A Multiresolution Approach for Video Texture Registration

Robert J. Bonneau, Mark E. Novak, James S. Perretta, Sevgi Ertan
AFRL/SNRT
26 Electronics Pkwy, Rome NY 13441

Abstract

Electro-optical imagery can have uniform characteristics that prevent it from being registered by conventional edge-based methods. Such uniform characteristics, if they have periodicity, can be exploited using multi-resolution texture extraction techniques. We will first use a multi-resolution Markov model to represent electro-optical textures and apply an auto-regressive statistical approach to find correspondence between two images. We will then demonstrate how this approach reduces the computational complexity of registering of two successive frames of video.

1. Introduction

These We will first describe the multiresolution Markov data structure and how it describes texture of the imagery. We will then describe our texture matching procedure and how it is applied to successive frames of video imagery. Finally we will show an example of the registration process and describe how it reduces the computational complexity of video registration over conventional techniques.

2. Wavelet Markov Data Structure

First we shall focus on multiresolution wavelet filter bank transform structures. The Mallat filter bank structure [2] shown in Figure 1 is the standard wavelet decomposition of most compression applications. The G and H filters are the high and lowpass filters respectively and each one is applied along the x and y axis alternately to extract the HH, HL, LH, and LL frequency band decompositions of the signal.

Recently there have been many studies that have shown the optimally of the wavelet transform domain for representation of texture. Other wavelet registration models have employed ad-hoc metrics for feature registration [3] which do not accurately represent texture. In order to work in this domain we must first represent our signal and noise process with an appropriate data model. Such a model is known as the wavelet Markov random field model

To represent this Markov random field [1] we define a given node in the quad tree structure as s, its children nodes as $s\alpha_{NW}, s\alpha_{NE}, s\alpha_{SE}, s\alpha_{SW}$ and its parent node as $s\bar{\gamma}$ where $\bar{\gamma}$ shifts the wavelet coefficients from parent $s\bar{\gamma}$ to child s as is shown in Figure 2. A Kth order model defined on the multiresolution structure is defined in either 1 or 2 dimensions with

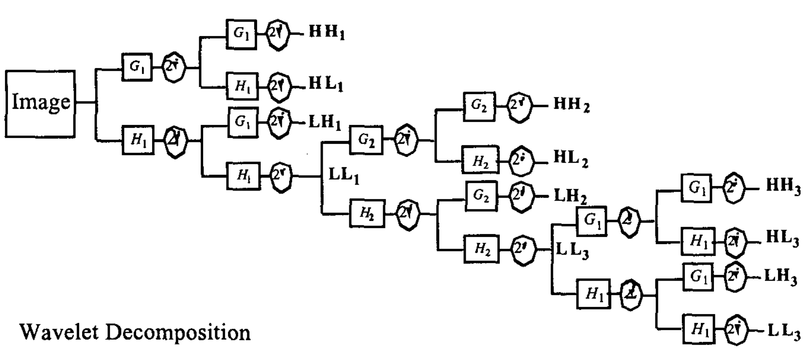


Figure 1 Wavelet Decomposition

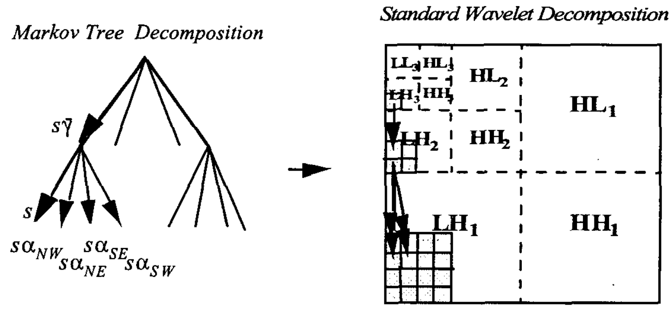


Figure 2 Markov Block Structure

$t \in \{1, 2, \ldots, K(T+1)\}$.

Now, defining a MRF on a $2^N \times 2^N$ lattice, a state at the mth level represents the values of the MRF at $16(2^{N-m}-1)$ points. This set of points is denoted as Γ_s and it is the union of 4 mutually exclusive subsets. In general we can divide Γ_s into four set sets of $4(2^{N-m(s)}-1)$ points in a similar fashion, and we denote these subsets as $\Gamma_{s,i}, i \in \{NW, NE, SE, SW\}$. Now if we have the random variable Z representing the current state of any Γ_s at any stage of the tree then we insert our local scale iterative relationship, the basic probabilistic Markov relationship is defined as

$$p_{Z_t, t \in \Gamma_{s\alpha_i} | Z_T, T \in \Gamma_s}(Z_t, t \in \Gamma_{s\alpha} | Z_T, T \in \Gamma_s) = p_{Z_t, t \in \Gamma_{s\alpha_i} | Z_T, T \in \Gamma_{s,i}}(Z_t, t \in \Gamma_{s\alpha} | Z_T, T \in \Gamma_{s,i}) \quad (1)$$

Once we have defined the Markov structure from the wavelet transform we next take the individual coefficient elements and represent them using an autoregressive set of equations as is shown in equation 2. Our desired texture is represented by the polynomial coefficients [1] A(s) added to a Gaussian noise component w(s) represented by the B(s) coefficients.

$$x(s) = A(s)x(s\bar{\gamma}) + B(s)w(s) \quad (2)$$

In the image context we can represent the elements of the image Markov structure in terms of the recursive scale structure shown in Figure 3 with where the superscripts R represent the scale and its associated coefficients.

$$I(s) = a_1(s)I(s\bar{\gamma}) + \ldots + a_R(s)I(s\bar{\gamma}^R) + w(s), a_i(s) \in \Re \quad (3)$$

Representing equation 2 in matrix form we have equations 4 and 5.

$$x(s) = [I(s) \quad I(s\bar{\gamma}) \quad \ldots \quad I(s\bar{\gamma}^{R-1})]^T \quad (4)$$

$$x(s) = \begin{bmatrix} a_{1,m(s)} & a_{1,m(s)} & \cdots & a_{1,m(s)} & a_{1,m(s)} \\ 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \ldots & 0 & 1 & 0 \end{bmatrix} x(s\bar{\gamma}) + \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} w(s) \quad (5)$$

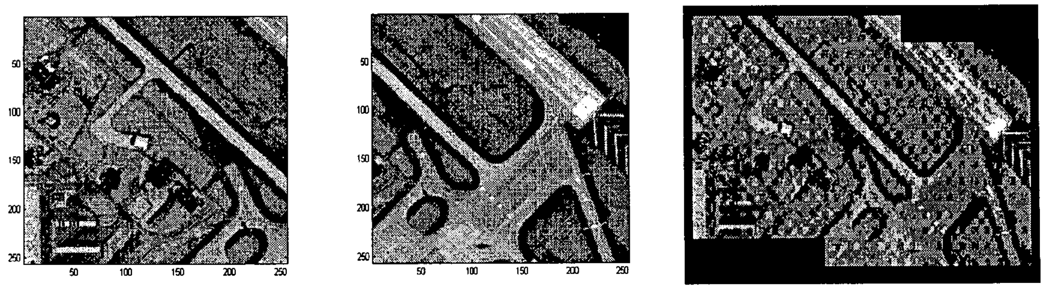
We can characterize a given signal or texture in our image by solving for the autoregressive coefficients a(s) equation 5 first for our representative texture or object. Then using these coefficients we attempt to use them to predict the target signal in a given input signal x(s). The residual between the target signal and the input signal is then w(s). This model assumes that our target signal is uncorrelated with the input signal.

3. Video Texture Registration Scheme

In order to register two successive frames of video we now apply a test statistic based on the residual between predicted coefficient and any given image coefficient as $w(s_c)$. We denote μ_c as the average of the expected residual for the object model and p_c is the standard deviation of those coefficients used to normalize the statistic $\varsigma(s_c)$

$$\varsigma(s_c) = \frac{w(s_c) - \mu_c}{\sqrt{p_c}} \quad (6)$$

Equation 6 indicates where the best match in texture occurs such that the residual $w(s_c)$ and thus the test statistic is at a minimum.



Frame 1 Frame 2 Registered Frames

Figure 3 Video Registration Output

Our registration procedure begins by decomposing two successive frames of video images using our wavelet transform. This registration process consists of computing how closely our texture from frame 1 corresponds to the texture in frame 2 using our multiscale Markov texture matching process described in equation 6. The registration point for both images is then the point that corresponds to the best texture match.

The algorithm is applied by subsampling the image to its lowest resolution scale, texture matching the imagery on the lowest two scales of resolution, then upsampling the imagery and matching on the next two successive scales. Since the registration is accurate to within 2x2 pixel regions after the first match is performed, the actual number of pixels which are necessary to search is dramatically reduced. The results are shown in Figure 3.

Thus by registration in this multiresolution manner the grid over which the registration point search is conducted is limited thereby decreasing the computational complexity of the search. The number of operations in conventional methods frame to frame registration is typically $N^2 \log N$ where N is the number of pixels along a side in an NxN image. where our method because of its multiresolution technique is simply NlogN operations.

4. Conclusion

We have demonstrated improved performance in video frame registration by casting the problem in terms of a multiresolution texture matching approach. Because of the multiresolution structure, we are able to focus on texture features rather than edges which are not always available in imagery and significantly improve the speed of the registration process. This method is useful for a wide range of applications where fast accurate frame registration is essential.

References

1. Luettgen, M., Karl, W., Willsky, A., Tenny, R., "Multiscale Representations of Markov Random Fields", IEEE Transactions on Signal Processing, Vol 41, No 12, pp. 3377-3396, December 1993.

2. S. Mallat, "Multiresolution Approximations and Wavelet Orthonormal Bases of L2R", Transactions of the American Mathematical Society, Volume 315, Number 1, September 1989.

3. Zheng, Q, Chellapa, R., "A Computational Vision Approach to Image Registration", IEEE Transactions on Image Processing, Vol 2., No 3., pp. 311-326 July 1993.