

# 基于复值小波分解的图象拼合\*

徐丹<sup>1</sup> 鲍歌<sup>2</sup> 石教英<sup>1</sup>

<sup>1</sup>(浙江大学CAD&CG国家重点实验室 杭州 310027)

<sup>2</sup>(香港理工大学电子计算系 香港)

E-mail: csdanxu@comp.polyu.edu.hk

**摘要** 图象拼合是一种基于图象的场景编码方法,它被很多基于图象的绘制(IBR(image based rendering))系统采用,用来建立复杂的虚拟场景表示(例如,360°球面和柱面全景图、环境映照及高分辨率图象等)。基于复值小波多分辨率分解(Complex Wavelet Multiresolution Decomposition)提出了一种有效的图象拼合方法,它能同时地、逐步求精地对图象进行匹配和整合。首先,采用复值小波变换不仅可以保证全局优化的结果,还能够满足图象整合的规模不变和平移不变性。其次,基于多分辨率分析可以实现由粗到精的图象匹配和整合,从而使系统与传统的图象整合方法相比具有较高的性能。另外,通过改进图象整合算法中相似距离的测量,降低了计算复杂度。最后,系统可以直接拼合从数码相机捕获的图象,而无需知道相机运动及其他内部参数。

**关键词** 图象拼合,复值小波变换,图象匹配,图象整合,基于图象的绘制(IBR)。

**中图分类号** TP391

基于图象的绘制(IBR(image based rendering))是虚拟现实应用中一项引人注目的新技术。它不依赖于几何模型,而是利用预先获得的一组图象对环境进行编码,并通过适当地组合这些图象生成位于不同视点的新视图,最终实现虚拟环境的完全漫游。

图象拼合是指,将一个场景的各个相重叠的子画面对齐,并拼合成一个较大画面的表示绘制方法。图象拼合方法已成为许多IBR绘制系统中的一项基本技术,用于虚拟现实应用中高分辨率的及宽视场的图象生成,如全景图象。图象拼合方法的核心问题是图象整合(Image Registration),即将图象间相重叠的部分对齐所需的变换,其中涉及到的一个问题是寻找图象间的对应匹配。

寻找对应匹配最简单的方法是手工指定4个对应点;一般的拼合算法依赖于图象间的相互关系(Cross-Correlation)<sup>[1,2]</sup>。这些方法大都采用寻找局部对应或抽取重要特征的方法来进行匹配,很多都要求校正相机,而且只能得到局部优解,同时还面临特征附近的噪声等问题。基于相关(Phase-Based Correlation)的方法利用Fourier变换虽然可提供全局解,然而,正如我们所知,Fourier变换具有很差的空间变比性,与之相反,小波分析在信号处理中具有很强的频域和空域的局部性和规模不变性。更进一步,如果采用复值小波,还可保证平移不变性,并且使插值计算较为方便。

图象拼合算法所面临的另一个问题是怎样依据图象的对应关系来寻找整合变换。有的文章把这个问题独立出来讨论,假设已经知道图象间点的对应关系<sup>[3]</sup>,但是,直接从原图拟合变换比较复杂,问题规模大,性能低。相比之下,层次算法可快速得到拟合结果,还能逐层精化。<sup>[4-6]</sup>本文提出的方法是利用小波分解的特点把匹配和变换拟合二者结合起来同时考虑。

在本文所提出的方法中,首先,输入图象经复值小波分解表示成多分辨率的层次结构。接着,本文提出一种基于复值小波的层次匹配和整合算法,算法在大多数情况下能自动完成,并且

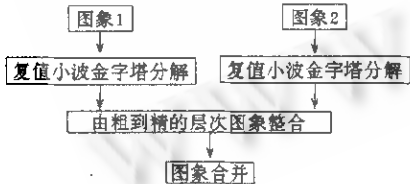


图1 两幅图象拼合的算法流程

\* 本文研究得到国家自然科学基金资助。作者徐丹,女,1968年生,讲师,主要研究领域为计算机图形学,虚拟现实,多媒体。鲍歌,1959年生,博士,副教授,博士生导师,主要研究领域为图形,图象编码,图象绘制,小波分析,虚拟现实。石教英,1937年生,教授,博士生导师,主要研究领域为计算机图形学,虚拟现实,体系结构。

本文通讯联系人:徐丹,香港,香港理工大学电子计算系

本文1998-02-28收到原稿,1998-04-20收到修改稿

不需要知道任何相机参数(如焦距等),另外,该算法可以得到全局解,并可根椐统计结果由粗到精的逐步优化.最后,经过两两整合的图象对可以拼合在一个较大的平面上,或者重新投影到柱面或球面上,构成全景图.图1描述了两幅图象拼合的算法结构.

### 1 复值小波分解表示

众所周知,小波分解是一项主要用于信号分解的技术.<sup>[7,8]</sup>一般地,如果已知图象  $f(x,y)$  的离散表示,实值小波金字塔分解是通过计算每一层上4个矩阵:  $A_j, D_{j,p}, p=1,2,3$  的滤波系数构成的,如图2(a)所示.

不同的小波有各自的特点.图象整合技术往往要求小波分解满足规模、平移及旋转不变性.但是,通常的实值小波分解不能满足平移和旋转不变性.<sup>[8]</sup>考虑到在采用图象拼合技术的 IBR 应用中往往待整合的两幅图象或是同视点,或是视点仅沿一个基轴平移,并只有很小的旋转,因此,可集中考虑解决平移问题.在此前提下,复值小波是较合适的表示,因为空间域的平移可表示为复值相域的旋转.而且这种表示还便于对小波变换插值.图2(b)说明了复值小波分解的过程,其中  $\{A_j^*, D_{j,p}^*, p=1,2,3\}$  是系数矩阵集  $\{A_j, D_{j,p}, p=1,2,3\}$  的共轭镜像.这两部分都包括了重要的图象信息.

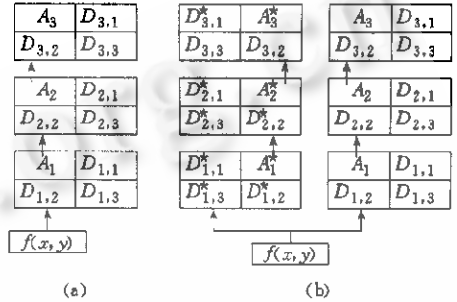


图2 小波多分辨率分解的实值及复值表示

这里,我们的算法采用特定的复小波;Magarey & Kingsbury 复小波.<sup>[7]</sup>相应的小波和规模滤波器  $g, h$  定义为 Gabor 滤波<sup>[9]</sup>

$$g_k = w(k+0.5)e^{i\pi(A+0.5)k}, \quad k = -N, -N-1, \dots, N-1$$

$$h_k = \tilde{w}(k+0.5)e^{i\pi(A+0.5)k}$$

其中  $w, \tilde{w}$  是一对由 Gaussian 窗口构成的实值低通窗口

$$w_1(k) = \exp\left(-\frac{k^2}{2\sigma_1^2}\right), \quad \tilde{w}_1(k) = \exp\left(-\frac{k^2}{2\tilde{\sigma}_1^2}\right)$$

最紧凑的窗口长度是4,即  $N=2$ .一种可选的  $g$  和  $h$  是 MKC-4

$$h = [1-i \quad 4-i \quad 4+i \quad 1+i]/10, \quad g = [-1-2i \quad 5+2i \quad -5+2i \quad 1-2i]/14$$

为了保证精度,需要估计非整数下标的系数值.可插值性是复小波的一个优点.子采样的复值小波系数可以通过调制一个低通插值核到带通滤波器的频率中心,并与调制核作卷积来求得.

$$D_{j,p}(k+u, l+v) \approx D_{j,p}(k, l) e^{i \cdot a_{j,p} \cdot (u,v)}$$

### 2 层次的图象匹配和整合算法

一般地,图象整合问题可以形式化定义为  $IR = (G, M, I_1, I_2)$ , 它满足

$$I_2(k', l') = I_1(G(M_1(k, l))),$$

其中  $G$  是一个平面投影变换,完成从图象  $I_1$  到图象  $I_2$  的变换,  $M$  是变换参数矢量:  $M = (m_0, m_1, \dots, m_7, m_8)$ . 如果用齐次坐标表示,  $G$  可以写为

$$\begin{pmatrix} u \\ v \\ \mu \end{pmatrix} = \begin{bmatrix} m_0 & m_1 & m_2 \\ m_3 & m_4 & m_5 \\ m_6 & m_7 & m_8 \end{bmatrix} \cdot \begin{pmatrix} k \\ l \\ 1 \end{pmatrix}, \quad \begin{pmatrix} k' \\ l' \\ 1 \end{pmatrix} = \begin{pmatrix} u/\mu \\ v/\mu \\ 1 \end{pmatrix}$$

其中  $(u, v, \mu)$  是齐次坐标点,它对应的图象笛卡尔坐标是:  $k' = u/\mu, l' = v/\mu$ . 这种表达可以满足大多数 IBR 的应用要求.当两个图象有共同的视点时,只需要恢复8个系数( $m_8=1$ ),于是式(6)重写为

$$\begin{cases} k' = \frac{m_0k + m_1l + m_2}{m_3k + m_7l + 1} \\ l' = \frac{m_4k + m_5l + m_6}{m_3k + m_7l + 1} \end{cases}$$

#### 2.1 相似性测量

层次匹配和整合算法的第1个任务是寻找参考图象对之间的部分匹配.这里采用的方法与文献[10]相似,不同之

处是,我们根据小波分解后每个子域的物理特性重新定义了模板矢量(Pattern Vector).因为相似距离计算是匹配过程中最密集的计算,在每个象素处的执行次数可达几百次.这个定义的简化大大减少了相似距离的计算复杂性,提高了算法效率.记  $PA_{j,q}(k,l), q=1,2$  和  $PD_{j,p}(k,l), p=1,2,\dots,6$  分别为基域和细节域中某象素中心的模板矢量.

$$PA_{j,q}(k,l)=[A_j(k+r,l+c);(r,c)\in N] \quad q=1,2, \tag{8}$$

$$PD_{j,p}(k,l)=[B_{j,p}(k+r,l+c);(r,c)\in N_p] \quad p=1,2,\dots,6, \tag{9}$$

其中 
$$B_{j,p}(k,l)=\frac{D_{j,p}(k,l)}{|A_j(k,l)|} \quad p=1,2,\dots,6, \tag{10}$$

称为小波分析的规范化细节域.  $N=\{(0,0),(-0.5,-0.5),(-0.5,0.5),(0.5,-0.5),(0.5,0.5)\}$ ,  $N_1,N_4=\{(0,0),(-0.5,0),(0.5,0)\}$ ,  $N_2,N_5=\{(0,0),(0,0.5),(0,-0.5)\}$ ,  $N_3,N_6=N$  是象素中心的小邻域.模板矢量的定义考虑到局部偏差的连贯性,即与该象素相邻的点应具有与该象素相同的偏移量.不同邻域的模板矢量定义针对不同频率带通滤波器的特性.例如,在水平低通垂直高通频率子域,将邻域定义为  $N_1,N_4$ ;而在垂直低通水平高通的子域,邻域定义为  $N_2,N_5$ .顶点的值可用插值公式(4)计算,其中  $\Omega_{j,q}$  和  $\tilde{\Omega}_{j,p}$  表示在  $j$  层上对子域  $p,q$  沿  $x$  和  $y$  方向的频率调制

$$(\tilde{w}_j, w_j), (w_j, \tilde{w}_j), (w_j, w_j), (-\tilde{w}_j, w_j), (-w_j, \tilde{w}_j), (-w_j, w_j) \quad \text{for } p=1,2,\dots,6, \\ (\tilde{w}_j, \tilde{w}_j), (-\tilde{w}_j, \tilde{w}_j) \quad \text{for } q=1,2$$

其中  $w_j$  和  $\tilde{w}_j$  满足

$$w_j = \frac{5\pi}{6}, \tilde{w}_j = \frac{\pi}{6}, \tag{11}$$

$$w_j = 3\tilde{w}_j, w_j = \frac{w_{j-1}}{2}, \tag{12}$$

两个对应象素  $(k,l)$  和  $(k',l')$  间的相似距离定义为

$$S_j((k,l),(k',l')) = \sum_{q=1}^2 SA_{j,q}((k,l),(k',l')) \sum_{p=1}^6 SD_{j,p}((k,l),(k',l')), \tag{13}$$

其中  $SA_{j,q}, q=1,2$  和  $SD_{j,p}, p=1,2,\dots,6$  表示各子域的相似距离,定义为

$$SA_{j,q}((k,l),(k',l')) = 1 - \frac{|PA_{j,q}(k,l) \cdot \overline{PA_{j,q}(k',l')}|}{|PA_{j,q}(k,l)| |PA_{j,q}(k',l')|}, \quad q=1,2, \tag{14}$$

$$SD_{j,p}((k,l),(k',l')) = |PD_{j,p}(k,l) - PD_{j,p}(k',l')|. \quad p=1,2,\dots,6. \tag{15}$$

### 2.2 对应匹配

按照前面的相似距离定义,对应匹配算法在大多数情况下可以自动完成.如果存在大的偏移,需要一些交互操作.匹配算法从两参考图象的最粗分辨率层开始,由里向外螺旋式进行.首先,利用式(13)寻找图象1的中心区域在图象2中的对应区域.其次,偏差值由中心向边缘方向一圈圈传递,对每一圈都用式(13)再作偏差微调.这个过程一直进行到对应边缘停止.于是得到该层的对应匹配关系.最后,利用前面两步所得的对应关系,可以拟合出一个较粗层次上的整合变换,这个变换又可用于较精细的下一层,以求得其初始对应关系.具体算法将在下节讨论.

应注意到图象整合与计算机视觉中的立体匹配问题<sup>[10]</sup>是不太相同的.这里考虑的参考图象间偏差较大,因而不仅仅是简单的完全匹配,需要考虑部分的对应匹配问题,即要标记匹配区域的边界.

### 2.3 层次图象整合

层次匹配和整合算法的第2个任务是拟合从图象  $I_1$  到图象  $I_2$  的整合变换.显然,这是一个非线性拟合问题.

由公式(7)可导出误差定义

$$e_i = \|(x'_i, y'_i) - G(M_j(x_i, y_i))\|^2, \\ \min_M E = \sum_i e_i. \tag{16}$$

式(16)中  $\|\cdot\|$  表示求欧几里德模数.为了最小化误差和,我们采用 Levenberg-Marquardt 算法求解,其特点是可在很少的迭代次数后收敛.<sup>[11]</sup>拟合结果根据式(13)可以由粗到精统计地优化.

**算法.** ① 从最初层次开始,用上节介绍的螺旋匹配算法找到最初层上的对应关系:  $x'_i = \delta x_i + x_i, y'_i = \delta y_i + y_i$ , 并由此得到一个初始的整合变换估计.

② 用 Levenberg-Marquardt 非线性最小算法拟合全局变换式(7),使其满足式(16).

③ 把第②步得到的全局变换用于下一个较精确的层,获得该层初始的匹配.这些对应接着用式(13)进行调整,以得到更优化的式(7)的解.

④ 重复迭代过程②、③,直到误差小于指定的阈值,或达到一定的迭代次数.

获得全局变换式(7)之后,就可以完成两幅参考图象  $I_1$  和  $I_2$  的混合. 为了减少视觉走样,给靠近图象中心的象素分配较大的权重,而靠边缘的象素分配给较小的权重. 这样,中心象素对组合的结果就有较大的贡献. 权重函数可表示为双线性函数

$$W(x,y)=W(x) \cdot w(y), \tag{17}$$

其中  $w$  是一个帽函数,它在图象边缘部分降为零. 图 3 是对办公室的两幅图象进行整合的结果,其中(a)、(b)是两幅参考图象,(c)是它们的整合结果.

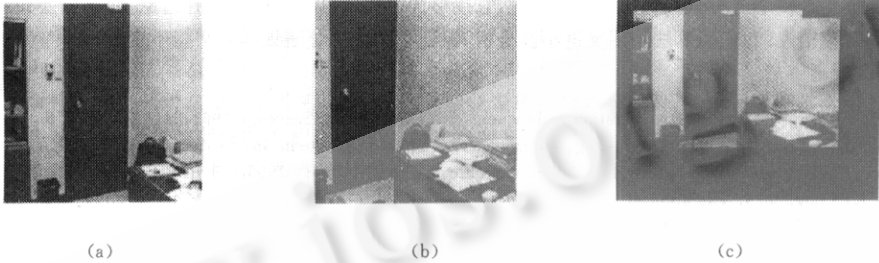


图 3 两幅图象整合

### 3 拼合图象的生成

这里,我们用一组整合变换来表达图象拼合,每一个变换对应一个输入序列中的图象,并表示由该图象象素到视域方向的变换. 这里我们考虑基于平面投影的图象拼合和基于圆柱面的图象拼合.

基于平面投影的图象拼合是最简单的拼合形式,利用前面讨论的匹配和整合算法可以直接生成. 由于图象从同一视点拍摄,我们只需恢复式(7)中的 8 个参数,并用迭代算法不断更新变换矢量  $M \leftarrow M + \delta m$ . 图 4 是一个简单的平面拼合图的例子. 如要拼合来自不同视点的图象,需要替代式(7)中的条件  $m_8=1$  为  $m_6^2+m_7^2+m_8^2=1$ ,详细讨论见文献 [12,13].



图 4 基于平面投影的多幅图象拼合



图 5 图象拼合的圆柱面表示

为了生成基于圆柱面的拼合图,每一个投影图象都需首先变换到圆柱面坐标. 从世界坐标点  $p=(x,y,z)$  到 2D 圆柱面坐标  $(\theta,\tau)$  的映射为

$$\theta=\tan^{-1}(x/z),\tau=y/\sqrt{x^2+z^2}, \tag{18}$$

其中  $\theta$  是水平摇动的角度, $\tau$  是扫描线. 然后,整个匹配整合和混合算法只需考虑纯粹的平移变换,即只需估计移动矢量中的  $m_5$  和  $m_6$ .

上述变换实现时要求知道焦距. 必要时,它可以从图象整合算法 8 参数估计结果中恢复,具体讨论可参阅文献 [12]. 图 5 是基于圆柱面的图象拼合例子.

### 4 实验结果

我们将上述算法用于一组由数码相机拍摄的图象数据,这些相片拍自香港理工大学校园. 从图 3 和图 4 的结果来看,拼合基本是无缝的. 图 5 是将图 4 的拼合图重新投影到圆柱表面上的结果.

考虑到相似距离的计算是整个匹配和整合算法中最密集的计算,每个象素处可多达几百次. 因此,我们在相似测量中简化了相似距离的计算,与文献[9]相比较,每个象素处少了  $N \times 16$  次加法操作和  $N \times 8$  次乘法操作. 其中, $N$  代表一个象素上相似距离的计算次数. 这一简化特别有利于加快高分辨率层上的处理速度.

为了测量算法偏差,我们定义了代价函数

$$err = \sum_i \| I(G(M; (x_i, y_i))) - I'(x'_i, y'_i) \| ^2, \quad (19)$$

其中  $I$  和  $I'$  表示象素的亮度值, 代价函数累积所有对应象素的亮度差. 事实上, 如果选择一个视觉容忍值  $T$ , 可只累计大于  $T$  的亮度偏差. 在我们的实验中, 当选择  $T=10$  时, 式(19)的误差均值不到 0.5 (对于 256 的灰度级).

## 5 总结

图象拼合是在基于图象的虚拟现实系统中建立复杂场景表示的一种基本方法. 本文所讨论的图象拼合方法为生成拼合图和全景图提供了一种全局优化的求解途径, 它能够同时完成自动匹配和整合图象, 而不需要知道任何相机参数以及深度和对应信息. 我们的算法利用复值小波分析, 保证了图象拼合是规模和平移不变的.

## 参考文献

- 1 Brown L G. A survey of image registration techniques. *ACM Computing Surveys*, 1992, 24(4): 325~376
- 2 McMillan L, Bishop G. Plenoptic modeling: an image-based rendering system. In: *Proceedings of SIGGRAPH '95, Computer Graphics, Annual Conference Series*, Los Angeles, California: ACM Press, 1995. 39~46
- 3 Szeliski R, Coughlan J. Hierarchical spline-based image registration. In: *Proceedings of IEEE Computer Vision and Pattern Recognition (CVPR '94)*, Los Angeles, California: IEEE Computer Society Press, 1994. 194~201
- 4 Mallat S. Zero-crossings of a wavelet transform. *IEEE Transactions on Information Theory*, 1991, 37(4): 1019~1033
- 5 Olivo J C, Deubler J, Boulin C. Automatic registration of images by a wavelet-based multiresolution approach. In: Andrew F Laine, Michael A Unser, Mladen V Wickerhauser eds. *Proceedings of SPIE (the international society for optical engineering), Wavelet Applications in Signal and Image Processing II*. SPIE Press, 1995, 2569: 234~244
- 6 Turcajova R, Kautsky J. A hierarchical multiresolution technique for image registration. In: Michael A Unser, Akram Aldroubi, Andrew F Laine eds. *Proceedings of SPIE (the international society for optical engineering), Wavelet Application in Signal and Image Processing N*. Denver, 1996, 2825: 686~696
- 7 Magarey J, Kingsbury N G. Motion estimation using complex wavelets. Technical Report, Cambridge University, 1995
- 8 Pan H P. Uniform full-information image matching using complex conjugate wavelet pyramids. In: *Proceedings of 18th ISPRS Congress (international society for photogrammetry and remote sensing)*, Vol. 31. Part of International Archives of Photogrammetry and Remote Sensing, Vienna, Austria, ISPRS Press, 1996
- 9 Gabor D. Theory of communication. *Journal of the IEE*, 1946, 9(3): 429~457
- 10 Pan H P. General stereo image matching using symmetric complex wavelets. In: Michael A Unser, Akram Aldroubi, Andrew F Laine eds. *Proceedings of SPIE (the international society for optical engineering), Wavelet Applications in Signal and Image Processing N*, Denver, 1996, 2825: 697~720
- 11 Press W, Flannery B, Teukolsky S *et al.* *Numerical Recipes in C: The Art of Scientific Computing*. 2nd edition, England: Cambridge University Press, 1992
- 12 Szeliski R. Video mosaics for virtual environments. *Computer Graphics and Applications*, Mar. 1996. 22~30
- 13 Szeliski R, Shum H Y. Creating full view panoramas image mosaics and environment maps. In: *Proceedings of SIGGRAPH '97, Computer Graphics, Annual Conference Series*. Los Angeles, California: ACM Press, 1997. 251~258

## Image Mosaics Based on Complex Wavelet Decomposition

XU Dan<sup>1</sup> BAO Ge<sup>2</sup> SHI Jiao-ying<sup>1</sup>

<sup>1</sup>(State Key Laboratory of CAD & CG Zhejiang University Hangzhou 310027)

<sup>2</sup>(Department of Computing The Polytechnic University of Hong Kong Hong Kong)

**Abstract** Image mosaics is one of the scene encoding approaches and is very popular among many IBR (image based rendering) systems in creating complex virtual environment based on photogeometrics, for example, 360° sphere or cylindrical panoramas, environment maps as well as high resolution images. In this paper, the authors present a robust panoramic image mosaicing scheme which employs complex wavelet pyramid techniques. It addresses the problems of both image matching and registration automatically and simultaneously. Complex wavelet transform guarantees not only a global optimal solution, but also scale and translation invariance for image alignment. The results can be progressively refined on the multiresolution decomposition. This feature guarantees that the scheme has higher performance than the traditional mosaicing techniques. The simplification of similarity measure decreases the complexity of computing. Additionally, the scheme registers images taken directly from digital camera without knowing camera motion and any intrinsic parameters of camera.

**Key words** Image mosaics, complex wavelet transform, image matching, image registration, IBR (image based rendering).