

基于多尺度均值和小波变换的 Internet 图象 可分级压缩编码传输技术

马社祥 刘贵忠 曾召华

(西安交通大学信息与通信工程研究所, 西安 710049)

摘要 图象具有直观、具体和生动等特点,同时,包含十分丰富的信息而被人们所重视.然而,利用浏览器在 Internet 网上进行信息搜索时,其图象信息的传输消耗了大量的传输时间.针对该问题,提出了基于多尺度均值的图象分级技术,即按分块取均值生成不同空间分辨率的图象,并采用基于小波理论的图象分层压缩编码技术来分级压缩编码,从而进一步降低了比特数.传输时可按空间分辨率由低到高的次序进行,这样,不仅大大加速了信息访问速度,同时为用户提供了多种选择的灵活性.

关键词 图象分级 图象编码 多尺度均值 小波变换 国际互联网络 信息传输

中图分类号: TN919.81 **文献标识码:** A 1006-8961(2000)11-0942-06

Image Coding and Transport Technique for Internet Based on Multiscale Mean Values and Wavelet Transform

MA She-xiang, LIU Gui-zhong, ZENG Zhao-hua

(Institute of Information & Communication Engineering, Xi'an Jiaotong University, Xi'an 710049)

Abstract Images are lifelike, specific and directly perceived through the sense. They contain a great deal of information to be regarded. However, a lot of time is spent on transporting image-information, when you scan on Internet for information. This paper introduces a novel image scaling technique by multiscale mean values. The images of various spatial resolution are generated by the mean values of blocks in the original image. Combined with wavelet transform for image-layered coding gives high compression ratio and increase flexibility. The speed of information downloaded will be accelerated based on this scheme. It will also be convenience for users.

Keywords Image scaling, Image coding, Multiscale mean values, Wavelet transform, Internet, Information transport

0 引言

国际互联网(Internet)的发展,已经深入到当今世界的每一个角落,而且联网用户正在以指数级速度增长.它的发展使人类的工作、学习和生活方式发生了根本性的变化,而且使未来信息化社会已初见端倪.但由于用户剧增以及当前 Internet 自身的技术所限,目前的商用 Internet 已经十分拥挤,无法满足人们,特别是以拨号方式上网用户的需要.上网

用户都知道,图象以其包含的信息量大、直观而著称,因而易于从中获取大量的信息.然而,图象信息的下载需要消耗大量的时间,因此许多人都关闭了图象下载模式,当特别需要看图象时再下载.即使如此,下载时也许还需要等待相当长的时间,随着图象自上而下地显示出来,使用者才得到一幅完整的图象,在此之前使用者对图象的概貌无法了解.对用户来说,既希望看到图象的全貌,又要快速下载,这就成了需要解决的问题.要解决这个矛盾,在现有的条件下,只有减少图象的比特数:一是缩小图象的大

小, 即降低图象分辨率; 二是图象压缩. 由于在绝大多数情况下, 使用者只需了解图象概貌, 而不是图象的大量细节和修饰成分, 只有在特殊需要和必需时 (即对图象非常感兴趣) 才下载全部信息, 这说明上述两种方法是可行的, 因此需要一种新的图象编码、传输模式来满足实际应用的需要.

本文引入了基于块的多尺度均值 (Block-Based Multiscale Mean Values, 简称 BMMV) 技术, 来实现图象空间大小 (即分辨率) 的分级, 即首先将图象分成大小不同的若干个图象, 然后在小图象编码的基础上增加少量的编码就可得到大图象的编码, 即大小图象之间具有兼容性, 再结合利用零树小波 (EZW) 编码^[1] 技术来实现图象的质量和空间频率的分级压缩编码. 其与 DCT 相比, EZW 提供了非常高的压缩效率, 同时还具有可分级特性^[2].

1 图象的 BMMV 分级和 EZW 编码

1.1 BMMV 分级特性

图象虽然包含的信息量很大, 但也存在着很强的相关性. 也就是说, 图象的大部分点在一个相当大的邻域 (或块) 内, 各点具有强相关性, 即在该邻域内的图象强度变化不大, 可以用一个点的值来近似代替. 从频域角度来看, 即图象的低频成分占有图象的大部分能量, 而高频能量很少. 据统计^[3], 在进行 3 次小波分解后的 10 个频段中, 大小只是原图象 1/64 的低频 LL₃ 子图象的能量就占原图象能量的 90% 以上. 鉴于上述事实, 本文提出将原图象分成大小相等的小块, 每一块按强度求均值, 用一个均值来近似代替该小块各点的值, 其误差应该是很小的. 由这些均值就组成一个尺寸上较小的子图象, 然后再对原图象所剩余的差值图象, 再用更小的块并用前述方法来生成尺寸上介于二者之间的子图象. 如此进行若干步, 可得到一个子图象序列, 记为 $S(k)$, 最后剩余的差值图象为 $d(K)$. 在分级方面, 如果要存贮、传输和接收不同大小的图象, 只需将上述 $S(k)$ 和 $d(K)$ 依次组合即可.

为了说明 BMMV 方法, 可以先看一维曲线的一个实例. 假设原曲线 f 如图 1(a) 所示, 它含有 200 个点, 首先将其等距分割为 10 等份, 对应地有 10 个小区间 $\Delta\alpha_1, \Delta\alpha_2, \dots, \Delta\alpha_{10}$, 每个小区间上有 20 个点, 然后对 f 在每个小区间 $\Delta\alpha$ 上求均值, 并以此均值作为函数 f 在该区间上各点的近似值, 则得到阶梯

状的曲线 f_1 , 将其作为 f 的近似 (如图 1(b) 所示), 其误差曲线为 d_1 (如图 1(c) 所示), 即 $f_1 + d_1 = f$. 对于曲线 f_1 , 一方面, 虽然它也包含 200 个点, 但实际上只需存储 10 个值即可; 另一方面, 从视觉效果来看阶梯状的曲线 f_1 并不好看, 且和原曲线 f 相差较大, 因此, 就可用其 10 个均值重新组成一曲线 f_1 作为 f 的低分辨率近似 (如图 1(d) 所示). 从点数上来说, f_1 是 f 的 1/20, 图虽然小, 但比阶梯状的大图要好看, 更象原曲线, 所以称其为低分辨率图. 如果想用少量的数据了解原曲线图 f 的概貌, 那 f_1 是最好的了, 而且由于 $f_1 + d_1 = f$, 而 f_1 也基本上代表了 f_1 , 所以由 f_1 和 d_1 也能合成 f , 且仍记 $f_1 + d_1 = f$, 这里的“+”表示依 d_1 的元素数将 f_1 按阶梯状放大, 再和 d_1 相加 (下文的“+”意义相同).

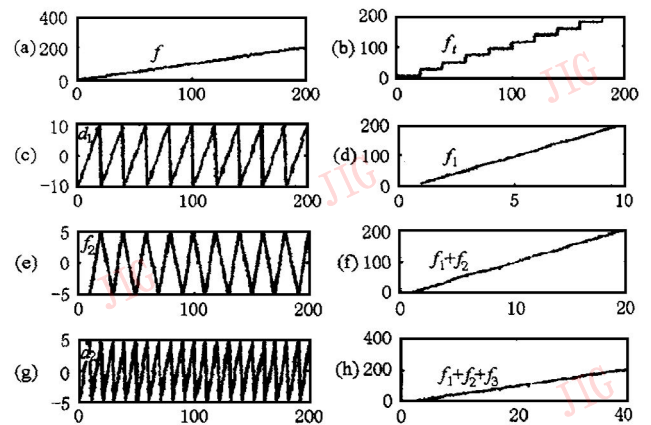


图 1

如果希望得到较高分辨率曲线图形 (即比 f_1 的点再多一些), 由于 $f_1 + d_1 = f$, 因此对于曲线 d_1 (即图 1(c)), 若将其等距分割为 20 等份, 对应地也有 20 个小区间 $\Delta b_1, \Delta b_2, \dots, \Delta b_{20}$, 则每个小区间上有 10 个点 (即原来的一个小区间 $\Delta\alpha$ 被细分为两个小区间), 然后对 d_1 在每个小区间 Δb_j 上求均值, 并按和前面相似的方法由其均值生成一个含 20 个点的曲线 f_2 (图 1(e)) 以及差值曲线 d_2 (图 1(g)), 即 $f_2 + d_2 = d_1$. 由 $f_1 + f_2$ 即可得到含 20 个点 (即是 f_1 的两倍) 的较高分辨率近似图形 (如图 1(f) 所示), 其误差为 d_2 , 即有 $f_1 + f_2 + d_2 = f$. 对于 d_2 , 可将上述小区间 Δb_j 再细分, 求得均值曲线 f_3 和差值曲线 d_3 , 由 $f_1 + f_2 + f_3$ 可得到更高分辨率的近似图形 (图 1(h)), 且 $f_1 + f_2 + f_3 + d_3 = f$. 如此进行若干步, 可得到 f 的不同分辨率的近似图形, 分别为 $f_1, f_1 + f_2, f_1 + f_2 + f_3, \dots$, 其分辨率越来越高 (点数越来越多), 误差分别为 d_1, d_2, d_3, \dots , 且误差越来越小.

对于一个二维图象 g , 和一维曲线情况相同, 假设有 $N \times N$ 个像素, 将其分割成大小相同的块, 假设块的大小为 $M \times M$, 然后对每一块求均值, 并由这些均值组成一个大小为 $\frac{N}{M} \times \frac{N}{M}$ 的小图象 $S(1)$, 作为 g 的低分辨率图象, 其差值图象记为 $d(1)$, 即有 $S(1) + d(1) = g$. 对于 $d(1)$ 按块的大小为 $\frac{M}{2} \times \frac{M}{2}$ 进行细分, 然后对每一块求均值, 并由这些均值组成一个大小为 $\frac{2N}{M} \times \frac{2N}{M}$ 的图象 $S(2)$, 其差值图象记为 $d(2)$, 即有 $S(2) + d(2) = d(1)$. 由 $S(1) + S(2)$ 可得到一个大小为 $\frac{2N}{M} \times \frac{2N}{M}$ 的较高分辨率图象, 且 $S(1) + S(2) + d(2) = g$. 对于 $d(2)$ 再按更小的块细分, 同样的方法可得到均值图象 $S(3)$ 和差值图象 $d(3)$, 即有

$S(3) + d(3) = d(2)$. 由 $S(1) + S(2) + S(3) + d(3) = g$. 如此进行若干步, 可得到的不同分辨率的近似图象分别为 $S(1), S(1) + S(2), S(1) + S(2) + S(3), \dots$, 其分辨率越来越高(图象越来越大), 误差分别为 $d(1), d(2), d(3), \dots$, 且误差越来越小.

若块的大小分别取 $8 \times 8, 4 \times 4, 2 \times 2$, 如果需要大小为原图象 $1/64$ 的图象, 则只取 $S(1)$ 即可; 如果需要大小为原图象 $1/16$ 的图象, 则取 $S(1) + S(2)$; 如果需要大小为原图象 $1/4$ 的图象, 则取 $S(1) + S(2) + S(3)$; 如果需要和原图象大小相同的图象, 则取 $S(1) + S(2) + S(3) + d(3)$. 这样就可以保证大小图象的兼容, 同时, 小图象能够很好地满足人们的视觉要求(如图2所示).

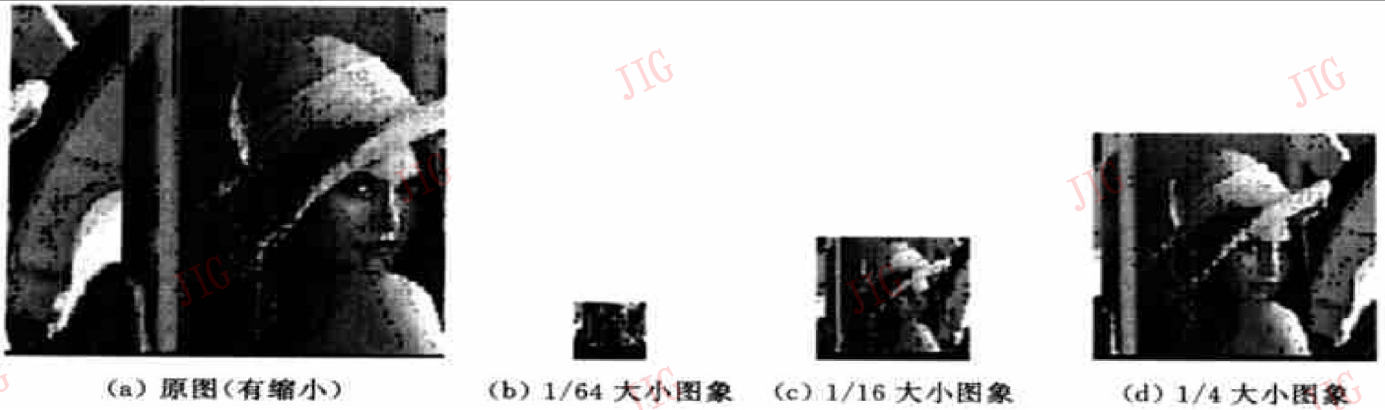


图2 基于BMMV的图象分级

1.2 EZW 编码

EZW 编码是一个简单的、被证明是非常有效的图象编码算法^[1]. 由这种算法得到的比特流中的比特是按其重要性排序的. 使用这种算法, 编码者能够在任一点结束编码, 这样就可使编码精确地达到一个目标比特率或目标失真率, 而这时仍能够确切地产生同样的图象. 这种算法的特点是绝对不要求训练, 不要求预先存储格式码书, 也不要求图象源的任何预先的知识.

EZW 编码基于以下3个主要思想, 即, (1) 利用小波变换^[4~7]在不同尺度间固有的自相似性来预测重要信息的位置; (2) 逐次逼近量化小波系数; (3) 使用自适应算术编码来实现无损数据压缩.

经过离散小波变换(即 Mallat 算法), 把一个图象分解成若干个级, 构成如图3所示的金字塔形结构. 若分成了 k 级, 则依子图象的幅值分布, 对于同一级子图象, 低频子图象 LL_j 最重要, 其次是 HL_j 与 LH_j , 高频子图象, 而 HH_j 不重要. 对于不同级来

说, 级高者重要, 级低者不重要. 所以, 子图象按其重要性总体的排序为 $LL_k, HL_k, LH_k, HH_k, HL_{k-1}, LH_{k-1}, HH_{k-1}, \dots, HL_1, LH_1, HH_1$ (如图3所示).

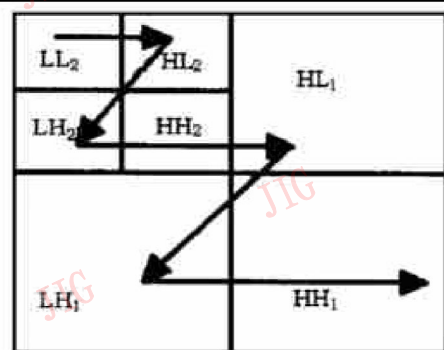


图3 小波分解两次及其子图重要次序

EZW 还引入了一个称为零树的数据结构, 即一个小波系数 x , 对于一个给定的门限 T , 如果 $|x| < T$, 则称小波系数 x 是不重要的. 如果一个小波系数在一个粗的尺度上, 关于给定的门限是不重要的, 且之后在较细尺度上, 对同样空间位置中的所有小波系数也关于给定的门限 T 是不重要的, 则称小波系

数形成了一个零树. 这时, 在粗的尺度上的那个小波系数称为母代小波系数, 它是树根, 在较细尺度上相应位置上的小波系数称为子代小波系数. 如果一个子代小波系数关于门限 T 是不重要的, 但它的子代小波系数中存在关于门限 T 是重要的, 则称这个小波系数是孤立零. 由此可见, 所有的小波系数只是下述 3 种情形之一: (1) 零树根 (ZTR), (2) 孤立零 (IZ); (3) 重要系数. 当然, 为了编码的需要还可把重要系数分为正重要系数 (POS) 和负重要系数 (NEG).

EZW 使用逐次逼近量化技术. 所谓逐次逼近量化就是逐次使用门限 $T_0, T_1, T_2, \dots, T_{N-1}$ 以决定重要系数, 其中门限序列的选取是 $T_i = T_{i-1}/2$, 而初始门限 T_0 的选取, 要使得对于所有系数 x_j 有 $|x_j| < 2T_0$. 为了计算方便, 一般取 $T_0 = M2^E$, 其中 E 是整数, M 是固定常数.

编码过程需要两个表, 一个称为主表, 用来存放还未确定是否是重要系数的小波系数坐标位置, 在每次量化过程中, 主表的初值为所有小波系数的坐标; 另一个称为辅表, 用来存放在逐次编码过程中小波系数的剩余值, 即剩余图, 其在初始化时的初值就为所有的小波系数. 在某步量化过程中, 若门限为 T , 则以 T 为阈值, 按图 3 所示的次序逐个扫描主表中的每一个坐标位置, 以及确定在辅表中相应位置的系数 x_j 的重要性. 若 x_j 关于 T 是正重要系数, 则用 POS 作为它的编码, 且编码后辅表中该系数的剩余值为 $x_j - T - T/2$; 若 x_j 关于 T 是负重要系数, 则用 NEG 作为它的编码, 且编码后, 辅表中该系数的剩余值为 $x_j + T + T/2$; 若 x_j 关于 T 是孤立零, 则用 IZ 作为它的编码, 且编码后辅表中该系数的值不变; 若 x_j 关于 T 是零树根, 则用 ZTR 作为它的编码, 编码后辅表中该系数及其所有子代小波系数结点的值不变, 且在本次扫描过程中, 主表中该结点的所有子代小波系数结点不必再扫描. 如此直到主表扫描完为止, 然后取门限 T 值为 $T/2$, 以及对主表初始化, 再逐个扫描编码. 重复上述编码过程, 直到满足要求为止. 数据传输时, 首先传输 T_0 , 然后按编码的先后次序传输所有的编码.

解码时, T_0 已知, 且记为 T , 按接收到码的先后次序逐个解码. 解码时也需要两个表, 一个是主表, 和编码时主表相同, 用于存放将要扫描的小波系数坐标, 以确定解码时的每一个码所对应的小波系数; 另一个表是辅表, 用来存放解码后小波系数的值, 初始化时, 辅表中各元素值为零. 对主表按图 3 所示的

扫描次序, 将接收到的码元逐个与主表中的坐标位置对应起来, 若码为 POS, 则在辅表中相应位置上加 $T + T/2$; 若码为 NEG, 则在辅表中相应位置上加 $-T - T/2$; 若码为 IZ, 则辅表在相应位置上的值不变 (即加 0); 若码为 ZTR, 则主表该位置以及其所有子代小波系数在本次的扫描中都已通过, 且在辅表中的所有这些结点的值不变 (即加 0). 如此逐个地将编码和主表中的坐标位置对应起来, 以确定辅表中的那些值作如何修改. 当主表扫描完一次时, 将门限值 T 修改为 $T/2$, 且对主表初始化, 重新开始扫描. 如此直到结束.

上述扫描过程中, 实际上主表每扫描完一次, 辅表也扫描了一次, 即辅表中每个元素的值都作了相应的修改, 在辅表中每次增加的这些值也组成了一个图, 称为重要图, 也简称位图, 这些位图的重要性依生成的次序递减, 将它们叠加起来, 即为解码后重构的小波系数. 随着位图数的增加, 剩余图各结点的值趋于零, 即位图数决定着重构图的质量, 因此也称位图为质量层.

2 图象可分级压缩传输与重构

2.1 基于 BMMV 的可分级压缩编码过程

图象首先经过 BMMV, 将原始图象分解成不同大小的子图 $S(1)$ 、 $S(2)$ 、 $S(3)$ 和 $S(4)$, 其中 $S(4)$ 就是差值图象 $d(3)$ (假设我们分解了 4 层, 实际上最多 4 层就够用了); 然后对 4 个子图分别进行离散小波变换 (记 DWT); 最后进行 EZW 编码, 输出比特流. 其编码过程如图 4 所示.

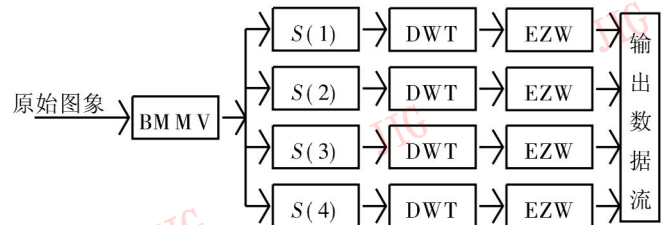


图 4 基于多尺度均值技术的零树小波压缩编码示意图

2.2 Internet 图象的传输和解码

由于子图 $S(1)$ 包含原图象中最重要的信息, 其次为 $S(2)$ 和 $S(3)$, 而 $S(4)$ 中包含原图象的信息已经很少. 所以它们的编码传输次序是:

$$S(1) > S(2) > S(3) > S(4)$$

对每一个 $S(i)$ 的 EZW 编码中, 由于幅值大的先编码, 幅值小的后编码, 而幅值大的信息比幅值小

的信息更重要,所以对每一个 $S(i)$ 的 EZW 编码按其编码的先后次序传输.

解码过程与编码过程相反,其解码过程如图 5 所示. 其中 EZW^{-1} 表示 EWT 编码的解码过程, DWT^{-1} 表示逆小波变换. 由于编解码过程中存在着误差,记 $\hat{S}(i)$ 是 $S(i)$ 的近似. 按传输的先后次序,如果只需要 $1/64$ 大小的图象,则只需下载第一部分中的部分数据,经过逆 EZW 和逆 DWT 即可;如果需要 $1/16$ 大小的图象,则下载前两部分中的部分数据,分别经过逆 EZW 和逆 DWT,之后再叠加可得;如果需要 $1/4$ 大小的图象,则下载前三部分中的部分数据,分别经过逆 EZW 和逆 DWT,之后再叠加可得;如果需要和原图象大小相同的图象,则下载 4 部分中的部分数据,分别经过逆 EZW 和逆 DWT,之后将它们叠加即可.

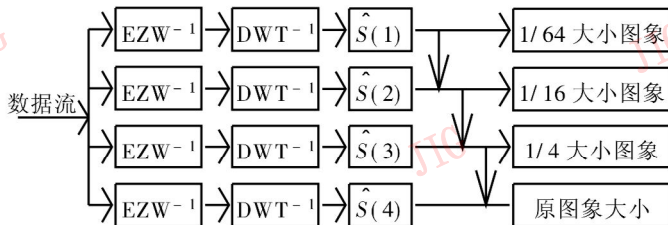


图 5 分级解码示意图

3 计算机仿真实验及讨论

为了验证本文提出的 Internet 图象分级压缩编

码、传输模式的有效性,因而在计算机上进行了仿真实验,实验中采用了双正交 B9/7 小波^[8],对标准灰度 Lena 图象(256× 256 象素大小,8bit 深度)做分级压缩编码、解码实验. 如图 6 所示,其中,(a)为原始图象,比特数为532 288bit;(b)为和原图象大小相同的重构图象,比特数为17 742bit,其压缩比为30,PSNR 值为32. 836 2(dB);(c)为 $1/4$ 大小的重构图象,比特数为12 288bit;(d)为 $1/16$ 大小的重构图象,比特数为8 192bit;(e)为 $1/64$ 大小的重构图象,比特数为4 096bit. 由于篇幅所限,其中(a)和(b)有缩小.

如从数据传输的时间考虑,若传输大小相同的图象则所需时间仅为常规方法的 $1/30$,若传输 $1/4$ 、 $1/16$ 、 $1/64$ 大小的图象,则所需时间分别为常规方法的 $1/42$ 、 $1/64$ 、 $1/128$. 由此可见,大大缩短了图象数据的传输时间. 同时,用户在得到小图象后,如果需要,则再增加少量的数据就可以得到较大的图象,由于它保证了图象之间的兼容,因而为用户提供了多种选择的灵活性. 这就为充分利用 Internet 网络资源,在有限时间条件下获取更多的有用信息提供了一个切实可行的方案. 同时从恢复图象的主观和客观标准来看,进一步证明了本文方法的可行性,为高效压缩图象数据提供了一个切实可行的方法.

从以上理论分析和仿真实验可看出,本方案简单可行、有效. 它为加速当今 Internet 信息下载,提高网络资源利用率,提供了一条新的有效途径.



图 6 Internet 图象分级压缩编码、传输的不同大小重构图象

参 考 文 献

- 1 Shapiro J M. Embedded image coding using zerotree of wavelet coefficients. IEEE Trans. Signal Processing, 1993, 41(12): 3445~ 3462.
- 2 Sodagar I, Lee H J, Hatrack P et al. Scalable wavelet coding for synthetic/natural hybrid images. IEEE Trans. Circuits System Video Technology, 1999, 9(2): 244~ 254.
- 3 张正阳. 高性能图象编码研究[博士论文]. 西安: 西安电子科技大学, 1999.
- 4 Daubechies I. Orthonormal bases of compactly supported wavelets. Commun. Pure and Appl. Math. 1988, XLI: 909~ 996.
- 5 Daubechies I. The wavelet transform, time-frequency localization and signal analysis. IEEE Trans. Information Theory, 1990, 36(5): 961~ 1005.

- 6 Burt P J. The laplacian pyramid as a compact image code. IEEE Trans. Communication, 1983, 31(4): 523~ 540.
- 7 刘贵忠, 邸双亮. 小波分析及其应用. 西安: 西安电子科技大学出版社, 1992.
- 8 Cohen A, Daubechies I, Feauveau J C. Biorthogonal bases of compacted supported wavelets. Commun. Pure and Appl. Math., 1992, XLV: 485~ 560.



刘贵忠 1962 年生, 现为西安交通大学电子与信息工程学院教授, 博士生导师. 主要研究领域为非平稳信号的表示和自适应信号处理等.



马社祥 1962 年生, 现为西安交通大学信息与通信专业博士生. 主要研究领域为视频压缩、多媒体通信、小波理论等.



曾召华 1972 年生, 现为西安交通大学信息与通信专业博士生. 主要研究领域为移动通信、信号处理、图象压缩编码等.

Intra3D

OPENGL 的对象化封装

- 分层设计, 每层少而精, 扩展灵活
- 提供 C++ 类库与 COM 库, 支持 VC++, VB, Delphi, Java, J++ 编程
- 免费软件, 所有源代码向用户开放

——适合各类图形编程人员

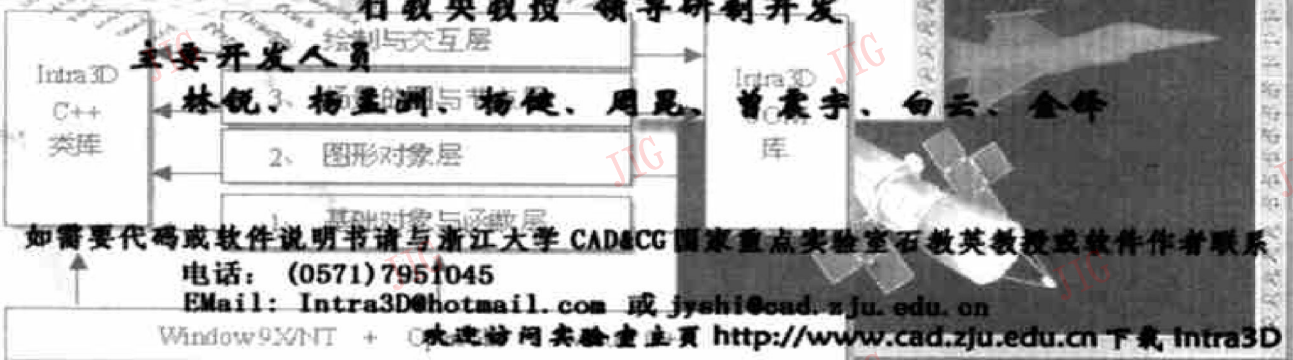
此软件包由

浙江大学 CAD&CG 国家重点实验室

石教英教授 领导研制开发

主要开发人员

林锐、杨孟洲、杨健、周昆、曾震宇、白云、金峰



如需要代码或软件说明书请与浙江大学 CAD&CG 国家重点实验室石教英教授或软件作者联系

电话: (0571)7951045

E-mail: Intra3D@hotmail.com 或 jyshi@cad.zju.edu.cn

Window9X/NT + 欢迎访问实验室主页 <http://www.cad.zju.edu.cn> 下载 Intra3D