3D Reconstruction Approach based on Wavelet Analysis and Neural Network in Neuro-vision System

YINGEN XIONG AND FRANCIS K H QUEK
Department of Computer Science and Engineering
Wright State University
348 Russ Engineering Center, 3640 Colonel Glenn Hwy, Dayton, OH 45435
USA

Abstract: - In this paper, a 3D reconstruction approach based on wavelet analysis and neural networks is presented. It can be used in neuro-vision system. The approach can be divided into two parts. First, the stereo matching problem is solved with wavelet analysis. Dyadic discrete wavelet analysis is adopted in this process and stereo matching process is realized with global optimization. A coherent hierarchical matching strategy is constructed, so that the stereo matching process can be accomplished with coarse to fine techniques. Second, a 3D object reconstruction neural network is constructed by using BP neural network. By feeding the image corresponding points between the left image and right image in a stereo image pair, the 3D coordinates of points on object surface can be obtained using this neural network and the configuration and shape of the object can be reconstructed. With multiple 3D reconstruction neural networks the 3D reconstruction processes can be performed in parallel. The examples for both synthetic and real images are shown in the experiment, and good results are obtained.

Key-Words: - 3D reconstruction, Stereo matching, Neural network, Wavelet analysis, Computer vision.

1 Introduction

Computer vision system takes an important part in intelligent manufacturing environment. Whether the system is performing inspections, locating parts, or deciding what parts are in the field of view, it is recognizing a particular pattern in an intensity image. While many applications could be solved by performing only 2D object recognition. More complex applications and the need for added flexibility are requiring the acquisition of three dimensional (3D) information along with 3D object recognition. Stereo vision is needed in this system.

Stereo vision is a technique for building a 3D description of a scene observed from several viewpoints. In the passive stereo vision, usually two cameras are used to observe the same scene from two slightly different viewpoints and the stereo image pairs are obtained. After the left image and right image in the stereo image pair are matched. The 3D description of the scene can be built. Stereo matching process takes an important part in the 3D description.

The objective of image matching is to establish correspondence between the selected primitives, so that the matched primitive arises from the same element in the scene. For the stereo vision, the stereo matching problem is to establish the correspondence between left image and right image in a stereo image pair. There have been several attempts to formulate solutions to the problem which may be categorized into feature and area based methods[1][2].

Many of problems encountered in stereo matching in the spatial domain can be avoided by treating the stereo matching in frequency domain by using local phase. The local phase has been used for stereo matching process[3], and this is the third category named phase based approach. Recently, wavelet analysis has been used for stereo matching process[4][5]. Wavelet decomposition provides a mathematically precise definition to the concept of multiresolution. The multiresolution representations derived from wavelet frames provide a mathematically coherent hierarchical framework to accomplish a coarse to fine incremental matching strategy.

In this paper, a 3D object reconstruction approach based on wavelet analysis is presented. First, a coherent hierarchical matching strategy is constructed, so that the stereo matching process can be accomplished with coarse to fine techniques. The dyadic discrete wavelet analysis is adopted in this process and the stereo matching process is realized with global optimization. Second, a 3D object reconstruction neural network is constructed by using BP neural network. By feeding the image corresponding points between the left image and right image in a stereo image pair, the 3D coordinates of points on object surface can be obtained using this
neural network and the configuration and shape of the object can be reconstructed. With multiple 3D reconstruction neural networks the 3D reconstruction processes can be performed in parallel.

2 Problem Formulation

2.1 2D wavelet analysis

The 2D wavelet transform of the function \( f(x, y) \in L^2(R) \) with a given mother wavelet \( \Psi \) is defined by

\[
T_\Psi f(s,u,v) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) \Psi_{s,u,v}(x,y) dx dy \quad (1)
\]

After discretizing both the scale factor \( s \) and the translation \((u, v)\), we can obtain the dyadic scale space.

\[ s = \frac{1}{2^j} \quad \text{and} \quad (u = k, v = l) \quad \text{with} \quad (j, k, l) \in \mathbb{Z}^3 \quad (2)\]

where \( \mathbb{Z} \) denotes the set of integers.

Let \( D_j f \) denote the difference between two approximations \( A_{j-1} f \) and \( A_j f \),

\[
D_j f = A_{j-1} f - A_j f \quad j = 1, 2, \ldots, n \quad (3)
\]

where \( A_0 \) is an identity operator. \( A_j f \) is the approximation of a given function \( f(x, y) \) at a scale \( s = \frac{1}{2^j} \), \( j = 0, 1, 2, \ldots, n \).

There are three components for the difference between two approximations \( A_{j-1} f \) and \( A_j f \)

\[
D_j f = A_{j-1} f - A_j f = D_{j,1} f + D_{j,2} f + D_{j,3} f \quad (4)
\]

so the multiresolution analysis of a function \( f(x, y) \) can be written

\[
f(x, y) = A_0 f + D_{1,1} f + D_{1,2} f + D_{1,3} f + D_{j,1} f + D_{j,2} f + D_{j,3} f + \sum_{j=1}^{n} \[D_{j,j} f + D_{j,j} f + D_{j,j} f] \quad (5)
\]

Given a discrete image \( f(x, y) \) with a limited support \( x, y = 1, 2, \ldots, 2^n \), the actual procedure for constructing the wavelet pyramid involves computing the coefficients \( a_{j,k,l}, d_{j,p,k,l} \), which can be grouped into four matrices \( A_j, D_{j,p}, p = 1, 2, 3 \), on each level \( j \)

\[
A_j = (a_{j,k,l}) \quad D_{j,p} = (d_{j,p,k,l}) \quad \text{for} \quad k, l = 1, 2, \ldots, 2^{n-j} \quad (6)
\]

Let \( h \) and \( g \) be the impulse response of the filter \( \phi \) and \( \psi \), the coefficients \( d_{j,k,l} \) and \( d_{j,p,k,l} \) \( p = 1, 2, 3 \) can be computed via an iterative procedure. The wavelet pyramid of image \( f(x, y) \) and its reconstructing process are illustrated in figure 1 and figure 2.

\[
\text{Figure 1 Wavelet pyramid of an image } f(x, y)
\]

2.2 Image matching with global optimization

Stereo matching process is to find correspondence between left image and right image in stereo image pair. For any image point \((k, l)\) on the reference image, its approximate correspondence \((k_0, l_0)\) on the matched image may be obtained through some general strategies, such as the spiral and hierarchical parallax propagation. The simplest way to catch the precise correspondence \((k', l')\) is the discrete search in a small neighborhood of \((k_0, l_0)\) which is defined by a distance threshold \(T\)

\[
\min_{(k', l')} S_j((k, l), (k', l')) \quad \forall(k', l') : \| (k', l') - (k_0, l_0) \| \leq T \quad (7)
\]
The distance threshold \( T \) should be defined in such a way that the allowed errors in the parallax propagated from the last higher level can be corrected. In general \( 1 \leq T \leq 2 \). Note that the search takes place not only on normal integer positions, but also on diagonal positions. The computational structure and the data flow of wavelet frame processing for stereo matching is shown in figure 3.

\[ A_j \quad D_{j,1} \quad D_{j,2} \quad D_{j,3} \]
\[ y \downarrow 2 \quad y \downarrow 2 \quad y \downarrow 2 \quad y \downarrow 2 \]
\[ h(y) \quad g(y) \quad h(y) \quad g(y) \]
\[ x \downarrow 2 \quad x \downarrow 2 \]
\[ h(x) \quad g(x) \]
\[ A_{j-1} \]

**Figure 2 The flowchart for the analysis from level \( j-1 \) to level \( j \)**

3 Three Dimensional Reconstruction

In the conventional 2D to 3D points conversion process, an elaborate system calibration to determine the relationships between the two cameras and the global co-ordinate system is required. Using 3D points known in 3D space and corresponding 2D points on images to create the relationship between the cameras and global co-ordinate system is usually the approach. Then each three dimensional point is found by feeding the two matching points into equations governing the transformations. Usually the Collinear is required. Instead of conventional methods, here we construct a four layers back propagation neural network to perform the operation. The input layer has four neurons for X and y values of the matched points on the left and right images of a stereo image pair. The network has two hidden layers. Each contains 50 neurons for storing the complex relationship between 2D image and 3D space. The output layer has 3 neurons for the 3D points. The architecture of the neural network is given as Figure 4.

\[ X \]
\[ Y \]
\[ Z \]

**Figure 4 The architecture of 2D to 3D point conversion neural network**

By feeding the matched points into the input layer, the neural network can output the 3D points of the object. If multiple networks are available points can be converted in parallel.
4 Experiment Results and Discussion
We have tested the approach presented in this paper with a series of synthesized images and nature images. All images shown here are composed of about 512x512 pixels. During the stereo matching processes, the epipolar lines are used to provide useful constrain for neural networks.

The example is a stereo image pair of a turbine blade (shown in figure 5). The left and right images are taken at parallel position. Figure 6 shows the 3D plot of the turbine blade reconstructed from the results of stereo matching. It can be seen that the 3D shape and surface are well recovered. This example is designed to test the ability of the approach to be used in practical situation. From this test we can see that the results can be used in the 3D reconstruction of object.

4 Conclusion
In this paper, a 3D reconstruction approach in neuro-vision system is presented. The stereo matching problems are solved based on wavelet analysis. In the dyadic scale space, the image can be decomposed into a set of frequency channels having the same bandwidth on a logarithmic scale. By adaptation to scale and translation a course to fine image matching process can be constructed. With the image matching strategy the corresponding points between left image and right image in stereo image pair can be obtained. A 3D object reconstruction neural network is constructed by using BP neural network, so the 3D configuration and shape can be reconstructed by this neural network. With multiple neural networks the 3D reconstruction processes can be performed in parallel. Some examples and applications for this approach are also presented. We have tested this approach with some synthesize and nature images. From the results of the examples and applications, we can see that the shapes and surface of objects are well recovered from the results of stereo matching. The approach presented in this paper can be used in computer vision and intelligent assembly system.

Please, follow our instructions faithfully, otherwise you have to resubmit your full paper. This will enable us to maintain uniformity in the conference proceedings as well as in the post-conference luxurious books by WSES Press. The better you look, the better we all look. Thank you for your cooperation and contribution. We are looking forward to seeing you at the Conference.

References: