Multiresolution Mosaic Images by using Laplacian of Gaussian method:A REVIEW

M.K.NEHARKAR[1], PROF.S.K.SUDHANSU[2] Dr.VEERESH G.K[3]
[1, 2, 3] P. G. Department, MBES College of Engineering, Ambajogai, India

ABSTRACT

Image mosaicing is the act of combining two or more images. It aims to combine images such that no obstructive boundary exists around overlapped regions. Emphasis is given on to create a mosaic image that contains as little distortion as possible from the original images, as well as preserving the general appearance of the original images. Multiresolution representation technique is an effective method for analyzing information contents of signals, as it processes the signals individually at finer levels, to give more accurate results that contains much less distortion. Laplacian pyramid, Gaussian pyramid and Wavelet transform are types of Multiresolution representations. In this work, we use Laplacian pyramid using Gaussian pyramid which superiors other transforms in context of simplicity and working satisfactorily in real time domain. Application areas of our subject are widespread in the fields like signal analysis, image coding, image processing, computer vision and still counting. The work on this project will be focused on designing a model which balances the smoothness around the overlapped region and the fidelity of the blended image to the original image.

Keywords – **Image processing, Multi-resolution image, Mosaicing, Laplacian, Gaussian method**.

1.INTRODUCTION

When two or more images are overlapped to form a single mixed image, finding an ideal image combination can be difficult. An image mosaic processing technique can be applied to greatly reduce this difficulty. To mosaic an image is to combine overlapped images so that the mixed image contains no obstructive boundaries in the transition region while care is taken to preserve the general appearance of the original images. An image mosaic is typically completed in two stages. In the first stage, the corresponding points in the two, to-be-combined images are identified and registered. This stage is usually referred to as image registration. Not all applications of image mosaicing require registration, such as in movie special effects. In the second stage, the intensities of the images are blended after the corresponding points have been registered [1].

1.1NEEDFOR MULTIRESOLUTION MOSAICING

In psychophysics and the physiology of human vision, evidence has been gathered showing that the retinal image is decomposed into several spatially oriented frequency channels. This explains the use of Multiresolution decomposition in computer vision and image processing research and why Multiresolution Spline approach works well for image mosaic [3]. The basic concept is to decompose the signal spectrum into its Sub spectra, and each sub spectrum component can then be treated individually based on its characteristic. For example, most nature signals will have predominantly low frequency components, thus the low-band components contain most of significant information, while for a texture the most significant information often appears in its middle-band component. Thus each channel can be processed separately to obtain more precise results in the technique of Multiresolution. In this work, design of a module that can decompose a given image into number of levels depending on the image size has been discussed. The Multi-resolution process is carried out with the help of Laplacian pyramid using Gaussian pyramid [4]. A blending technique needs to be implemented for combining two or more images into a larger image mosaic. In this procedure, the images to be blended are first decomposed into a set of band-pass filtered component images. Next, the component images in each spatial frequency hand are assembled into a corresponding band pass mosaic. In this step, component images are joined using a weighted average within a transition zone which is proportional in size to the wave lengths represented in the band. Finally, these band-pass mosaic images are summed to obtain the desired image mosaic. When coarse features occur near borders, these are blended gradually over a relatively large distance without blurring or otherwise degrading finer image details in the neighborhood of the border[1,6,7].

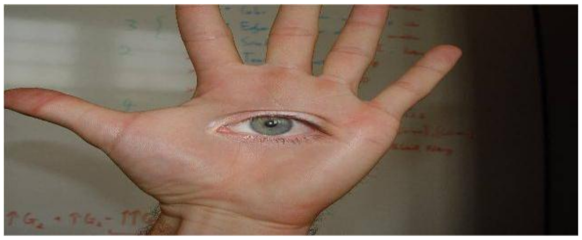


Fig1. A typical multiresolution mosaic image

2.MULTIRESOLUTION

The frequency and time information of a signal at some certain point in the time-frequency plane cannot be known. In other words, we cannot know what spectral component exists at any given time instant. The best we can do is to investigate what spectral components exist at any given interval of time. This is a problem of resolution, and it is the main reason why WT is popular than STFT. Since the previously used transforms like short time Fourier transform (STFT), Wigner distributions were not able to give good time and frequency resolution. Instead they were giving fixed resolution. To be simpler, every spectral component is not resolved equally in STFT. Thus the information provided by them was highly redundant in nature as far as reconstruction of the signal is considered. The Fig. 2 clearly shows how actually two images are combining to form mosaic image.

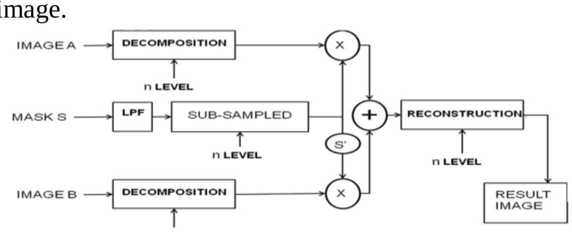


Fig.2 N level decomposition of image

First image (consider it as A) is take, then it is decomposed up to N level as per requirement of user. Similarly we have taken second image (consider it as B).Now we need to design mask with same size as that of the image size. Mask is nothing but binary representation of image in to be combine images. This dummy image is used as mask for hiding appropriate part of image, i.e. Mask is a outer part of 'A' image & inner part of 'B' image. To get multiresoved format mask for each level of decomposition I have to use low-pass filter and then sub-sampled [5].

Image is nothing but matrix of values, hence direct multiplication of mask with image is taken. Then two masked images are obtained, which are then combine to form the resultant image at each level of resolution. Now Using these entire components original image is reconstructed. After reconstruction we get the resultant mosaic image.

3. MULTIRESOLUTION TECHNIQUES

Following are the two important pyramid structures. Hierarchical representation of an image is shown in Fig. 3.

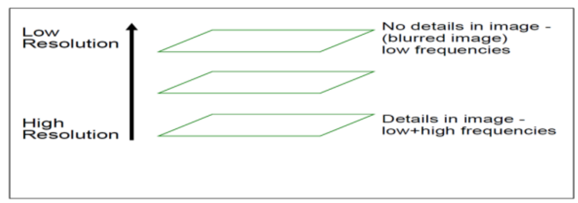


Fig. 3 Hierarchical representation of an image

3.1 GAUSSIAN PYRAMID

The first step in Laplacian pyramid coding is to low-pass filter the original image $g0$ to obtain image g_1. g_1 is a "reduced" version of g_0 in that both resolution The same 5-by-5 pattern of weights w is used to generate each pyramid array from its predecessor. This weighting pattern, called the generating kernel, is chosen subject to certain constraints. For simplicity we make w separable

$$W(m, n) = \hat{w}(m)\,\hat{w}(n). \qquad (1)$$

The one-dimensional, length 5, function $\hat{w}$ is normalized on and sample.

$$\sum_{m=-2}^{2} \hat{w}(m) = 1 \qquad (2)$$

And symmetric

$$\hat{W}(i) = \hat{w}(-i) \text{ for } i = 0, 1, 2.$$

An additional constraint is called equal contribution. This stipulates that all nodes at a given level must contribute the same total weight (=1/4) to nodes at the next higher level [4]. Let $\hat{w}(0) = a$, $\hat{w}(-1) = \hat{w}$

(1) = b, and wˆ (–2) = wˆ (2) = c in this case equal contribution requires that a + 2c = 2b. These three constraints are satisfied when

$$\hat{w}(0) = a \quad (3)$$

$$\hat{w}(-1) = \hat{w}(1) = 1/4 \quad (4)$$

$$\hat{w}(-2) = \hat{w}(2) = 1/4 - a/2 \quad (5)$$

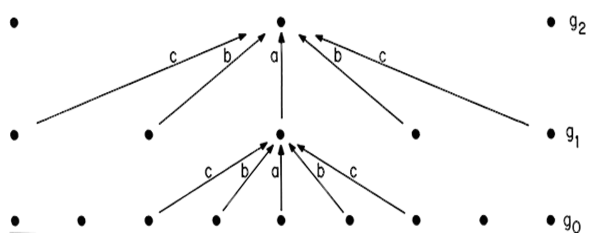


Fig 4.One-dimensional graphic representation of the process which generates a Gaussian pyramid

Fig.4illustrates the contents of a Gaussian pyramid generated with a = 0.4. The original image, on the far left, measures 257 by 257. This becomes level 0 on the pyramid. Each Higher level array is roughly half as large in each dimension as its predecessor, due to reduced sample density.

3.2 LAPLACIAN PYRAMID

Our purpose for constructing the reduced image g_1 is that it may serve as a prediction for pixel values in the original image g_0. To obtain a compressed representation, we encode the error image which remains when an expanded g_1 is subtracted from g_0. This image becomes the bottom level of the Laplacian pyramid. The next level is generated by encoding $g1$ in the same way [3, 4].

The Laplacian pyramid is a sequence of error images L_0, L_1, L_N. Each is the difference between two levels of the Gaussian pyramid. Thus, for 0 < 1 < N,

$$L_l = g_l - \text{EXPAND}(g_l + 1)$$
$$= g_l — g_l + 1.\ 1. \quad (6)$$

Since there is no image g_N + 1 to serve as the prediction image for g_N, we say $L_N = g_N$. The original image can be recovered exactly by expanding, then summing all the levels of the Laplacian pyramid.

$$go = \sum_{l=o}^{N} Li, i - \quad (7)$$

A more efficient procedure is to expand L_N once and add it to $L_N - 1$, then expand this image once and add it to $L_N - 2$, and so on until level 0 is reached and g_o is recovered. This procedure simply reverses the steps in Laplacian pyramid generation.

$$g_N = L_N$$

And for $l = N - 1, N - 2\ldots 0$,

$$g_l = L_l + \text{EXPAND}(g_l + 1). \quad (8)$$

The Laplacian pyramid is a versatile data structure with many attractive features for image processing. It represents an image as a series of quasi-band passed images, each sampled at successively sparser densities. The resulting code elements, which form a self-similar structure, are localized in both space and spatial frequency. By appropriately choosing the parameters of the encoding and quantizing scheme, one can substantially reduce the entropy in the representation, and simultaneously stay within the distortion limits imposed by the sensitivity of the human visual system.Fig.5summarizes the steps in Laplacian pyramid coding. The first step, shown on the far left, is bottom-up construction of the Gaussian pyramid images g_0, g_1, g_N. The Laplacian pyramid images L_0, L_1, L_N are then obtained as the difference between successive Gaussian levels [3, 8]. These are quantized to yield the compressed code represented by the pyramid of values C_l (i, j). Finally, image reconstruction follows an expand and sum procedure.

It should also be observe that the Laplacian pyramid encoding scheme requires relatively simple computations. The computations are local and may be performed in parallel, and the same computations are iterated to build each pyramid level from its predecessors. Envision of performing Laplacian coding and decoding in real time using array processors and a pipeline architecture.

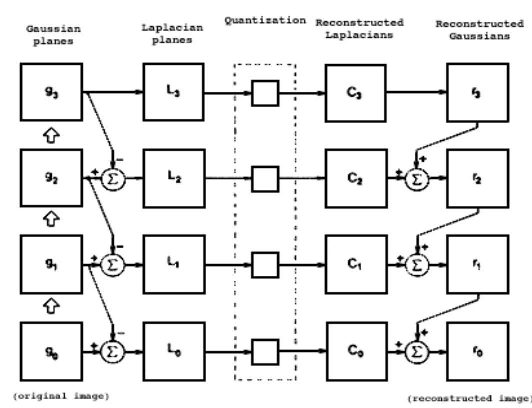


Fig. 5 Summary of the steps in Laplacian pyramid coding and decoding

4. MASK DESIGN

In psychophysics and the physiology of human vision, evidence has been gathered showing that the retinal image is decomposed into several partially oriented frequency channels. This explains why Multiresolution decomposition methods are so popular in computer vision and image processing research and why Multiresolution Spline approach works well for image mosaic [3, 7, 9]. Since the low-frequency content of a signal are often sufficient in many instances (such as the content of an image), and the detail information resembles the high frequency components (such as edge of an image), thus, the width of the transition zone *T* is chosen according to the wave length represented in each band. That is, for lower frequency components, the width of transition zone *T* is chosen to be larger than that of higher frequency components. This implies that low-frequency components "bleed" across the boundary of mosaic region further than High-frequency components do. Actually as described in the above section, the signal $c_{m+1,n}$ at resolution $m+l$ is a smoothed down-sampled approximation of $c_{m,n}$ at resolution m, and $d_{m+l,n}$ is just the detail (or difference) information between $c_{m,n}$ and $c_{m+1,n}$. Therefore , using the same width of transition zone between detail component of resolution m and its down-sampled components in resolution m+l means the actual transition zone of the low-frequency components is larger than that of the high-frequency[3].

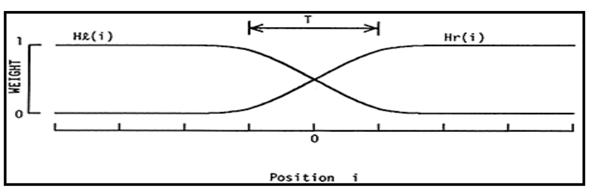


Fig. 6 The weighted average function W(x)

To simplify and generalize arbitrary shape mosaic both for 1-D and 2-D signals [2], the transition zone *T* and the weighting function are not explicitly expressed, instead, another Multiresolution structure of mask signal is introduced. The mask signal is a binary representation which describes how two signals will be combined. For example, two signals A and B will be combined to form a mosaic signal, and the mask signal *S* is a binary signal in which all points inside the mosaic region are set to 1 and those outside the mosaic region are set to 0. As the way to generate a sequence of lower resolution signal (not the detail signal) describe in the above section, the mask signal S is low-pass filter and sub sampled to construct its Multiresolution structure $c'_{Mn}(S)\ldots c'_{1n}(S)\ c'_{2n}(S)$ and then each smoothed version of the mask signal will be used as the weighting function in its corresponding resolution level. The low-pass filter used to construct Multiresolution structure of the mask signal need not be the same as the DWT used, which is why here we use c'Mn(S) instead of cMn(S).Now, both 1-D and 2-D DWT-based signal mosaics will be described.

4.11-D SIGNAL MOSAIC

Suppose two I-D signals A(x) and B(y) will be combined together to form a new signal C(x), and complete overlapping of the two signals is assumed. A mask signal S(x) is designed to describe the mosaic region.

4.2EXTENSION TO 2-D IMAGE MOSAIC

There are various expressions of I-D wavelet transform to higher dimensions and Mallat's method is adopted in this works.

The 2-D wavelet basis function can then be expressed by the tensor product of two I-D wavelet basis functions along the horizontal and vertical directions. Then, the scaling function is [2],

$$\Phi(x, y) = \Phi(x)\ \Phi(y) \qquad (9)$$

And the three 2-D wavelets are defined as,

$$\Psi''(x, y) = \Phi(x)\ \psi(y) \qquad (10)$$

$$\Psi v(x, y) = \psi(y)\ \Phi(x) \qquad (11)$$

$$\Psi_D(x, y) = \psi(x)\ \psi(y) \qquad (12)$$

And the corresponding 2-D image decomposition and Reconstruction. A 2-D biorthogonal Multiresolution analysis can be obtained by means of a tensor product, from two one-dimensional (1-D) biorthogonal Multiresolution analyses, which obtains the scaling functions Ø and Ψ[1,2],

$$\Phi_{j,k,m}(x) = 2^{-j}\ \Phi\ (2^{-j}\ x-k)\ \Phi\ (2^{-j}\ y-m) \qquad (13)$$

$$\Psi^{H}_{j,k,m}(x) = 2^{-j}\ \Phi\ (2^{-j}\ x-k)\ \psi\ (2^{-j}\ y-m) \qquad (14)$$

$$\Psi^{V}_{j,k,m}(x) = 2^{-j}\ \psi\ (2^{-j}\ x-k)\ \Phi\ (2^{-j}\ y-m) \qquad (15)$$

$$\Psi^{D}_{j,k,m}(x) = 2^{-j}\ \psi\ (2^{-j}\ x-k)\ \psi\ (2^{-j}\ y-m) \qquad (16)$$

The nth Multiresolution analysis space can be projected into subspace images and which is given as,

$$f(x, y), \Phi_{N,k,m}(x,y) = \iint_{-\infty}^{\infty} f(x,y)\ \phi N,k,m(x,y)\,dxdy \qquad (17)$$

The original image can then be reconstructed by summing all of the subspace images, i.e

$$F(x, y) = \sum_{j=0}^{N} Fj(x, y) \tag{18}$$

We can be expressed as the following polynomial weighting formula to merge the corresponding points in each image [1],

$M_{\epsilon j}(I_j(x,y), r_j(x,y) = c_{j,0}(x,y) + c_{j,11}(x,y)I_j(x,y) + C_{j,12}(x,y)r_j(x,y) + C_{j,21}(x,y)I_j^2(x,y) + C_{j,22}(x,y)I_j^2(x,y) + C_{j,23}(x,y)I_j(x,y)\, r_j(x,y) + \text{Higher order terms}(x,y)$

5. BLENDING

The images to be joined overlap so that it is possible to compute the gray level value of points within a transition zone as a weighted average of the corresponding points in each image [3]. Suppose that one image, *Fl*(*i*), is on the left and the other, *Fr*(*i*), is on the right, and that the images are to be blended at a point ˆ*i* (expressed in one dimension to simplify notation). Let *Hl* (*i*) be a weighting function which decreases monotonically from left to right and let *Hr* (*i*) = 1 – *Hl* (*i*).

Then, the blended image *F* is given by $F(i) = Hl(i - \hat{i})\, Fl(i) + Hr(i - \hat{i})\, Fr(i)$[3]. (19)

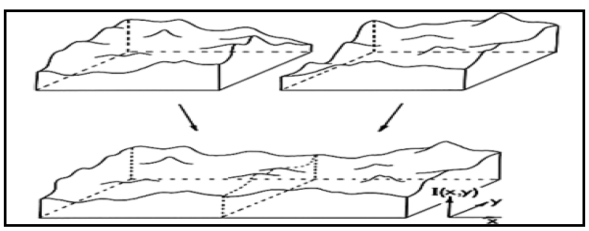


Fig. 7 A pair of images may be represented as a pair of surfaces above the (*x*, *y*) plane

It is clear that with an appropriate choice of H, the weighted average technique will result in a transition which is smooth. However, this alone does not ensure that the location of the boundary will be invisible. Let *T* be the width of a transition zone over which *Hl* changes from 1 to 0. If *T* is small compared to image features, then the boundary may still appear as a step in image gray level, albeit a somewhat blurred step. If, on the other hand, T is large compared to image features, features from both images may appear superimposed within the transition zone, as in a photographic double exposure. The size of the transition Zone, relative to the size of image features, plays a critical role in image blending. To eliminate a visible edge the transition width should be at least comparable in size to the largest prominent features in the image [3].

To minimize the image value variation, we impose a constraint that allows the pixel values of a blended image to be as close as possible to the corresponding pixel values of the to-be-combined images. To minimize the first derivative variation we impose the constraint that requires the first derivative of the mosaic images to consistently agree with that of the to-be-combined image. We formulate our energy functional at scale 2^j as,

$$E^*_{blend}(f_j, lj, rj) = \int_{-\epsilon j}^{\epsilon j} Eblend(f_j, lj, rj)\, dx$$

$$\int_{-\epsilon j}^{\epsilon j} Eimage(f_j, lj, rj) + \lambda^2_j E_{deriv}(f_j, lj, rj)\, dx \tag{20}$$

In general, the left side of a mosaic image should be similar to the left to-be-combined image, and the right side to the right To-be-combined image and is define as [1],

$$P^{l,\epsilon j} l_j(x,y) = \begin{cases} L_j(x,y) & \text{if } x < -\epsilon j \\ P\epsilon_j(x,y)\, l_j(x,y), & \text{if } -\epsilon j = x = \epsilon j \\ 0, & \text{if } x > \epsilon j \end{cases} \tag{21}$$

Then, we can write the term of the image value variation

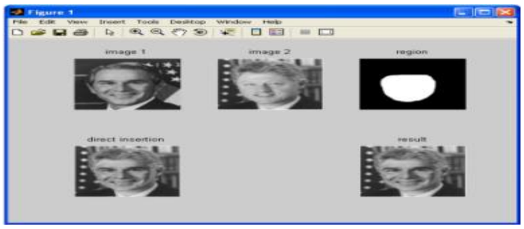
$$E_{\text{image}}(f_j, l_j, r_j) = \left[P^{l,\epsilon_j} f_j(x,y) - P^{l,\epsilon_j} l_j(x,y)\right]^2 + \left[P^{r,\epsilon_j} f_j(x,y) - P^{r,\epsilon_j} r_j(x,y)\right]^2 \tag{22}$$

6. RESULTS AND DISCUSSIONS

This section gives brief idea about what how mosaicing can be practically implemented. For this the image used is having the dimensions of 512x512. Due to its size it is decomposable unto 2^ (9) (9-1) =8 stages. At first the image is decomposed into two stages with coefficient from all the stages are saved. Then designing the mask, we have selected the mask size same as image size.

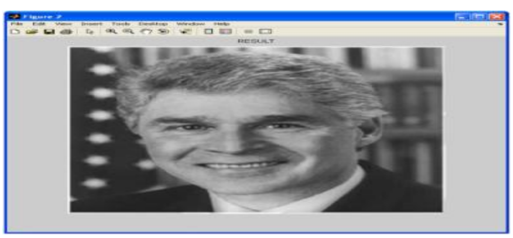
Now a mask has to be designed as per the requirement. In this case, we want to place the face of image shown in Fig. 8 onto the face of some other person. So, we have designed a mask which has matrix in which all the coefficients are 1(white) for the face region. Then the original image is multiplied with the mask to obtain the image in result (i.e. result B).

Now we take other image from which we just have to extract the face of the person which in this case is Mr. George Bush. So, while designing the mask, we need to take complement of the mask that design for previous section. Actual Mosaicing is done with result images from stages one (result A) and two (result B) which gives final result image.



(a)Mask design

(b)Combined image



(c) Multi-resolved image

Fig. 8 Image mosaicing using multiresolution.

7. CONCLUSIONS

The Laplacian pyramid using Gaussian pyramid have been discussed which are used as tools for image mosaic. A binary mask image is used to simplify precise shape mosaic. Special mosaic effects can also be achieved by combining components with specific frequency.

REFERENCES

[1] Ming Shing Su, Wen-Liang Hwang and Kuo-Young Cheng, “Analysis of Image Mosaics”, ACM Transactions, July 2004.

[2] Chiou Ting Hsu and Ja-Ling Wu, “Multiresolution Mosaic”, IEEE Transactions, Nov 1996.

[3] P.J.Burt and E.H.Adelson, “A Multiresolution Spine with Application to Image Mosaics”, ACM Transactions, Nov1983.

.[4] J.Burt and E.H.Adelson, “The Laplacian Pyramid as a Compact Image Code”, ACM Transactions, April 1983.

[5] Minh N. Doy and Martin Vetterliyx, “Frame reconstruction of the Laplacian pyramid”, IEEE Transactions, Nov 2001.

[6] Nadege Rebiere, Marie-Flavie, Auclair-Fortier, “Image Mosaicing using Local Optical Flow Registration”, IEEE Transactions, August 2008.

[7] Xianyong Fang, Bin Luo, Jin Tang, Haifeng Zhao, Biao He, Hao Wu, “Registration of blurred images for image mosaic”, IEEE Transactions, September 2009.

[8] Ruwei Luo, Xiaowei Chen, “Registration of blurred images for image mosaic”, IEEE Transactions, September 2009.

[9] Guosheng Yang, Huanlong Zhang, YuLin Yang, “Study of Image Mosaic Based on the Method of Finite Difference”, IEEE Transactions, Augest 2008.