不产生精度截断及数据膨胀的图像金字塔

谢耀华　汤晓安　孙茂印

(国防科技大学电子科学与工程学院, 长沙　410073)

摘　要　图像金字塔是处理和分析数字图像的重要工具。以图像金字塔为研究对象,旨在减小图像金字塔的存储量,同时解决恢复数据时的精度截断问题。从同一思想出发,提出了两种新的金字塔结构:类均值金字塔与类高斯金字塔;它们都是通过对子像素的简单运算来求得父像素,而这种运算不存在精度截断问题,也不需要额外的比特来保存小数。由于从父像素能够反推出它的子像素之一,所以,在存储金字塔时可以舍弃这部分子像素,从而使所需存储的像素数仅与原始图像相同。对精度、存储量以及构建速度的理论分析与实验结果表明,该金字塔结构具有优良的总体性能。

关键词　精度截断　数据膨胀　图像金字塔　均值金字塔　高斯金字塔

中图法分类号: TP391　　**文献标识码:** A　　**文章编号:** 1006-8961(2009)06-1070-05

Image Pyramid Structures without Truncation and Data Expansion

XIE Yao-hua, TANG Xiao-an, SUN Mao-yin

(*School of Electronic Science and Engineering*, *National University of Defense Technology*, *Changsha* 410073)

Abstract　Image pyramid is an important tool for processing and analyzing digital images. This study focuses on image pyramid, and finds out an approach to reduce the data volume of them and solve the problem of truncation in precision. Based on a same idea, we propose two kinds of image pyramid: the mean-like pyramid and the Gaussian-like pyramid. In them, father-pixels can be got by simple calculations of their son-pixels; there is no problem of truncation in such calculations, and no extra bits are needed to store decimals. Besides, as some son-pixels can be recovered from their father-pixels, they can be dropped when storing a pyramid. Thereby, the number of pixels need to be stored is equal to that of the original image only. It's shown that the proposed pyramid structures have good performance in general by theoretic analysis and experimental on precision, data volume and speed of creation.

Keywords　truncation, data expansion, image pyramid, mean pyramid, Gaussian pyramid

1　引　言

图像金字塔是多分辨率图像处理中的一项重要技术[1]。它实质上是根据原始图像构建出来的一个图像序列;序列中的每个图像称为一个层,它们尺寸递减并均为原始图像的一个低分辨率表示[2];金字塔的相邻层之间,分辨率一般相差两倍。不同应用领域中的金字塔结构各有特点;而本文所讨论的,其各层图像在视觉上均为原始图像的近似,或者能够快速复原成为原始图像的近似。图像金字塔已经在各种领域中得到成功的应用,比如:图像融合[3],图像编码[4]与渐进传输[5],海量图像/地形的实时显示[6]等等。

为了满足不同的应用需求,几十年来各国学者提出了多种图像金字塔结构[7-9]。最简单的一种是子抽样金字塔,每一层图像由其前一层沿行、列两个方向等间隔抽样得到;其优点是构建速度快,但生成

收稿日期:2007-11-27;**改回日期:**2008-04-07

第一作者简介:谢耀华(1979～　),男。国防科技大学信号与信息处理专业博士研究生。主要研究方向为遥感影像存储、管理与处理。E-mail: fjpnxyh2000@163.com

的图像混淆显著,像素对所在区域没有很好的代表性。均值金字塔则具有良好的视觉效果,运算量也较小。高斯金字塔也具有良好的视觉效果,但运算量略大。拉普拉斯金字塔由高斯金字塔派生而来,需要多读取部分数据才能够复原到高斯金字塔的效果。上述 4 种金字塔都比原始图像多出约 1/3 的像素,造成数据膨胀。Reduced-Sum 金字塔则解决了数据膨胀问题,它舍弃了部分像素,从而使金字塔中的像素数等于原始图像。但 Reduced-Sum 金字塔存在精度截断的问题,因为计算得到的低分辨率像素带有小数,这些小数在舍入过程中被丢弃了,使得原始图像无法精确恢复;否则,每个像素需要多加两个比特来储存小数。实际上,精度截断现象在现有各种金字塔普遍存在;只是像均值金字塔之类的冗余结构因为完整包含了原始图像,所以允许小数被舍弃。

在本文提出的金字塔结构中,父像素由它的某个子像素加上一个整数得到,此整数由该子像素的邻域像素确定;该子像素在存储时将被舍弃,而其邻域则保留。反之,被舍弃的子像素可由父像素减去该整数而得以精确恢复,不存在精度截断问题。这样,存储金字塔时可以舍弃约 1/4 的子像素,所需存储的像素数仅与原始图像相等。此外,因为父像素不含小数,每个像素又能节省两个比特。

2　本文提出的金字塔结构

设原始图像为 $I(u,v)$, $u,v=0,1,\cdots,N-1$,且 $N=2^M$。构建的金字塔均为 L 层;分辨率最高的为第 0 层;第 1 层分辨率为第 0 层的一半;依此类推,第 $L-1$ 层仅含一个像素。易知 $M=L-1$。图 1 为金字塔某一层的示意图;图中,每个方格代表一个像素,灰色方格则表示要被舍弃的像素。

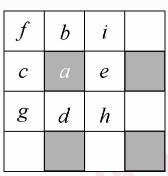


图 1　金字塔某一层的示意图

Fig. 1　Part of a level of a pyramid

以图 1 为例说明构建类均值金字塔的基本思想:首先利用像素 b,c,d 和 e 来估计要舍弃的像素 a;然后用估计值 $\hat{a}$ 与像素 b,c 和 f 求均值,得到父像素的估计值 $\hat{p}$;再将 $\hat{p}$ 减去 $\hat{a}$ 得到父像素 p 与被舍弃像素 a 之差的估计值 Δ;最后,将父像素 p 的修正为 $(a+\|\Delta\|)$。类高斯金字塔的基本思想与此相似,只是求 $\hat{p}$ 的方法类似于高斯滤波。

2.1　类均值金字塔结构

本金字塔结构是由均值金字塔改进而来。其构建规则如式(1)所示:

$$
\begin{aligned}
& I_0(u,v)=I(u,v);u,v=0,1,\cdots,N-1 \\
& I_l(u,v)=I_{l-1}(2u,2v)+\|\Delta_l(u,v)\| \\
& \Delta_l(u,v)=\hat{I}_l(u,v)-\hat{I}_{l-1}(2u,2v) \\
& \quad =\frac{1}{4}\cdot[\hat{I}_{l-1}(2u,2v)+I_{l-1}(2u,2v+1)+ \\
& \quad I_{l-1}(2u+1,2v)+I_{l-1}(2u+1,2v+1)]- \\
& \quad \hat{I}_{l-1}(2u,2v) \\
& \quad =\frac{1}{4}\cdot[I_{l-1}(2u,2v+1)+I_{l-1}(2u+1,2v)+ \\
& \quad I_{l-1}(2u+1,2v+1)]-\frac{3}{4}\cdot\hat{I}_{l-1}(2u,2v) \\
& \hat{I}_{l-1}(2u,2v)=\frac{1}{4}\cdot[I_{l-1}(2u-1,2v)+I_{l-1}(2u,2v-1)+ \\
& \quad I_{l-1}(2u,2v+1)+I_{l-1}(2u+1,2v)] \\
& \quad l=1,\cdots,L-1;u,v=0,1,\cdots,\frac{N}{2^l}-1
\end{aligned}
\tag{1}
$$

式中,$I_l(u,v)$ 为第 $l(l=0,1,\cdots,L-1)$ 层图像在点 (u,v) 处的灰度值;$\Delta_l(u,v)$ 为与 $I_l(u,v)$ 对应的 Δ 值;$\hat{I}_{l-1}(2u,2v)$ 是 $I_{l-1}(2u,2v)$ 的估计值;为避免冲突,本文用 $\|\Delta_l(u,v)\|$ 表示对 $\Delta_l(u,v)$ 取整。算法首先根据 4 邻域像素求出 $\hat{I}_{l-1}(2u,2v)$;然后用 2×2 求均值的方法计算 $\hat{I}_l(u,v)$,从而得到 $\Delta_l(u,v)$;虽然 $\Delta_l(u,v)$ 近似等于 $I_l(u,v)$ 与 $I_{l-1}(2u,2v)$ 之差,却不依赖于它们,仅用保留的像素就可求得;利用 $\|\Delta_l(u,v)\|$ 在 $I_l(u,v)$ 与 $I_{l-1}(2u,2v)$ 之间建立起关系,是本金字塔的特色所在。

由式(1)可知,当被舍弃像素的估计值完全准确,即 $\hat{I}_{l-1}(2u,2v)=I_{l-1}(2u,2v)$ 时,$I_l(u,v)$ 与均值金字塔中的非常接近(只有取整方式的差异)。

若已知 $I_{l-1}(2u,2v)$ 的邻域像素及其父像素 $I_l(u,v)$,按式(2)即可反推出:

$$
I_{l-1}(2u,2v)=I_l(u,v)-\|\Delta_l(u,v)\| \tag{2}
$$

因此,在存储金字塔时可以将 $I_{l-1}(2u,2v)$, $l=1,2,\cdots,L-1$;$u,v=0,1,\cdots,\frac{N}{2^l}-1$ 全部舍弃,即每

个 2×2 的区域内舍弃一个像素。

2.2 类高斯金字塔结构

本金字塔结构是由高斯金字塔改进而来。与高斯金字塔类似,首先定义一个权值矩阵:

$$\boldsymbol{W}=\begin{bmatrix}\omega_{-1,-1} & \omega_{-1,0} & \omega_{-1,1}\\ \omega_{0,-1} & \omega_{0,0} & \omega_{0,1}\\ \omega_{1,-1} & \omega_{1,0} & \omega_{1,1}\end{bmatrix}$$

则金字塔的构建规则如式(3)所示:

$$\begin{aligned}
&I_0(u,v)=I(u,v);u,v=0,1,\cdots,N-1\\
&I_l(u,v)=I_{l-1}(2u,2v)+\|\Delta_l(u,v)\|\\
&\Delta_l(u,v)=\hat{I}_l(u,v)-\hat{I}_{l-1}(2u,2v)\\
&\quad=\omega_{-1,-1}\cdot I_{l-1}(2u-1,2v-1)+\\
&\quad\omega_{-1,0}\cdot I_{l-1}(2u-1,2v)+\\
&\quad\omega_{-1,1}\cdot I_{l-1}(2u-1,2v+1)+\\
&\quad\omega_{0,-1}\cdot I_{l-1}(2u,2v-1)-\\
&\quad(1-\omega_{0,0})\cdot \hat{I}_{l-1}(2u,2v)+\\
&\quad\omega_{0,1}\cdot I_{l-1}(2u,2v+1)+\\
&\quad\omega_{1,-1}\cdot I_{l-1}(2u+1,2v-1)+\\
&\quad\omega_{1,0}\cdot I_{l-1}(2u+1,2v)+\omega_{1,1}\cdot I_{l-1}(2u+1,2v+1)\\
&\hat{I}_{l-1}(2u,2v)=\frac{1}{4}\cdot[I_{l-1}(2u-1,2v)+I_{l-1}(2u,2v-1)+I_{l-1}(2u,2v+1)+I_{l-1}(2u+1,2v)]\\
&l=1,\cdots,L-1;u,v=0,1,\cdots,\frac{N}{2^l}-1
\end{aligned}\tag{3}$$

各种符号的定义同前。在根据 4 邻域像素求出 $\hat{I}_{l-1}(2u,2v)$ 之后,算法用高斯滤波的方法计算 $\hat{I}_l(u,v)$,从而得到 $\Delta_l(u,v)$;最后利用 $\|\Delta_l(u,v)\|$ 在 $I_l(u,v)$ 与 $I_{l-1}(2u,2v)$ 之间建立起关系。

由式(3)可知,当被舍弃像素的估计值完全准确时,$I_l(u,v)$ 与高斯金字塔中的非常接近(只有取整方式的差异)。

若已知 $I_{l-1}(2u,2v)$ 的邻域像素及其父像素 $I_l(u,v)$,按式(4)即可反推出:

$$I_{l-1}(2u,2v)=I_l(u,v)-\|\Delta_l(u,v)\|\tag{4}$$

同样的,在存储金字塔时将 $I_{l-1}(2u,2v)$, $l=1,2,\cdots,L-1;u,v=0,1,\cdots,\frac{N}{2^l}-1$ 全部舍弃。

3 主要性能的理论分析

3.1 精度

现有的各种金字塔结构,比如:均值金字塔、高斯金字塔、拉普拉斯金字塔、Reduced-Sum 金字塔等,都存在精度截断的问题;即:根据子像素计算出的父像素带有小数。比如:均值金字塔的父像素计算式如公式(5)所示:

$$I_l(u,v)=\frac{1}{4}\cdot[I_{l-1}(2u,2v)+I_{l-1}(2u,2v+1)+I_{l-1}(2u+1,2v)+I_{l-1}(2u+1,2v+1)]\tag{5}$$

假设有 4 个邻域像素值分别为:126, 131, 135 和 143,则均值为 133.75,需要多用两个比特来存储小数的取值(4 种可能);如果舍弃小数,则无法精确恢复子像素,造成精度截断。

本文提出的金字塔则不存在此问题。从式(2)和式(4)可以看出,$I_{l-1}(2u,2v)$ 和 $I_l(u,v)$ 只相差一个整数 $\|\Delta_l(u,v)\|$,根本不存在精度截断问题。只要反推时对 $\Delta_l(u,v)$ 采用的取整策略和构建时相同,$I_{l-1}(2u,2v)$ 就能被精确恢复。这样,既不影响精度,又为每个像素节省了 2 bits 的存储空间。

3.2 存储量

在本文的两种金字塔结构中,除了第 $L-1$ 层之外,其余各层均有 1/4 的像素被舍弃。因此所存储的像素数都如式(6)所示:

$$\begin{aligned}Q&=\left(N^2+\left(\frac{N}{2}\right)^2+\cdots+\left(\frac{N}{2^{L-2}}\right)^2\right)\times\frac{3}{4}+1\\&=\frac{N^2\cdot\left[1-\left(\frac{1}{4}\right)^{L-1}\right]}{1-\frac{1}{4}}\times\frac{3}{4}+1=N^2\\&(\because N=2^{L-1})\end{aligned}\tag{6}$$

由此可见,所需存储的像素数恰好与原始图像相同,即完全没有数据膨胀;这与 Reduced-Sum 金字塔相同。而常见的金字塔,如均值金字塔、高斯金字塔与拉普拉斯金字塔等,所需存储的像素数均为

$$Q'=\frac{4}{3}\cdot N^2\cdot\left(1-\frac{1}{4^L}\right)\approx\frac{4}{3}\cdot N^2\tag{7}$$

3.3 构建速度

构建金字塔时,$I_0(u,v)$ 不需要计算,余下各层所需求解的像素总数为

$$T = \left(\frac{N^2}{4} + \cdots + 4 + 1\right) = \frac{N^2}{3} \cdot \left(1 - \frac{1}{4^{L-1}}\right) \tag{8}$$

对于类均值金字塔结构,由式(1)可知,每像素需要进行 3 次乘法,7 次加(减)法。

对于类高斯金字塔结构,由式(3)可知,每像素需要进行 10 次乘法,12 次加(减)法。

而对于均值金字塔,每像素需要进行 1 次乘法,3 次加(减)法;对于高斯金字塔,当采用 3×3 的高斯核时,每像素需要进行 9 次乘法,8 次加(减)法。由此可见,本文的金字塔结构在解决精度截断问题,并消除数据膨胀的同时,一定程度上牺牲了构建速度;但增加这些运算量并不对其实用性造成本质影响。

4　实验结果

对于金字塔结构而言,除了上述分析的各种性能之外,视觉效果也是一个重要的衡量指标。采用标准测试图像 Lake 进行实验,该图像为 256×256,8 bpp。实验对现有的均值金字塔、高斯金字塔和本文提出的类均值金字塔、类高斯金字塔进行了测试。其中,高斯金字塔与类高斯金字塔的高斯核(权值矩阵)均为

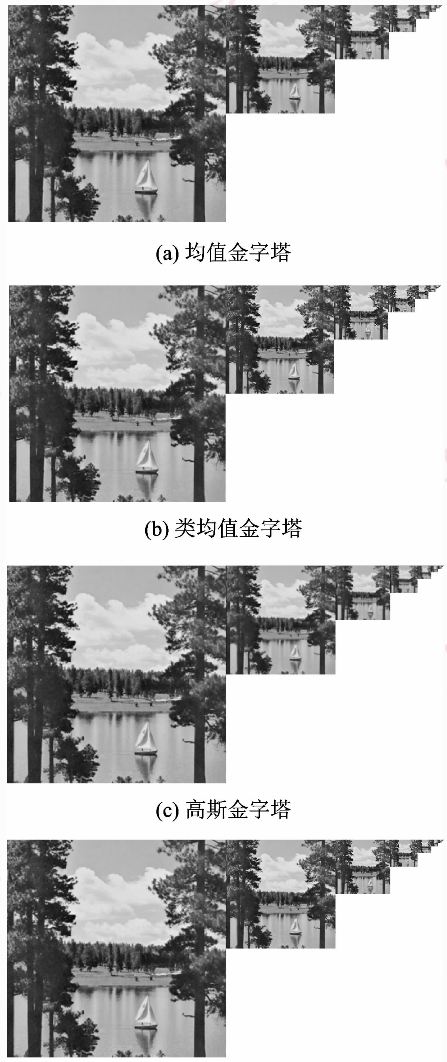
$$\boldsymbol{W} = \frac{1}{16} \cdot \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

针对 Lake 图像分别建立了上述 4 种金字塔,如图 2 所示:

从图中可以看出,本文提出的类均值金字塔、类高斯金字塔分别与现有的均值金字塔、高斯金字塔具有相近的视觉效果。不过现有两种金字塔的特点是比较平滑,而本文所提出两种金字塔则略微增强了边缘。

进一步,定量比较本文金字塔与经典金字塔的差异,得到类均值金字塔与均值金字塔的平均像素差为 2.079 4,类高斯金字塔与高斯金字塔的平均差为 1.790 5。其中,金字塔 I 与金字塔 J 的平均像素差定义为

$$D(I,J) = \frac{\sum_{l=0}^{L-1} \sum_{u=0}^{\frac{N}{2^l}-1} \sum_{v=0}^{\frac{N}{2^l}-1} \left[I_l(u,v) - J_l(u,v) \right]}{\sum_{l=0}^{L-1} \left(\frac{N}{2^l}\right)^2} \tag{9}$$



(a) 均值金字塔

(b) 类均值金字塔

(c) 高斯金字塔

(d) 类高斯金字塔

图 2　金字塔视觉效果对比

Fig. 2　Comparison of visual quality of pyramids

5　结　论

本文从同一思想出发,提出了两种图像金字塔结构,分别由均值金字塔与高斯金字塔改进而来。这两种金字塔不存在因舍弃小数而带来精度截断问题,解决了常见金字塔的通病;并且它们不产生数据膨胀,存储量小于一般的金字塔结构;从视觉效果上看,这两种金字塔分别与均值金字塔、高斯金字塔相近。但在获得这些性能提升的同时,构建金字塔所需的运算量略有增大。

在构建本文的金字塔时,对舍弃像素的估计是个重要环节。在后续的工作中,可以寻找更合适的像素估计方法,进一步提高金字塔的视觉效果,同时

又尽可能减小运算量。此外,从本文的思想出发还可以设计更多的金字塔结构,以满足不同的应用需求。

参考文献(References)

1 Gonzalez R C, Woods R E. Digital Image Processing[M]. 2nd ed. Translated. by Ruan qiu-qi. Beijing: Publishing House of Electronics Industry, 2003. [冈萨雷斯著. 数字图像处理[M]. 第二版. 阮秋琦译. 北京:电子工业出版社, 2003.]

2 Rashwan M A, Elsherif M S, Elsayad A M. Pyramid data structures for on-line image progressive transmission[A]. In: Proceedings of the 36th Midwest Symposium on Circuits and Systems[C]. Detroit, MI, USA, 1993: 103-106.

3 Miao Qi-guang, Wang Bao-shu. A novel image fusion algorithm based on PCNN and contrast[A]. In: Proceedings of the International Conference on Communications, Circuits and Systems[C]. Guilin Guangxi, China, 2006: 543-547.

4 Flierl M, Vandergheynst P. An improved pyramid for spatially scalable video coding[A]. In: Proceedings of IEEE International Conference on Image Processing (ICIP)[C], Genoa, Italy, 2005: 878-881.

5 Qiu Guo-ping. A progressively predictive image pyramid for efficient lossless coding[J]. IEEE Transactions on Image Processing, 1999, **8**(1): 109-115.

6 Dai Chen-guang, Zhang Yong-sheng, Deng Xue-qing. An organization and management approach of data for real-time visualization of massive terrain dataset[J]. Journal of System Simulation, 2005, **17**(2): 406-409, 413. [戴晨光,张永生,邓雪清. 一种用于实时可视化的海量地形数据组织与管理方法[J]. 系统仿真学报, 2005, **17**(2): 406-409, 413.]

7 Goldberg M, Wang L. Comparative performance of pyramid data structures for progressive image transmission[J]. IEEE Transactions on Communications, 1991, **39**(4): 540-548.

8 Rosiles J G, Smith M J T. A low complexity overcomplete directional image pyramid[A]. In: Proceedings of the International Conference on Image Processing[C], Barcelona, Catalonia, Spain, 2003: 1049-1052.

9 Duchesnay E, Montois J J, Jacquelet Y, *et al.* Irregular adaptive pyramid of agents for segmentation to interpretation of image[A]. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics[C], Nashville, Tennessee, USA, 2000: 1574-1580.