Defect Detection on Inclined Textured Planes Using the Shape from Texture Method and the Delaunay Triangulation

Justin Plantier  
Cognitive Science Department, IMASSA, BP 73, 91223 Brétigny-sur-Orge, France  
Email: jplantier@imassa.fr

Laurent Boutté  
CEMIF-LSC, Université d'Evry, 40 rue du Pelvoux, CE 1455 Courcouronnes, 91020 EVRY Cedex, France  
Email: lboutte@inforie.fr

Sylvie Lelandais  
CEMIF-LSC, Université d'Evry, 40 rue du Pelvoux, CE 1455 Courcouronnes, 91020 EVRY Cedex, France  
Email: s.lelandais@iut.univ-evry.fr

Received 20 July 2001 and in revised form 13 March 2002

We present one method for detecting defects on an inclined textured plane. This method uses a combination of a shape from texture (SFT) method with the Delaunay triangulation technique. The SFT method provides the theoretical equation of the plane orientation in two steps. First, a wavelet decomposition allows us to build an image of the inverse of the local frequency, that is the scale, that we call the local scales map. Then we perform an interpolation of this map using the equation of the theoretical variation of the scales. With the interpolation parameters it is possible to extract the texels by the use of an adaptive thresholding for each pixel of this map. Then we compute the centers of each texel in order to match a mesh on it after processing a Delaunay triangulation. When there is a defect, the regularity of the triangulation is disturbed, so one hole appears in the mesh.

Keywords and phrases: defect detection, inclined textured planes, wavelet decomposition, local scales, texels extraction, shape from texture method, Delaunay triangulation.

1. INTRODUCTION

It is not easy to detect defects on an inclined plane because the relations between the different points are modified according to the orientation of the plane. In this paper, we present one way to detect defects on an inclined plane which is covered by a regular macrotexture. Two steps are necessary for this method. First, we compute the local scale of each pixel of the image by a wavelet decomposition. An interpolation of the local scales map gives the equation of the plane orientation. Then we threshold the local scales map in order to extract the texels of the original image. With the previous equation it is possible to obtain a threshold value for each pixel. After that, we compute the gravity centers of the texels and map a mesh on it using a Delaunay triangulation. The triangulation is regular when there is no defect. When there is a defect, the regularity of the triangulation is disturbed and it is not possible to build the mesh. Then, to look for defects is the same as looking for holes in the mesh. Some results of this fully automatic method, obtained on synthetic and real textures, will be shown.

Much work has been performed with the aim to analyze textures [1]. Since Haralick’s work with the co-occurrence matrix which allows classification of microtextures [2], several authors have proposed methods for texture segmentation [1, 3, 4, 5]. In our application, we want to detect defects on inclined textured planes. Previous techniques of texture analysis and segmentation are not useful because the orientation of the plane and the defects modify the texture parameters. We have to find a method which takes the variations tied to the orientation into account, so we use a shape from texture (SFT) method followed by a Delaunay triangulation in order to perform defect detection. This technique allows us to find defects on regular macroscopic texture (Figure 1).

Defect detection on inclined textured planes poses a difficult problem because of the texture distortion induced by its projection on the plane image. In this article, a perspective projection model is used. This hypothesis takes the three
660 EURASIP Journal on Applied Signal Processing

In the following, first we describe our method to recover the equation of the plane. This method uses a wavelet decomposition, a computation of local scales for each pixel, and an interpolation of the local scales map. To perform the interpolation, we use an improvement of a method developed by Lu et al. [12]. Then we explain how it is possible to build a mesh which will be used as a model of the texture. Finally, we present the performance of our method on different kinds of defects.

2. EQUATIONS OF THE INCLINED PLANE

The texture can be characterized by its spatial frequencies. This description allows us to find the orientation of the planar surface from the variation of the local frequencies in the image. The local frequency of a point can be defined as the most important frequency around this point. This feature is obtained from a wavelet decomposition of the image. In fact, the convolution of the image and a wavelet gives the spatial representation of the frequency filtered by the wavelet (the value given by the convolution is greater when the frequency around a point is near the frequency of the wavelet). So we can split all the frequencies present in the image with a set of wavelets which covers the whole spectrum. In our study, the images will be decomposed with DOG (difference of Gaussians) functions.

In the following, first we describe our method to recover the equation of the plane. This method uses a wavelet decomposition, a computation of local scales for each pixel, and an interpolation of the local scales map. To perform the interpolation, we use an improvement of a method developed by Lu et al. [12]. Then we explain how it is possible to build a mesh which will be used as a model of the texture. Finally, we present the performance of our method on different kinds of defects.

The texture can be characterized by its spatial frequencies. This description allows us to find the orientation of the planar surface from the variation of the local frequencies in the image. The local frequency of a point can be defined as the most important frequency around this point. This feature is obtained from a wavelet decomposition of the image. In fact, the convolution of the image and a wavelet gives the spatial representation of the frequency filtered by the wavelet (the value given by the convolution is greater when the frequency around a point is near the frequency of the wavelet). So we can split all the frequencies present in the image with a set of wavelets which covers the whole spectrum. In our study, the images will be decomposed with DOG (difference of Gaussians) functions.

So we can define the local frequency of a pixel as the frequency of the wavelet that yields the maximum value at this point [15]. If we use this definition, a problem called step effect appears in the resulting image. In fact the number of different frequencies in the resultant image is less than or equal to the number of wavelets used for the decomposition. To remove this phenomenon, we must find the real value of the local frequency from the sampling given by the decomposition. So we must perform, for each point of the image, an interpolation of the discrete curve given by the wavelets, giving the frequencies around the studied point. With this method, we obtain an image which is the map of the local scale for each pixel [16], the local scale being defined as the inverse of the local frequency. Figure 3 shows the local scales map (Figure 3b), which we call $E_L(x, y)$, obtained from an original image (Figure 3a).

The computation of the orientation of a textured tilt plane from an image requires us to know theoretically the frequency variation. Consequently, the opposite problem of our study must be analyzed, that is, the local spatial frequency variation must be known for a given set of incline parameters. With this knowledge, the inclination parameters of a plane can be computed from the variation of the local spatial frequency in the image [12].

To recover the orientation of an inclined textured plane from one image, we must study the deformation of the texture caused by the projection. It is possible to describe the incline of a plane in the 3D space by the use of a normal vector of this plane, with two rotations. We have chosen a representation by the tilt and the slant angles as shown on Figure 4. The tilt angle ($\tau$) is the angle between the projection

Figure 1: Two examples of regular macroscopic textures.

Figure 2: The three cues of deformation.

These three cues, shown on Figure 2, will allow us to recover the orientation of a surface. Before looking for defects, it is necessary to obtain the theoretical equation of the inclined textured plane. Actually, this equation gives us a threshold value for extracting the texels. This is very close to computing the orientation of the plane. The problem of computing the orientation of a planar surface can be solved in several different ways [1, 7, 8, 9, 10]. These methods can be classified according to the description of the texture of the plane and the way of using this description to recover the orientation of the surface. In this work, we choose to describe the texture by local frequencies [11, 12, 13, 14]. The advantage of this method is to have a very large quantity of usable data to recover the orientation of the plane.
of the normal vector \((Z'_p)\) of the textured plane \((\pi)\) and the horizontal axis \((X)\). This angle gives the direction of the incline of the plane \((\pi)\). The slant angle \((\theta)\) is the angle between the normal vector to the textured plane \((Z')\) and the optical axis \((Z)\). This angle gives the importance of the incline of the plane \((\pi)\). In order to find the tilt and the slant angles, we define three reference marks:

- the first is the object reference mark; the people who calibrate the camera choose this reference mark. It is made up of the plane \((\pi)\) and the axis \((Z')\) (Figure 4);
- the second is the camera reference mark; this reference mark is fixed with respect to the camera. Its center is at the intersection of the optical axis and the lens of the camera. One of its axes is the optical axis and the two others are parallel to the sensor matrix. It is the \(XYZ\) space of Figure 4;
- the third is the image plane reference mark. It is a bidimensional reference mark where point coordinates are given in pixels. The origin of this reference mark is the optical center of the camera. Its two axes are parallel with the \(X\) and \(Y\) axes of the previous reference mark.

The matrix \(\Omega\) is the rotation matrix which allows us to transform a point from the object reference mark in a point in the camera reference mark, on the condition that the two reference mark centers are the same

\[
\Omega = \begin{pmatrix}
\Omega_{11} & \Omega_{12} & \Omega_{13} \\
\Omega_{21} & \Omega_{22} & \Omega_{23} \\
\Omega_{31} & \Omega_{32} & \Omega_{33}
\end{pmatrix},
\]

where \(\Omega_{11} = \sin^2 \tau + \cos^2 \tau \cos \theta, \Omega_{12} = \sin \tau \cos \theta (\cos \theta - 1), \Omega_{13} = \cos \tau \sin \theta, \Omega_{21} = \sin \tau \cos \theta (\cos \theta - 1), \Omega_{22} = \cos^2 \tau + \sin^2 \tau \cos \theta, \Omega_{23} = \sin \tau \sin \theta, \Omega_{31} = - \cos \tau \sin \theta, \Omega_{32} = - \sin \tau \sin \theta,\) and \(\Omega_{33} = \cos \theta\).

A theoretical study using this representation by the tilt and slant angles allows us to obtain the equation that links the orientation with the local frequencies. To find these equations [16], we begin with a simple case, that is, an orientation where the tilt angle is equal to zero. Let \(P_\pi = (x_\pi, y_\pi, 0)^t\) be one point of the texture plane. The point \(P_c = (x_c, y_c, z_c)^t\) is the same point in the camera mark and \(P = (x, y)\) is the projection of \(P_c\) in the image mark. It is possible to write

\[
P_c = \Omega \cdot P_\pi = \begin{pmatrix}
x_\pi \\
y_\pi \\
-z_\pi \sin \theta
\end{pmatrix},
\]

where \(\Omega = \begin{pmatrix}
x_\pi \\
y_\pi \\
(z_\pi \sin \theta)
\end{pmatrix}.
\]

\[
P = \begin{pmatrix}
x_\pi \cos \theta \\
y_\pi \\
(f + Z) + x_\pi \sin \theta \cdot f
\end{pmatrix},
\]

where \(f\) is the focal length of the camera and \(Z\) is the distance between the center of the object mark and the optical center of the camera. With (2b), we obtain

\[
x = \frac{x_\pi \cos \theta}{(f + Z) + x_\pi \sin \theta \cdot f}.
\]

So

\[
x_\pi = \frac{(f + Z)}{f} \cdot \frac{1}{\cos \theta} \cdot \frac{x}{1 - (x/f) \tan \theta}.
\]

Let \(S(x_\pi, y_\pi)\) be a perfect sinusoidal texture in the plane \(\pi\). With the perspective projection hypothesis, the equation of the texture in the image coordinates in the plane \(XY\) becomes

\[
S(x, y) = \cos (y(x_\pi)),
\]

where

\[
y(x_\pi) = \omega \cdot x_\pi = \omega \left[ \frac{(f + Z)}{f} \cdot \frac{1}{\cos \theta} \cdot \frac{x}{1 - (x/f) \tan \theta} \right].
\]

The local frequency is defined by the following equation:

\[
F(x, y) = \frac{1}{2\pi} \frac{dy(x_\pi)}{dx} = k_1 \cdot \frac{x}{(f, Z)} \cdot \frac{1}{\cos \theta} \cdot \frac{1}{(1 - (x/f) \tan \theta)^2}.
\]

It is possible to change this expression to have the local scale...
in each point of the image. The local scales will be privileged in this study because the interpolation of a second degree polynomial is easier than the interpolation of a hyperbola

\[
E(x, y) = \frac{1}{F(x, y)} = \beta_2 x^2 + \beta_1 x + \beta_0 \tag{8}
\]

with

\[
\begin{align*}
\beta_0 & = \frac{2\pi}{\omega} \left( \frac{f}{f + z} \right) \cos \theta, \tag{9a} \\
\beta_1 & = \beta_0 \left( -\frac{2}{f} \right) \tan \theta, \tag{9b} \\
\beta_2 & = \beta_0 \left( \frac{1}{f^2} \right) \tan^2 \theta. \tag{9c}
\end{align*}
\]

We can find the slant angle with the following equation:

\[
\theta = \pm \arctan \left( \sqrt{\frac{\beta_2}{\beta_0}} \right). \tag{10}
\]

To generalize for all tilt angles, we rotate the image coordinates with the tilt angle. Equation (8) becomes

\[
E(x, y) = a_1 x^2 + a_2 x y + a_3 y^2 + a_4 x + a_5 y + a_6, \tag{11}
\]

with

\[
\begin{align*}
a_1 & = \beta_2 \cos^2 \tau; \quad a_3 = \beta_2 \sin^2 \tau; \quad a_5 = \beta_1 \sin \tau; \\
a_2 & = 2\beta_2 \cos \tau \sin \tau; \quad a_4 = \beta_1 \cos \tau; \quad a_6 = \beta_0. \tag{12}
\end{align*}
\]

To obtain \(E(x, y)\), we have to know the six parameters \(a_i\). We must interpolate the map of the local scales, \(E_L(x, y)\), with (11). In the following work, which has the aim to detect defects, only the six coefficients from (11) are necessary. The knowledge of the tilt and slant angles is not important in this way, but in another work we studied the accuracy of the method for these angles [17]. With these six values \(a_i\) we obtain the theoretical value of the local scale for each pixel: \(E(x, y)\). This equation, which uses a sinusoidal model, can be used for textures with one main frequency. This is the limitation of the method.

3. MESH COMPUTATION

Now that we have the theoretical value of the local scale for each pixel, we want to map a mesh on the texels which form the textures. First we have to extract the texels. Then we perform a Delaunay triangulation and we create the mesh.

To extract texels from the local scales map, we have to use a method adapting locally to the size of the texels. This problem is present for the inclined planes, because the perspective projection introduces a variation of the texel size due to the scaling cue [18]. In order to separate the points belonging to the texels from these of the remainder of the image, we have to threshold the local scales map, \(E_L(x, y)\), in two classes. This task can be done if we have a threshold value \(T(x, y)\) for each point of the image, because this value depends on the local size of the texels. Thanks to the interpolation of the local scales map by a paraboloid, by our SFT method, we know the size, that is the scale, of the texels in each pixel. The texels are represented by high scales in the local scales map, because they have relatively important surfaces in the image. Besides, the edges of texels introduce low scales in the map. Thanks to these observations, we can say that a point, in the local scales image, belongs to a texel if its scale is greater or equal to the theoretical scale of the texel.

Figure 5 illustrates this principle. In Figure 5b, we present a local scales map obtained from the original image (Figure 5a). Figure 5c shows the scales from a line drawn in Figure 5b. To find the resulting image, we use the following algorithm:

\[
\text{Resulting image} = \begin{cases} 
0 & \text{if } E_L(x, y) \geq T(x, y), \\
1 & \text{if } E_L(x, y) < T(x, y), 
\end{cases} \tag{13}
\]

with \(T(x, y) = E(x, y) P\). In this equation, \(E_L(x, y)\) represents the local scales map obtained after the wavelet decomposition, and \(E(x, y)\) the interpolated scales computed with (11) using the six \(a_i\) parameters. The factor \(P\) is the margin allowed to account for the lack of precision of the interpolation. \(P\) is experimentally fixed at 130% for all the images we have tested.

In Figure 6c, we show the texels extracted from the original image (Figure 6a) with this method. When we have a binary image with all the texels, we compute for each of them the coordinates of their gravity center.

With these centers of texels, it is possible to create a mesh using a Delaunay triangulation [19]. This triangulation describes the spatial organization of the texture. In Figure 6, we illustrate the main steps which allow us to obtain this triangulation from a textured image (Figure 6a). These steps are the following:

- to initialize the process, we build one triangle in the original image from three texels centers. It is the first triangulation (Figure 6d);
- then we integrate a new texel center into the triangulation by using the incremental method, so we obtain another triangulation;
- the last step is repeated for all the centers of the texels of the image (Figure 6e).

After these three steps, we have a Delaunay triangulation of the texels’ centers (Figure 6f). Now, we have to create a regular mesh from this triangulation. First we remove all the triangles which are created in order to obtain a convex set, because they do not have to belong to the final mesh. Then we arrange triangles, two by two, in order to create “quasi-parallelograms” whose sides are the axes of the mesh (Figure 6g).

4. DEFECTS EXTRACTION AND RESULTS

As we explained in the introduction, this method is applied to regular macroscopic textures (see Figure 1). For these
Defect detection on inclined textured planes

Figure 5: Principle of adaptive thresholding. (a) Original image; (b) local scales map from (a); (c) values of the local scales on line 128 of (b) and threshold value for each pixel.

Figure 6: Different steps in order to obtain a regular mesh from a natural texture.

Figure 7: Different steps in order to obtain a regular mesh from a natural texture. We disturb the regularity of this texture by shifting a set of texels. In Figure 7a, we see how this defect modifies the local scales map. If there is a defect in the texture, some texels disappear during the texels extraction stage because the local scales map is modified around the defect. In this area, the adaptive threshold is not suitable for these corrupted texels. Some other texels are detected, but their centers are not in line with their neighbors. So, the corrupted texels are eliminated when we build the mesh, because the Delaunay triangulation is not regular in the area of the defect (Figure 7c). To finish, there is a hole on the final mesh (Figure 7d). So, to detect possible defects, we look for holes in the mesh. All these areas with holes are considered as defect areas.

This method has been tested on synthetic and real textures. We have created different kinds of defects: a spot made by a uniform area, a piece of tissue of the same or different texture, a ball dressed with the same texture put on the plane. Figure 8 shows a result for a synthetic texture (Figure 8a) where the defect is a missing texel which is replaced by a set of high frequencies. In Figure 8b, there is a magnification of the defect area. We can see that only one texel is corrupted. The hole in the mesh is shown in Figure 8c and the defect position in Figure 8d. In Figures 9, 10, and 11, we present...
defect detections in real images. In Figures 9a and 10a, the defect is a piece of the same tissue as the background. Even if the texels of the defect are in line with the ones of the textured plane, we see the hole in the mesh (Figures 9b and 10b). The localization of the defect is very good as it is possible to observe in Figures 9c and 10c. In these two examples, the main frequency of the texture (background and defect) are very different: medium frequency for Figure 9 and high frequency for Figure 10. Even if a part of the defect is not detected (Figure 10), we can see the efficiency of the method. Figure 11a presents two defects simultaneously, a uniform spot and a pleat in the tissue dressing on the plane. The spot is a low frequency area, so there is a zone of high scales (white pixels in Figure 11b) in the local scales map. The pleat is very difficult to see. Figure 11c shows a zoom of this defect zone. Texels which are on the pleat are smaller than the others and the intertexel distance is slightly modified, so the threshold value was not good for these corrupted texels. Figure 11d shows the mesh obtained with the extracted centers of texels. Two holes are clearly present. We can see in Figure 11e that the two defects are perfectly detected and localized on the original image.

In these four examples, the frequencies of the defects are very different: high frequencies in Figure 8, the same frequencies as the background in Figures 9 and 10, and low frequencies in Figure 11. In these four cases, it is always possible to build the mesh and to automatically find the defect.
5. CONCLUSION

In this paper, we have shown the interest of the combination of a shape from texture (SFT) method with the Delaunay triangulation technique in order to detect defects on inclined textured planes. The accuracy of this method, which is fully automatic, is very good for defects having different frequencies and different sizes. In all the examples we have used to test this method, the defects are always detected. If only one texel is corrupted, it is possible to detect it. If the defect is smaller than the texels or the inter-texels distance, it is not possible to detect it. In addition to the detection, we obtain the position and the area of the defect. All these features allow us to take a decision about the importance of the defect and the processing. One of the applications of such a method is to perform textile control quality, but the tissue to be controlled has to show a regular macroscopic texture.

The position of the camera has no importance with respect to the tissue position. It is not necessary to calibrate the camera. When we interpolate the local scales map, we use (11), and we obtain for each pixel the interpolated value $E(x, y)$. If we do not compute the slant angle, camera parameters are not necessary.

This method assumes no knowledge about the tissue and its orientation. But, sometimes, it is difficult to detect the defect, when the defect area is too large. In order to solve this problem, it is possible to think of a processing which will include knowledge about:

- the six interpolation parameters of (11) will be obtained by using an image without a defect and will give a fixed threshold value for each pixel;
- it is also possible to compute the orientation of the tissue. Then we are able to compute the theoretical position of all the texels and to interpolate the whole mesh.
This method is not fast because of the computation of the local scales map. The computation time is related to the number of scales used in this step. With knowledge about texture, it is possible to use less scales. Then the computation time would be reduced.

REFERENCES