W.Schafer, *Theoretical aspects of morphological filters by reconstruction*, IEEE Trans.on Image Processing **47(2)** (1995), 201-225.
- [3] L.Vicent, *Morphological gray scale reconstruction in image analysis: Applications and efficient algorithms*, IEEE Trans. on Image Processing **2(2)** (1993), 176-201.
- [4] P.Salembier and A.Oliveras and L.Garrido, *Anti-extensive connected operators for image and sequence processing*, IEEE Trans. on Image Processing **7(4)** (1998), 555-570.
- [5] P.Salembier and L.Garrido, *Binary partition tree as an efficient representation for image processing, segmentation and information retrieval*, IEEE Trans. on Image Processing **7(4)** (2000), 561-576.
- [6] P.Monasse and F.Guichard, *Scale-Space from a Level Lines Tree*, Journal of Visual Com. and Image Rep. **11** (2000), 224-236.
- [7] Z.Wang and A.Conrad Bovik and E. P.Simoncelli, *Image Quality Assessment: from Error Visibility to Structural Similarity*, IEEE Trans. on Image Processing **13(4)** (2004), 600-612.
- [8] A.Efros and W.T.Freeman, *Texture synthesis by non-parametric sampling*, Proc. of ACM Conf. on Computer Graphics, SIGGRAPH (2001), 341-346.
- [9] M.Ashikhmin, *Synthesizing natural textures*, Proc. ACM Symposium on Interactive 3D Graphics (2001), 217-226.
- [10] A.Efros and T.Leung, *Texture Synthesis by non-parametric sampling*, Proc. of Int. Conf. on Computer Vision, ICCV (1999), 21033-1038.
- [11] F.Meyer, *The levelings*, Proc. of Int.Symposium on Mathematical Morphology, ISMM, vol. 2, 1997, pp. 211-214.
- [12] A.Oliveras and P.Salembier, *Generalized connected operators*, Proc. of Visual Communications and Image Processing, SPIE, vol. 2727, 1996, pp. 761-772.
- [13] L.Garrido and A.Oliveras and P.Salembier, *Motion analysis of image sequences using connected operators*, Proc. of Visual Communications and Image Processing, SPIE, vol. 3024, 1997, pp. 546-557.
- [14] G.K.Ouzounis and M.H.F.Wilkinson, *Second order attribute filters using Max-Trees*, Proc. of Int. Symposium on Mathematical Morphology, ISMM, 2005, pp. 65-74.
- [15] S.Masnou and J.-M.Morel, *Level-lines based disocclusion*, Proc. of Int. Conf. on Image Processing, ICIP, vol. 3, 1998, pp. 259-263.
- [16] H.Heijmans, *Connected Morphological Operators and filters for binary images*, Proc. of Int. Conference on Image Processing, ICIP, vol. 2, 1997, pp. 211-214.
- [17] C.Ballester and M.Bertalmio and V.Caselles and G.Sapiro and J.Verdera, *Filling-in by joint interpolation of vector fields and grey levels*, IEEE Trans. on Image Processing **10** (2001), 1200-1211.
- [18] M.Bertalmio and G.Sapiro and V.Caselles and C.Ballester, *Image Inpainting*, Proc. 27th Annual Conf.on Computer Graphics and Interactive Techniques, 2000, pp. 417-424.
- [19] T.F.Chan and J.Shen, *Mathematical models for local nontexture inpaintings*, SIAM Journal of Applied Mathematics **62** (2001), 1019-1043.
- [20] Z.Wang, *The SSIM index for Image Quality Assessment*, MATLAB implementation available online from: <http://www.cns.nyu.edu/~lcv/ssim/>.