



# Image inpainting with nonsubsampling contourlet transform



Rosa Rodríguez-Sánchez\*, J.A. García, J. Fdez-Valdivia

Departamento de Ciencias de la Computación e I. A., CITIC-UGR, Universidad de Granada, 18071 Granada, Spain

## ARTICLE INFO

### Article history:

Received 2 July 2012

Available online 14 June 2013

Communicated by A. Heyden

### Keywords:

Nonsubsampling contourlet transform

Inpainting

Multiscale representation

Broken edges

Object removal

## ABSTRACT

A new algorithm for image inpainting based on the searching of redundancy for corner points across different scales and orientations is proposed. The searching utilises the nonsubsampling contourlet transform (NSCT) of the original image. The target region is filled-in following the priority which is given to the corner points that accumulate a high energy at different orientations and scales, and also having a high level of confidence. In each iteration, the source patch chosen is copied onto the target patch producing the minimum dispersion across different scales and orientations of the NSCT within it. This approach – to firstly fill-in corner points of high energy and confidence which minimize the dispersion in the target patch across different scales and orientations – have been tested on two groups of images, one group with broken edges and objects and a second group with large objects to be erased. The proposed algorithm was compared with three other methods and results obtained indicate that the capability of this algorithm is sustainable.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

Several definitions of image inpainting can be found in the literature but the most intuitive might be as follows: the technique of modifying an image in an undetectable form (Bertalmio et al., 2000). Image inpainting has been applied on different areas from digital photography for restoration until compressed sensor theory (Kahge et al., 2012). In the movies context has been used as a mechanism to video post-production, e.g., stereo image inpainting (Hervieu et al., 2011; Arias et al., 2012) address the reconstruction of missing information in pairs of stereo images.

One the basis of the image inpainting comes from its relationship with texture synthesis employed to repair digitized images (Efros and Leung, 1999). In this form, these algorithms normally copy image patches, or pixels, from the original image while avoiding blurring. However, these texture synthesis algorithms are unable to extend to lineal structure.

The objective of image inpainting is to translate the information from the surroundings to the area to be filled (target region). These methods are based on a partial differential equation (PDE) which is objective to the extension of the lineal structure using the isophote information (Bertalmio et al., 2000; Criminisi et al., 2003; Chen et al., 2007; Chan and Shen, 2001). The importance of the order in filling the target region, has been related in different papers (Criminisi et al., 2003; Wu and Ruan, 2006). In these papers, two terms to establish the priority of the information to be filled are

used, the data term and the confidence term. The data term is a measure of the isophote strength hitting the boundary of the target region. The confidence term gives the amount of credible information that surrounds a pixel in the target region.

On the other hand, different algorithms of texture synthesis have used analysis multiresolution approaches, such as the Laplacian pyramid (De Bonet, 1997) or wavelet decomposition (Simoncelli and Portilla, 1998). In Yamauchi et al. (2003), the image inpainting and texture synthesis were combined in a multiresolution approach by using a discrete cosine transform and a Gaussian decomposition. Li et al. (2013) proposed a model of inpainting image using DCT-Haar multiresolution analysis with applications to impulsive noise removal and filling missing information on moderate-sized regions.

If we try to modify an image in an undetectable way, we can characterize the result of image inpainting in the target region as a smooth zone and/or with a small dispersion. Thus when the source patch is copied to the target patch, the production of a small dispersion was taken into account in Goyal et al. (2010). For this, the variance must be minimized when the source patch is copied to the target patch.

An inpainting method that can be used in images as well as in video was presented in Wexler et al. (2007). In this algorithm, the fill-in of the target region occurs by choosing the value of each pixel according to its neighbourhood. The accuracy of a pixel value depends on whether its local neighbourhood forms a coherent structure. Both Goyal et al. (2010) and Wexler et al. (2007) assert that the sum of squared differences (SSD) of colour information, which is so widely used for image completion, is not sufficient to establish the similarity between two patches.

\* Corresponding author. Tel.: +34 958 244019; fax: +34 958 243317.

E-mail addresses: [rosa@decsai.ugr.es](mailto:rosa@decsai.ugr.es) (R. Rodríguez-Sánchez), [jags@decsai.ugr.es](mailto:jags@decsai.ugr.es) (J.A. García), [jfv@decsai.ugr.es](mailto:jfv@decsai.ugr.es) (J. Fdez-Valdivia).

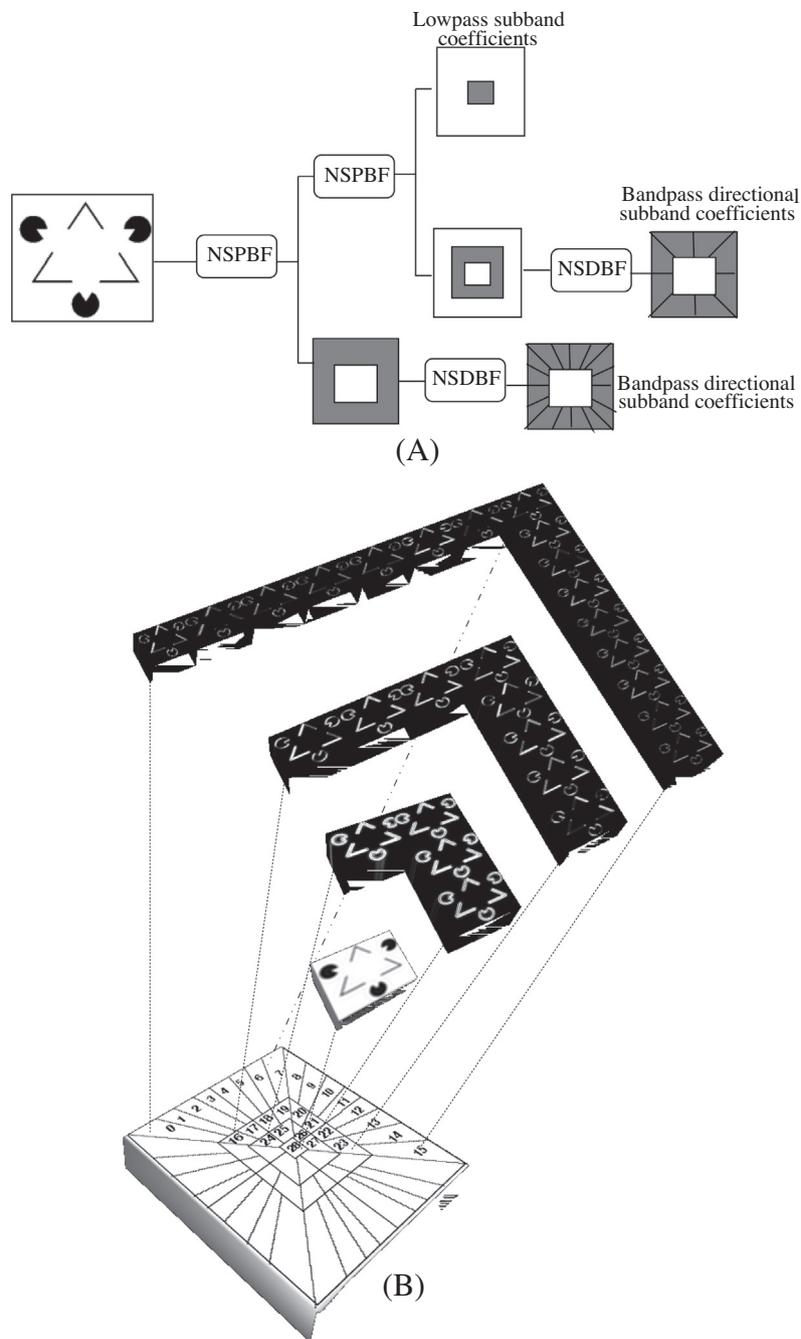


Fig. 1. Nonsampled contourlet transform: (A) scheme of the NSCT; and (B) idealized frequency partitioning together with the response for each NSCT band.

Here we show the importance of the order of filling of the target region, as seen previously in Criminisi et al. (2003) and Wu and Ruan (2006). But instead it is based on the corner points having a high enough confidence level and high energy across orientations and scales for the nonsampled contourlet transform.

Also we demonstrate that selection criteria which are aimed only to minimize the sum of squared differences between the source and the target patch can lead to results with many artefacts. To overcome this problem, instead our method chooses the source patch in two steps: First, a source-patch candidate set is selected; and second, the candidate that produces a minimal dispersion across different orientations and scales for the nonsampled contourlet transform is obtained.

This paper is organized as follows. Section 2 reviews the nonsampled contourlet transform (NSCT). Section 3 presents a novel technique for image inpainting using the nonsampled contourlet transform. A complete set of experiments is next presented in Section 4. There, the method proposed is compared with three different models: the Criminisi's algorithm (Criminisi et al., 2003), the Wexler's algorithm (Wexler et al., 2007) and the Goyal's algorithm (Goyal et al., 2010). Section 5 concludes.

## 2. Nonsampled contourlet transform

In order to recover a target region for a given image, we have to eliminate transients from some undesired object or

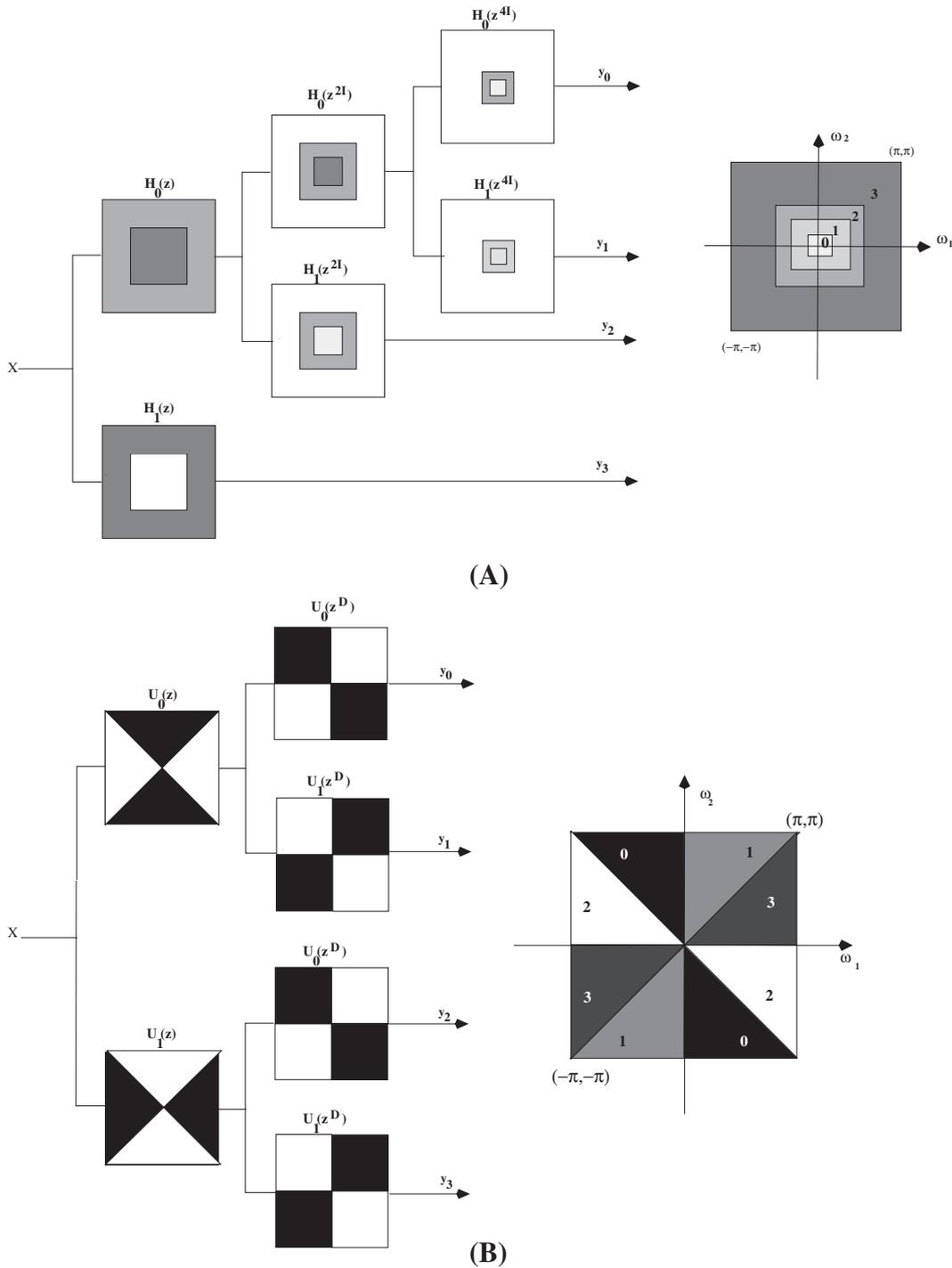


Fig. 2. Nonsampled pyramid and nonsampled directional filter bank. (A) Left: pyramidal decomposition. Right: frequency partitioning. (B) Left: four-channels NSDFB built with two-channel fan filter bank. Right: idealized frequency partitioning.

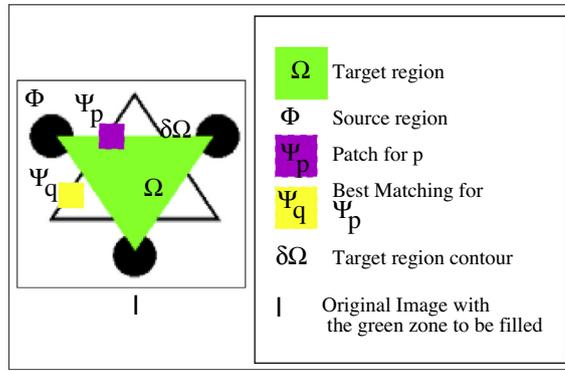
noise while producing a nearly stationary signal which maintains the number of transients at least in the target region.

In this context, signal theory offers a variety of transforms that allow us to distinguish where these transients occur and where they are absent. They include the Gabor (Daugman, 1988), the wavelets (Daubechies, 1992) and the contourlet (Do and Vetterli, 2005) transforms. All of these transforms offer an efficient representation in the sense that they have the ability to capture significant information of an object of interest given a small description.

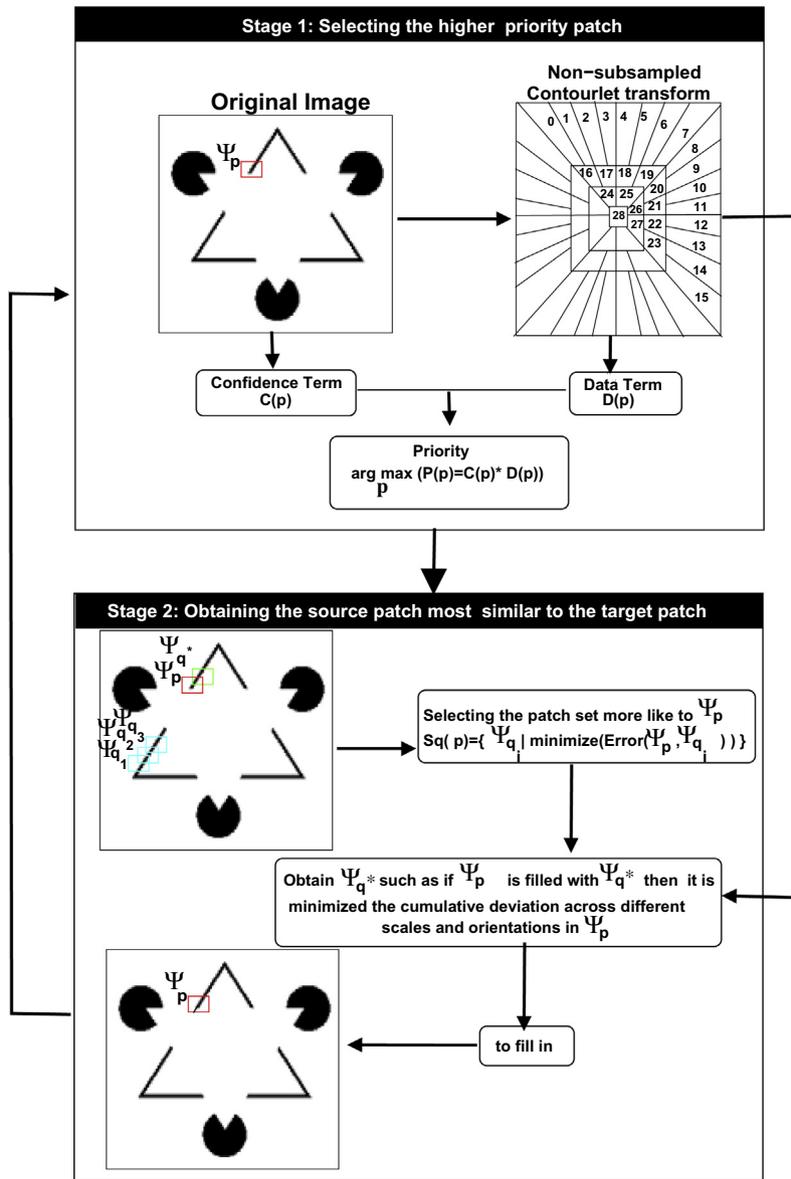
These representations also infer a congruence level for the information which represent, with the information being more congruent whenever it is present across different scales and orientations (Rodríguez-Sánchez et al., 1999).

With respect to the image inpainting process, this property is very interesting for two reasons:

- It allows to establish the order to fill-in the target region.
- Also, it can be used to decide what information from some source area is the candidate to replace undesired data that is present in the target region.



(A)



(B)

**Fig. 3.** (A) Elements of the Criminisi's algorithm. For an image  $I$  with target region  $\Omega$  and source region  $\Phi$ : the goal in each iteration is to find the target patch  $\Psi_p$  centred at the  $p$ -pixel with the highest priority which should be fill in using the source patch  $\Psi_q$  from  $\Phi$ . (B) Scheme for the proposed image inpainting algorithm.

A new multi-scale image decomposition called the contourlet transform (CT) was previously presented (Do and Vetterli, 2005). This transform combines the Laplacian pyramid (LP) (Do and

Vetterli, 2003) and the directional filter bank (DFB) (Bamberger and Smith, 1992). Compared to the discrete wavelet transform (DWT), the CT is a decomposition characterized by multi-direction

and anisotropy in addition to multi-scale and localization. Thus, the CT can represent edges and other singularities along curves much more efficiently (i.e., using fewer coefficients). However, the CT lacks the shift-invariance which is desirable in many image applications, such as image enhancement or image denoising.

The nonsubsampling contourlet transform (NSCT) was presented as an overcomplete transform (da Cunha et al., 2006), which has been successfully used in image denoising (da Cunha et al., 2006) and image enhancement (Zhou et al., 2005). This transform is characterized as a shift-invariant version of the contourlet transform (CT) (Do and Vetterli, 2005). The CT applies the Laplacian pyramid (LP) to obtain a multi-scale decomposition and a directional filter bank (DFB) (Bamberger and Smith, 1992). In the NSCT, any downsamplers and upsamplers during the decomposition and the reconstruction of the image are realized to avoid the frequency-aliasing of the CT and, at the same time, to achieve the shift-invariance. For this, the NSCT uses the nonsubsampling pyramids filter banks (NSPFB) and the nonsubsampling directional filter banks (NSDFB).

Fig. 1 shows a scheme for the NSCT. The same figure also illustrates the idealized frequency partitioning together with the response for each NSCT band.

The NSPFB is a two-channel nonsubsampling filter bank (NFB). To achieve the multi-scale decomposition, the filters for the next stage are obtained by upsampling the filters of the previous state with the sampling matrix.

$$D = 2I = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \quad (1)$$

which gives a multi-scale decomposition without need of additional filter design. In Fig. 2(A) it is shown the NSPFB decomposition with  $J = 3$  stages. The ideal support of the lowpass filter at the  $j$ -th stage is the region  $[-(\pi/2^j), (\pi/2^j)]^2$ . The complement of the lowpass defines the support of the high-pass filter  $[-(\pi/2^{j-1}), (\pi/2^{j-1})]^2 \setminus [-(\pi/2^j), (\pi/2^j)]^2$ . The filter in each level is calculated as:

$$H_n^*(z) = \begin{cases} H_1(z^{2^{n-1}}) \prod_{j=0}^{n-2} H_0(z^{2^j}), & 1 \leq n \leq J \\ \prod_{j=0}^{n-2} H_0(z^{2^j}), & n = J + 1 \end{cases} \quad (2)$$

where  $H_0(z)$  and  $H_1(z)$  are the lowpass filter and highpass filter at the first stage, respectively. Fig. 2(B) illustrates a four-channel directional decomposition. In the second level, the upsampled fan filters  $U_j(z^D)$   $j = 0, 1$  have checker-board frequency support and the sampling matrix  $D$  is the quincunx matrix:

$$D = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad (3)$$

When the filters  $U_j(z^D)$  are combined with the fan filters  $U_i(z)$  ( $i = 0, 1$ ) in the first level, we obtain the four-channel directional decomposition. The NSCT is obtained by combining the NSPFB and the NSDFB (see Fig. 1).

### 3. A new inpainting method

The proposed method is based on the Criminisi's algorithm (Criminisi et al., 2003). Thus it combines the advantages of "textures synthesis" algorithms to obtain a large image from sample textures and "inpainting" techniques for filling in small image gaps. The elements of the Criminisi's method are shown in Fig. 3(A).

The Criminisi's algorithm makes use of two different terms:

1. A data term  $D(p)$  which is a function of the strength of isophotes hitting the front  $\delta\Omega$ . This term is defined as:

$$D(p) = \frac{|\nabla I_p^{\perp} \cdot n_p|}{\alpha}$$

2. A confidence term  $C(p)$  which is a measure of the amount of reliable information surrounding the pixel  $p$  (which has to be filled-in)

$$C(p) = \frac{\sum_{q \in \Psi_p \cap \Omega} C(q)}{|\Psi_p|}$$

Using both terms, the priority of a point  $p$  is defined as:

$$P(p) = D(p) * C(p)$$

A scheme of this process is shown in Algorithm 1

---

#### Algorithm 1. Main steps of the Criminisi's algorithm

---

1. For each  $p \in \delta\Omega$ 
    - (a)  $C(p) = \frac{\sum_{q \in \Psi_p \cap (\Omega - \delta\Omega)} C(q)}{|\Psi_q|}$
    - (b)  $D(p) = \frac{|\nabla I_p^{\perp} \cdot n_p|}{\alpha}$
    - (c)  $P(p) = D(p) * C(p)$
  2.  $p = \arg \max_{q \in \delta\Omega} (P(q))$
  3. Define  $\Psi_p$  as the target patch centred at the  $p$ -pixel with the highest priority
  4.  $\Psi_q = \arg \min_{\Psi_q \in \Phi} SSD(\Psi_p, \Psi_q)$  with SSD being the sum of square differences.
  5. Copy the source patch  $\Psi_q$  onto the  $\Psi_p$  target patch.
  6. if  $\delta\Omega \neq \emptyset$  go to 1
- 

#### 3.1. Image Inpainting based on the contourlet transform.

In this section, the proposed algorithm is discussed in detail. This novel technique is based on:

1. Exemplar-based image inpainting which is aimed to fill in the target region following image texture (Criminisi et al., 2003; Efros and Freeman, 2001; Efros and Leung, 1999; Liang et al., 2001).

A multiresolution analysis to achieve an approximation for exemplar-based image inpainting.

The main stages of the algorithm are shown in Fig. 3(B):

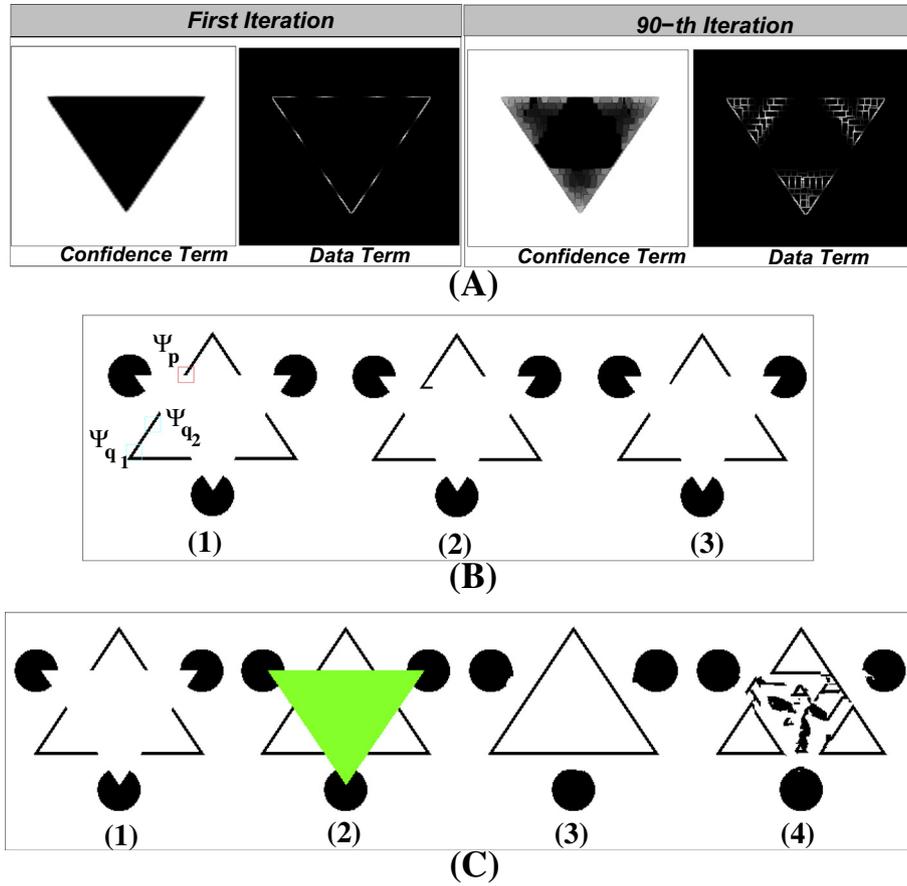
1. Selecting the higher priority patch with centre at  $p$ .
2. Obtaining the source patch to be copied onto the target patch.

#### 3.2. Selecting the higher priority patch with centre at $p$

At this stage the algorithm needs to select the target patch  $\Psi_p$ , with the point centre  $p$  to be filled in. To this aim, the confidence term is defined as given in the Criminisi's algorithm. However, the data term is defined as the strength of the isophotes hitting the front  $\delta\Omega$  across different scales and orientations of the contourlet domain. Thus, we first calculate a  $J$ -level NSCT for the original image by obtaining the corresponding coefficients at the patch  $p$ .

$$\{C_{j_0}(p), C_{j,l}(p)\} \quad (j_0 \geq j \geq 1, l = 1, 2, \dots, 2^j)$$

where  $lj$  denotes the number of levels in the NSDFB at the  $j$ -th scale.  $C_{j_0}$  represents the lowpass subband coefficients at the coarse scale and,  $C_{j,l}$  denotes the bandpass directional subband coefficients at the  $j$ -th scale and the  $l$ -th orientation. In the following we will note



**Fig. 4.** (A) The confidence and data term for the first and the 90-th iteration of the algorithm. (B) Problems with SSD: (1) original image and target patch  $\Psi_p$  in red colour and two source patches  $\Psi_{q_1}$  and  $\Psi_{q_2}$  having the same value of SSD compared with  $\Psi_p$ ; (2) output when the chosen source patch is  $\Psi_{q_1}$ ; and (3) output when the chosen source patch is  $\Psi_{q_2}$ . (C) Why to minimize the dispersion? (1) original image; (2) original image with target region in green; (3) results obtained following the Eq. (10); and (4) results obtained when it is maximized the dispersion accumulated across different orientations and scales. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

$C_{j,l}$  as  $C_{l(j)}$ . Next, we expose the different steps to obtain the data term:

1. *Corner points for each scale at different orientations.* In this step for each pixel  $(i, j) \in \delta\Omega$  we calculate:

$$\{C_{k(s),t(s)}(i, j) = |C_{s,k}(i, j)| \times |C_{s,t}(i, j)|\} \quad \forall k(s), t(s), \quad 1 \leq k(s) < t(s) \leq O(s) \quad (4)$$

where  $O(s)$  is the number of orientations at scale  $s$ . A corner point (at a particular scale) is produced when two bands in different orientations meet them in a particular spatial location. Therefore, a high value of  $C_{k(s),t(s)}(i, j)$  means that in the  $(i, j)$  pixel there is a concurrence of high energy at orientations  $k(s)$  and  $t(s)$ .

2. *Corner points across different scales.* This step searches for the corner points with different orientations which can be present at different scales. From the pyramidal structure of the nonsub-sampled contourlet transform, it follows that each coefficient at a given scale is associated with the same spatial location for certain coefficients at a higher scale. Thus Eq. (5) is formulated as:

$$P_{(o_1(s), o_2(s)), (o_1(s+1), o_2(s+1)), \dots, (o_1(J), o_2(J))}^S(i, j) = \prod_{n=s}^J C_{o_1(n), o_2(n)}(i, j) \quad (5)$$

3. *Data term.* In order to derive the data term, we accumulate at each point the values obtained in Eq. (5) for different scales (parameter  $s$ ). To compress the dynamic range of the values we apply the natural logarithm function.

$$D(i, j) = \log \left( \sum_{s=1}^J P_{(o_1(s), o_2(s)), (o_1(s+1), o_2(s+1)), \dots, (o_1(J), o_2(J))}^S(i, j) \right) \quad (6)$$

The data term should be normalized as given in Eq. (7) in order to achieve a trade-off between confidence and data terms, which is needed to correctly calculate the level of priority for a point in the target region.

$$D(i, j) = \frac{D(i, j)}{\max_{(m, n) \in \delta\Omega} (D(m, n))} \quad (7)$$

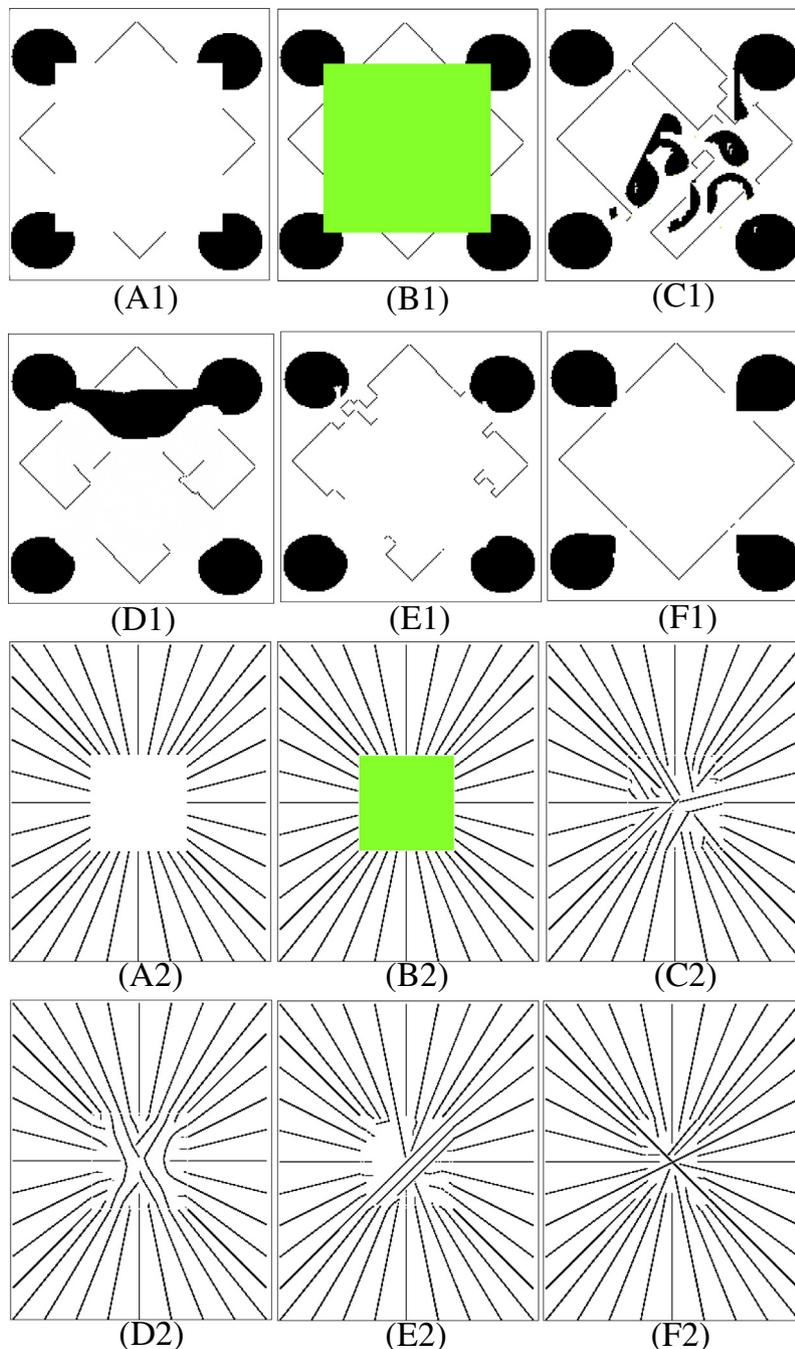
The priority of point  $p$  can be calculated following the same definition used in the Criminisi's algorithm:

$$P(p) = D(p) * C(p)$$

In Fig. 4(A), the image data and confidence at the first and at the 90-th iteration respectively are shown. A black pixel in the image implies the lowest confidence value (which is the initial value for the target region) and a white pixel has a high confidence value. On the other hand, with respect to the data term, the highest values correspond to points which have high energy across different scales and orientations.

### 3.3. Obtaining the source patch to be copied onto the target patch

After finding the point  $p$  with the highest priority to be fill-in, our algorithm searches the source image to find the patch that is



**Fig. 5.** Kanizsa square: (A1) original image; (B1) original image with target region in green colour; (C1) Criminisi's result; (D1) Wexler's result; (E1) Goyal's result; and (F1) our method. *Different orientations*: (A2) original image; (B2) original image with target region in green colour; (C2) Criminisi's result; (D2) Wexler's result; (E2) Goyal's result; and (F2) our method. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

most similar to the target patch. The sum of squared differences (SSD) is used to calculate the distance between the target patch  $\Psi_p$  and the source patch  $\Psi_q$ . The most similar patch  $\Psi_q$  satisfies:

$$\Psi_q = \arg \min_{\Psi_q \in \Phi} d(\Psi_p, \Psi_q) \quad (8)$$

When two patches  $\Psi_{q_1}$  and  $\Psi_{q_2}$  have the same SSD value compared with  $\Psi_p$  (see Fig. 4(B)), our algorithm selects the patch which minimizes the accumulated dispersion across different orientations and scales.

Following the Criminisi's algorithm, our method selects a set of the source patches  $S_q(p)$  which are most similar to the target patch

according to their SSD values. Thus, the steps to obtain the best source patch are:

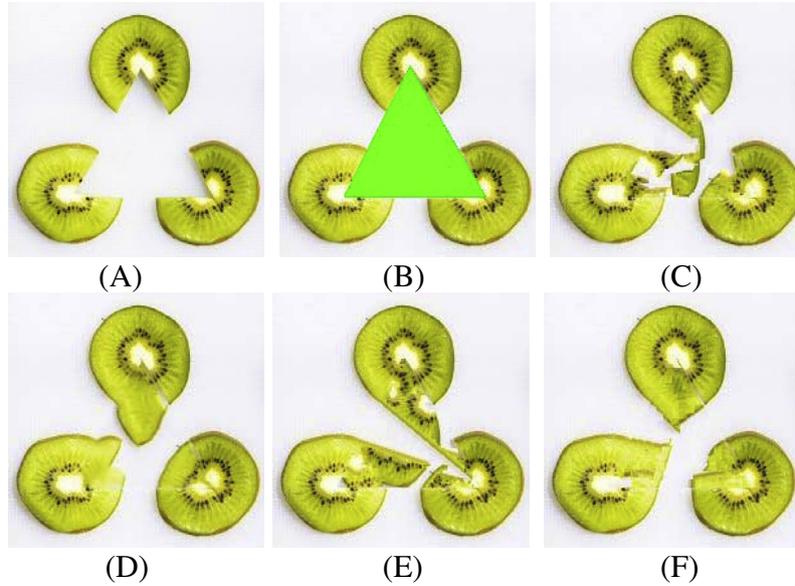
1. Obtain the source patch candidate set. We obtained  $S_q(p)$  as:

$$S_q(p) = \{\Psi_q | SSD(\Psi_p, \Psi_q) < \epsilon\} \quad (9)$$

where  $\epsilon$  is a small constant value. Thus,  $S_q(p)$  is composed of the source patches most similar to the target patch.

2. Select the best source patch from  $S_q(p)$ . To obtain the best source patch from  $S_q(p)$  we define

$$\{\sigma_{j_0}(p), \sigma_{j_l}(p)\} \quad j_0 \geq j \geq 1, l = 1, 2, \dots, 2^l$$



**Fig. 6.** Broken object: Illusory triangle using a kiwi. (A) Original image; (B) original image with target region in green colour; (C) Criminisi's result; (D) Wexler's result; (E) Goyal's result; and (F) our method. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

as the standard deviation associated to the coefficients in each band of the contourlet transform for  $\Psi_p$ .

$$\{\sigma_{j_0}(q \leftarrow p), \sigma_{j,l}(q \leftarrow p)\} \quad j_0 \geq j \geq 1, l = 1, 2, \dots, 2^j$$

is the standard deviation of the contourlet coefficients within each band, when the source patch  $\Psi_q$  is copied onto the target patch  $\Psi_p$  to fill in the empty area. Then, the best source patch,  $\Psi_{q^*}$  is obtained as:

$$\Psi_{q^*} = \arg \max_{\Psi_q \in S_q(p)} \left( e^{-\sigma_{j_0}(q \leftarrow p)} + \sum_{j,l} e^{-\sigma_{j,l}(q \leftarrow p)} \right) \quad (10)$$

With Eq. (10) we find the source patch that produces the lowest accumulated dispersion (when is copied in the target patch), across different orientations and scales in the NSCT domain. This property asserts that the selected source patch continues the broken features in the original image. On the other hand, the method would choose a source patch making new features appear in the image (i.e., new edges, or artifacts). From the point of view of signal analysis, these new features, drive to new transients characterized by high energy in some coefficients compared with the energy values of the neighbouring coefficients, driving to an higher dispersion. Thus, selecting the source patch with the lowest accumulated dispersion across different scales and orientations in the target patch, we do not generate new transients, keeping the transients of the original image.

An example illustrating the relationship between the minimization of the accumulated dispersion and the continuity of the broken features in comparison with the relationship between the maximization of the accumulated dispersion and the generation of new features, is shown in Fig. 4(C). For image 4(C1) and (C2) (with the target region in green), we have obtained 4(C3) (following the Eq. (10) minimizing the dispersion accumulated across different orientations and scales). On the other hand, in Fig. 4(C4) is shown the result obtained when the dispersion accumulated across different orientations and scales is maximized. In this last image there is a tendency to build new structures breaking the continuity of the original structures in the image.

The fundamental steps of the proposed methodology are given in Algorithm 2.

---

#### Algorithm 2. Fundamental steps of our algorithm

---

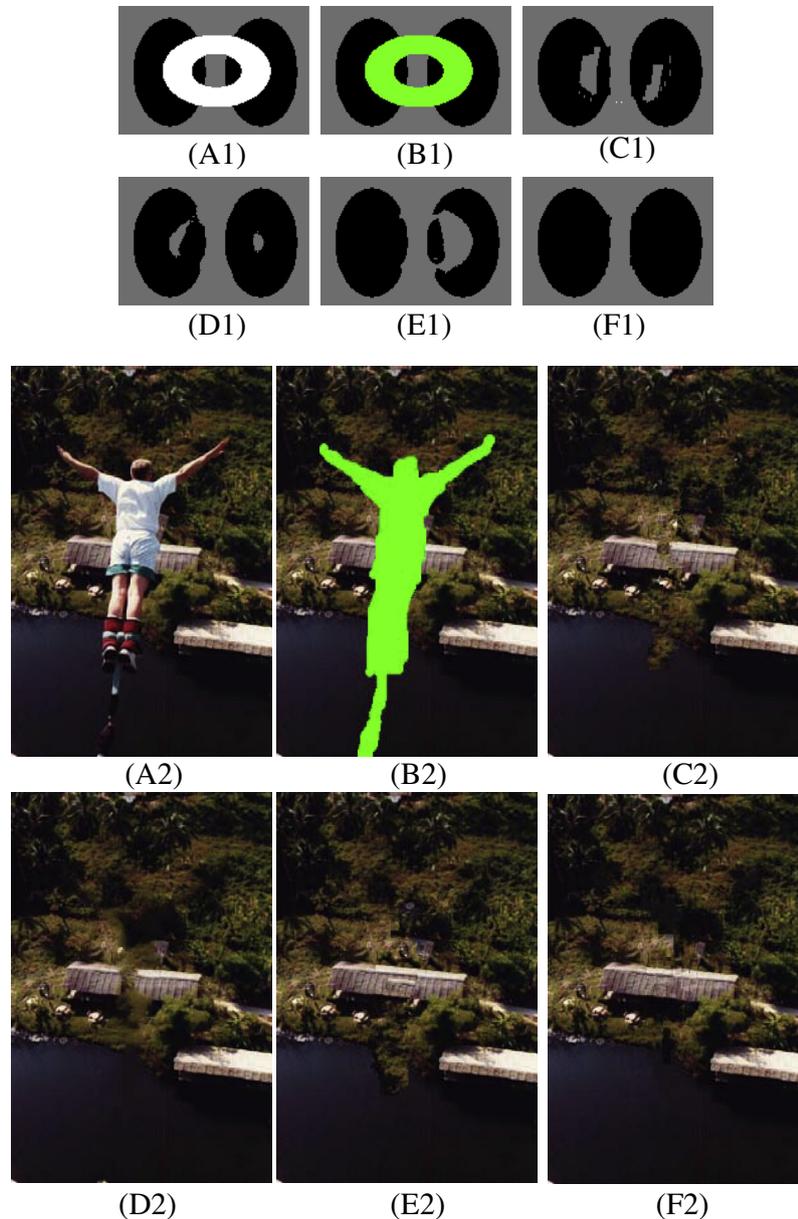
1. For each  $p \in \delta\Omega$ 
    - (a)  $C(p) = \frac{\sum_{q \in \Psi_p \cap (I-\Omega)} C(q)}{|\Psi_q|}$
    - (b)  $\tilde{D}(p) = \log \left( \sum_{s=1}^J P_{(o_1(s), o_2(s)), (o_1(s+1), o_2(s+1)), \dots, (o_1(J), o_2(J))}^s(p) \right)$
    - (c)  $D(p) = \frac{D(p)}{\max_{q \in \delta\Omega} (D(q))}$
    - (d)  $P(p) = D(p) * C(p)$
  2.  $p = \arg \max_{q \in \delta\Omega} (P(q))$
  3. Define  $\Psi_p$  as the target patch with highest priority
  4.  $S_q(p) = \{\Psi_q | SSD(\Psi_p, \Psi_q) < \epsilon\}$
  5.  $\Psi_{q^*} = \arg \max_{\Psi_q \in S_q(p)} \left( e^{-\sigma_{j_0}(q \leftarrow p)} + \sum_{j,l} e^{-\sigma_{j,l}(q \leftarrow p)} \right)$
  6. Copy the  $\Psi_{q^*}$  source patch onto the target patch  $\Psi_p$ .
  7. if  $\delta\Omega \neq \emptyset$  go to 1
- 

## 4. Results

### 4.1. Details of the implementation

The algorithm proposed has several parameters including number of scales, filters used in the NSCT decomposition, the threshold to define the  $S_q(p)$  cardinal and the dimension of the patch. For the examples illustrated in this section, the parameters were chosen as follows:

1. The filter used to implement the NSPFB and NSDFB is the 9/7 tap (Antonini et al., 1992).
2. The number of scales of the NSCT transform is 3. The first level has ( $j = 1$ ) four directional bands, the second level has eight directional bands and the third level with 16 directional bands together with the lowpass band.
3. The  $S_q(p)$  cardinal has a minimum value of 10. This number can increase when the 10th point in  $S_q(p)$  has a SSD-value equal to that of other points in  $S_q(p)$ . In this case, all points with an SSD-value equal to the SSD-value of the 10th point in  $S_q(p)$  are also included.



**Fig. 7.** Removing object 1: torus. (A1) Original image; (B1) original image with target region in green colour; (C1) Criminisi's result; (D1) Wexler's result; (E1) Goyal's result; and (F1) our method. Removing object 2: bungee. (A2) Original image; (B2) original image with target region in green colour; (C2) Criminisi's result; (D2) Wexler's result; (E2) Goyal's result; and (F2) result of our method. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. The dimension of the patch is  $13 \times 13$  pixels.

The software can be downloaded from [http://nautilus.ugr.es/research/inpainting\\_contourlet/](http://nautilus.ugr.es/research/inpainting_contourlet/).

Also in this web site are shown the results of the algorithm on a lot of examples including those of this paper. Finally, it is also possible to analyse the step by step execution of the algorithm for each image.

#### 4.2. Comparison

The method has been compared with three models, the Criminisi's algorithm (Criminisi et al., 2003), the Wexler's algorithm (Wexler et al., 2007) and the Goyal's algorithm (Goyal et al., 2010). The parameter for the Criminisi's algorithm is the dimension of the patch that is fixed to  $13 \times 13$  pixels. The parameters

of the Wexler's algorithm are number of levels, neighbourhood radius and number of rounds. By default the values are:

1. Number of levels is 4
2. Neighbourhood radius is 3
3. Number of rounds is 5

These values were applied to all examples in the following with the exception of images 5(A1) where the neighbourhood radius was 2 and 7(A1) where the number of levels was set to 3.

The software used to execute the different models has been obtained from:

- <http://www.cc.gatech.edu/sooraj/inpainting/>: the Criminisi's algorithm.
- <http://www.wisdom.weizmann.ac.il/vision/VideoCompletion.html>: the Wexler's algorithm.

**Table 1**  
Efficiency computational. Times CPU (s) by algorithm proposed.

Image	CPU time (s)
Fig. 6(A) Kanizsa triangle by kiwis	321
Fig. 7(A1) torus	360
Fig. 7(A2) bungee	630

- <http://pulkitgoyal.wordpress.com/2011/01/09/image-inpainting/>: the Goyal's algorithm.

Two groups of images have been used, a first group that contains broken edges, and a second group that contains a large object in the image that must be eliminated.

#### 4.3. Broken edges

A valid method of inpainting should be able to reconstruct an incomplete 2-D image in every one of its features. In this form, the image inpainting looks real when no obvious artefacts that cause it to look like an artificial image are present. One of the most important features of an image are the edges which are present at different scales and orientations. Several examples of broken edges are shown in Fig. 5(A1) and (A2).

In these images, only two levels of grey are used, black and white. For them, according to 'The Connectivity Principle' (Kanizsa, 1979) of the human disocclusion process (Chan and Shen, 2001), humans mostly seem to prefer the connected result.

In Fig. 5(A1), the results for the Criminisi's, Wexler's and Goyal's methods are shown in images (C1), (D1) and (E1) respectively. The result obtained by our algorithm is shown in (F1). Our method has maintained the image structure.

Another interesting image is shown in Fig. 5(A2), where broken edges need to be connected in different orientations. The outputs of the Criminisi's, Wexler's and Goyal's methods are shown in (C2), (D2) and (E2) respectively. The output of our method is shown in (F2). The result of our method has brought more lines closer to the centre of image than the other methods, which has kept a strong perception of the white square.

In Fig. 6(A), three pieces of a kiwi were broken to form an illusory triangle. The objective of this image was to fill in each piece of kiwi in some realistic form. The results obtained for the different methods can be seen in Fig. 6(C)–(F). The outputs more "realistic" are those obtained using the Wexler's method as well as our method.

#### 4.4. Removing an object from a scene

Fig. 7(A1) and (A2) show examples of object removing. In these examples, there is an object occluding others behind it. A synthetic image of this kind is shown in Fig. 7(A1), in which we have three levels of grey, that is, black, medium grey and white.

The objective is to eliminate the geometric form like a torus, in white, by filling in the hole to obtain two black ellipses. The output of the Criminisi's algorithm is shown in Fig. 7(C1).

A result similar to the output of the Criminisi's method was obtained by the Wexler's method (see Fig. 7(D1)). On the other hand, the Goyal's method fails to fill in the ellipse on the right. In Fig. 7(F1) the result obtained by our method can be seen. The Fig. 7(A2) is other example of real scenes. The results obtained by the different methods can be observed in Fig. 7(C2)–(F2) respectively.

#### 4.5. Computational efficiency

The software has been developed under Matlab and tested in the 2008a and 2011 versions. We show in Table 1 the CPU times

(CPU i7-860 2.80 GHz) for different images. The computational efficiency of the algorithm depends on two factors:

1. Area (percentage) of the target region.
2. Cardinal of source patch candidate set  $S_q(p)$ . In order to obtain a result coherent for the different images, this number must adopt a enough high value.

## 5. Conclusions

Here we propose an image inpainting algorithm using the non-subsampled contourlet transform (NSCT). The proposed approach is based on a new mechanism for the selection of the target patch as well as the selection of the source patch in each iteration.

In the selection of the target patch, which is to used to fill-in the target region, two terms are defined: a confidence term and a data term. While the confidence term is a measure of the amount of reliable information surrounding a pixel, the data term is the strength of isophotes hitting the boundary of the target region.

Here we proposed a new definition of the data term based on the redundancy, across orientations and scales in the NSCT domain, of the corner pixels located at the boundary of the target region.

It follows that a pixel with a high value for the data term means that: (i) There is a corner point at a particular scale, since the high energy level is achieved when two bands of equal scale at different orientations meet at this particular spatial location; and (ii) this same corner point must be present across different scales.

The interesting point is that pixels in the boundary of the target region which are part of broken or occluded features (e.g., broken edges) have a high value of the data term, and thus, they are identified as corner points.

Concerning the source-patch selection, we have shown that by only using the Square Sum Differences between target and source patch, it can produce results with artefacts.

To overcome this problem, a new mechanism of source patch selection was proposed. Firstly we obtain the set of the most similar source patches to the target patch by using the Square Sum Differences. Next, from this set, we select the patch achieving the minimum dispersion across orientations and scales when is copied onto the target patch, which maintains the transients of the original image while avoiding new ones. Hence, the selected source patch continues the broken features in the original image while avoiding the formation of artefacts.

## Acknowledgement

This research was sponsored by the Spanish Board for Science and Technology (MICINN) under grant TIN2010-15157 cofinanced with European FEDER funds.

## References

- Antonini, M., Barlaud, M., Mathieu, P., Daubechies, I., 1992. Image coding using wavelet transform. *IEEE Transactions on Image Processing* 1 (2), 205–220.
- Arias, P., Caselles, V., Facciolo, G., et al., 2012. Nonlocal variational models for inpainting and interpolation. *Mathematical Models & Methods in Applied Sciences* 22 (2).
- Bamberger, R.H., Smith, M.J.T., 1992. A filter bank for the directional decomposition of images: theory and design. *IEEE Transaction on Signal Processing* 40 (4), 882–893.
- Bertalmio, M., Sapiro, G., Caselles, V., Ballester, C., 2000. Image inpainting. In: *Proceedings of ACM Conference on Computer Graphics (SIGGRAPH'00)*, pp. 417–424.
- Chan, T.F., Shen, J., 2001. Non-texture inpainting by curvature driven diffusions. *Journal of Visual Communication and Image Representation* 12 (4), 436–449.
- Chen, Q., Zhang, Y., Liu, Y., 2007. Image inpainting with improved exemplar based approach. *Lecture Notes in Computer Science*, vol. 4577, pp. 242–251.
- Criminisi, A., Pérez, P., Toyama, K., 2003. Object removal by exemplar-based inpainting. In: *Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and, Pattern Recognition (CVPR'03)*, pp. 1–8.

- da Cunha, A.L., Zhou, J.P., Do, M.N., 2006. The nonsubsampling contourlet transform: filter design and application in denoising. *IEEE Transactions on Image Processing* 15 (10), 3089–3101.
- Daubechies, I., 1992. Ten lectures on wavelets. In: CBMS-NSF, Regional Conference Series in Applied Mathematics, SIAM: Society for Industrial and Applied Mathematics, vol. 61. 357p.
- Daugman, J.G., 1988. Complete discrete 2-D Gabor transform by neural networks for image analysis and compression. *IEEE Transactions on Acoustic, Speech, and Signal Processing* 36 (7), 1169–1179.
- De Bonet, J.S., 1997. Multiresolution sampling procedure for analysis and synthesis of texture images. In: Proceedings of the 24th Annual Conference on Computer Graphics and Interactive, Techniques (SIGGRAPH'97), pp. 361–368.
- Do, M.N., Vetterli, M., 2003. Framing pyramids. *IEEE Transaction on Signal Processing* 51 (9), 2329–2342.
- Do, M.N., Vetterli, M., 2005. The contourlet transform: an efficient directional multiresolution image representation. *IEEE Transactions on Image Processing* 4 (12), 2091–2106.
- Efros, A., Freeman, W.F., 2001. Image quilting for texture synthesis and transfer. In: Proceedings of ACM Conference on Computer Graphics (SIGGRAPH'01), pp. 341–346.
- Efros, A., Leung, T., 1999. Textures synthesis by nonparametric sampling. In: Proceedings of the International Conference on Computer Vision, Kerkyra, Greece, pp. 1033–1038.
- Goyal, P., Diwakar, S., Agrawal, A., 2010. Fast and enhanced algorithm for exemplar based image inpainting. In: Proceedings of the IEEE Fourth Pacific-Rim Symposium on Image and Video Technology, pp. 325–330.
- Hervieu, A., Papadakis, N., Bugeau, A., et al., 2011. Stereoscopic image inpainting using scene geometry. In: International Conference on Multimedia and Expo (ICME), pp. 1–6.
- Kahge, S., Luis, R., Cornolli, M.H., 2012. Removing high contrast artifacts via digital inpainting in cryo-electron tomography: an application of compressed sensing. *Journal of Structural Biology* 178 (2), 108–120.
- Kanizsa, G., 1979. *Organization in Vision*. Holt Rinehart and Winston, New York.
- Liang, L., Liu, C., Xu, Y.Q., Guo, B.N., Shum, H.Y., 2001. Real-time texture synthesis by patch-based sampling. *ACM Transactions on Graphics (TOG)* 20 (3), 127–150.
- Li, Yan-Ran, Shen, Lixin, Suter, Bruce W., 2013. Adaptive inpainting algorithm based on DCT induced wavelet regularization. *IEEE Transactions on Image Processing* 22 (2), 752–763.
- Rodríguez-Sánchez, R., Garcia, J.A., Fdez-Valdivia, J., Fdez-Vidal, Xose R., 1999. The RGFF representational model: a system for the automatically learned partitioning of “visual patterns” in digital images. *IEEE Transactions on PAMI* 21 (10), 1044–1073.
- Simoncelli, E.P., Portilla, J., 1998. Texture characterization via joint statistics of wavelet coefficient magnitudes. Proceedings of the IEEE Fifth International Conference on Image Processing 1, 62–66.
- Wexler, Y., Shechtman, E., Irani, M., 2007. Space-time completion of video. *IEEE Transaction on Pattern Analysis and Machine Intelligence* 29 (3), 463–476.
- Wu, J., Ruan, Q., 2006. Object removal by cross isophotes exemplar-based inpainting. In: Proceedings of the 18th International Conference on, Pattern Recognition (ICPR'06), pp. 810–813.
- Yamauchi, H., Haber, J., Seidel, H.-P., 2003. Image restoration using multiresolution texture synthesis and image inpainting. In: Proceedings of the Computer Graphics International (CGI'03), pp. 120–125.
- Zhou, J.P., da Cunha, A.L., Do, M.N., 2005. Nonsubsampling contourlet transform: filter design and application in enhancement. In: Proceedings of the IEEE International Conference on Image Processing (ICIP'05), pp. 469–476.