

# 图象视觉主分量的建模与提取

袁杰辉

王玉平 蔡元龙

(西安交通大学,西安 710049) (北京大学视听信息处理国家实验室,北京 100871)

**摘要** 边缘是图象的重要视觉信息。此文提出了新的视觉主分量模型(Visual Primary Component, VPC)。采用具有快速算法的B样条小波变换,根据边缘的强度(变化幅度、宽度及孤立性)、长度、均匀性等特性,设计了有效的VPC提取和编码算法,并给出了计算机模拟结果。

**关键词** 视觉感知,视觉主分量,边缘提取,多尺度B样条小波变换,压缩编码

## 1 引言

边缘是图象内容发生变化的地方。视觉实验表明,正是边缘引起人的视觉感知<sup>[1]</sup>,因而是图象非常重要的视觉信息。边缘信息可由强度变化及其所在位置来描述。不同性质的边缘在人的视觉系统中所产生的反应有较大区别,这种视觉差异可在图象处理的各种应用场合中加以利用。特别是 Xiaonong Ran 等<sup>[1]</sup>研究指出,强边缘更是视觉首要的。所谓强边缘,是指变化剧烈、宽度(边缘上下沿之间的距离)较窄且较孤立的边缘。

众所周知,在图象压缩编码系统中,追求以尽可能少的比特数保持恢复图象的高保真度。但在基于波形编码技术中(如DCT、子带、矢量量化等)常用的失真度标准最小均方差(MSE)并不能充分与视觉特性相匹配。在这类编码器的设计中,人类视觉特性的考虑是局部和次要的。通常在中等码率时,波形编码有高的恢复质量。当压缩比增加时,诸如方块效应、边缘模糊和振铃等失真逐渐明显。另一方面,基于特征的编码技术侧重于结构、几何或视觉的图象基本分量分析,通过边缘的检测进行区域分割,然后对区域轮廓和区域内部各自独立编码。因此,这种方法在利用图象的几何冗余时,还利用了轮廓和区域内部不同的统计与视觉特性,在整体上与HVS联

系更紧,可有较高的如50甚至100压缩倍数,但这类方法一般失之过于复杂。

为追求更高的压缩比而保持较低的视觉失真,综合地运用上述两种方法不失为一种有效途径<sup>[2,3]</sup>。作者给出了一种有效的编码方案<sup>[4]</sup>,其实质是将图象中首要的视觉信息提取出来并单独编码表达,而对剩余分量采用小波编码,从而显著地改善高压缩比下的主观效果。

## 2 多尺度B样条小波变换及快速实现

### 2.1 一维情形

在此仅对多尺度B样条小波变换作一简介。

$n$ 阶连续B样条定义为: $\beta^n(x) = [\beta^0(x) * \beta^0(x) * \dots * \beta^0(x)]_{n+1}$ , $\beta^0(x)$ 为0阶中心B样条。 $n$ 阶离散B样条定义为: $b_m^n(k) = \beta^n\left(\frac{k}{m}\right)$ , (或者 $b_m^n(k) = \beta^n\left(\frac{k}{m} + \frac{1}{2}\right)$ ) $m, k$ 为整数。考虑信号 $s(x)$ 的连续小波变换

$$WS(a, b) =$$

$$\frac{1}{a} \int_{-\infty}^{\infty} s(x) \Psi\left(\frac{b-x}{a}\right) dx = (\Psi_a * s)(b)$$

其中 $\Psi(x)$ 称为分析小波, $a, b$ 为连续变化的尺度和位移参数, $\Psi_a(x) = \frac{1}{a} \Psi\left(\frac{x}{a}\right)$ 。对 $a, b$ 离散取整,

\* 国家自然科学基金资助项目

收稿日期:1996-01-01;收到修改稿日期:1996-08-20

即  $a = m, b = k$ , 将信号和小波都表示为 B 样条的逼近, 用 B 样条的优良性质便可得到与尺度大小无关的快速计算方法。为方便计, 信号  $s$  取 0 阶 B 样条逼近。分析小波  $\Psi(x)$  用  $n$  阶离散 B 样条展开:

$$\begin{aligned} WS(m, k) &= \frac{1}{m} \left( \sum_i p(i) \beta^n \left( \frac{k}{m} - i \right) \right) * \left( \sum_l s(l) \beta^0(k - l) \right) \\ &= \frac{1}{m} (b^{n+1} * u_m^n * s * p_{\uparrow m})(k) \end{aligned}$$

其中  $b^{n+1}$  是尺度为 1 的  $n + 1$  阶离散 B 样条,  $u_m^n$  为  $n + 1$  个宽度为  $m$  的方波级联卷积,  $\uparrow m$  表示以  $m$  间隔上采样(upsampling)。上式可分为三步实现:

- (1) 初始化  $s_1(k) = (b^{n+1} * s)(k)$ , 可通过迭代算法实现。
- (2) 移动平均求和  $s_2(m, k) = u_m^n * s_1(k)$ , 可只用加法迭代更新实现。
- (3) 后处理  $s_3(m, k) = [p]_{\uparrow m} * s_2(m, k)$ 。若小波取光滑 B 样条,  $p \equiv 1$ , 此步可省略。二阶导数情形取  $p \equiv [-1, 2, -1]$ 。

### 2.2 二维情形

此窗函数  $\Psi(x)$  通常为二维可分 B 样条  $\beta^n(x) \beta^n(y)$  的 Laplacian 算子或方向导数。本文为前者情形。

$$\begin{aligned} \nabla^2 \beta^n \left( \frac{x}{a}, \frac{y}{a} \right) &= \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \beta^n \left( \frac{x}{a} \right) \left( \frac{y}{a} \right) = \\ &\Psi^n \left( \frac{x}{a} \right) \beta^n \left( \frac{y}{a} \right) + \beta^n \left( \frac{x}{a} \right) \Psi^n \left( \frac{y}{a} \right) \end{aligned}$$

其中  $\Psi^n \left( \frac{x}{a} \right) = \frac{\partial^2}{\partial x^2} \beta^n \left( \frac{x}{a} \right)$

这表明二维窗函数可分解成两个方向可分离的算子的和, 即可利用以上一维情形, 将图象分别沿水平和垂直方向与相应的窗函数作卷积, 然后相加就得到了最后的结果, 且与尺度大小无关。

## 3 视觉主分量(VPC)模型

### 3.1 强边缘轮廓的提取

设原始图象为  $I \equiv \{(i, j, x_{i,j})\}$ , 其中  $i, j = 0, \dots, M-1, M$  为图象大小,  $x_{i,j}$  表示图象灰度值。图象像素灰度在边缘处发生变化, 低灰度像素构成的轮廓称为边缘下边沿(lower brims), 高灰度构成的轮廓为边缘上边沿(upper brims)。传统上, 边缘的宽度(即上下沿距离)被认为只有一个像素宽。这种不太真实的假设使边缘检测对噪声敏感, 容易产生伪边缘。通过对自然图象的观察, 我们认为(强)边缘

$$\Psi(x) = \sum_{k \in z} p(k) \beta^n(x - k),$$

于是可推出:

的宽度为 2 个像素宽。实验表明这种假设更符合实际情况, 有利于保持边缘检测的一致性。

图象经小波变换后, 通过检测零交叉可确定边缘的位置, 其区域极大值反映了边缘的肩或顶所在, 区域极小值则是脚或底位置。为提取边缘上沿, 在零交叉区域内以正值为中心, 对  $+-, ++, --$  情况(后两者表示宽度为 2), 取沿最大斜度的最大最小值之差的一半为零交叉梯度, 该梯度值反应了边缘变化强度。于是得到一个零交叉梯度图象  $G \equiv \{(i, j, g_{i,j})\}$ , 其中  $g_{i,j}$  为梯度值(见图 1)。

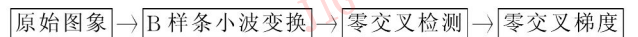


图 1 零交叉梯度图象 G 的产生。

Fig. 1 The generation of zero-crossing-gradient image G

由于零交叉梯度图象 G 中包含了各种各样错综复杂的边缘信息, 直接采用阈值方法提取 VPC 几乎是不可能的, 因而如何从 G 中有效提取 VPC 并利于编码是关键。根据强边缘的强度、长度和均匀性等特征, 设计了强边缘上沿轮廓(contours of upper brims)提取算法, 其思想是在尽可能保持边缘的连续性下, 对边缘分别处理, 达到尽可能消除弱边缘。具体步骤如下:

#### (1) 预处理及细化

① 阈值处理 如果  $g_{i,j} < T1$ , 则  $g_{i,j} = 0$ 。取较低的门限  $T1$  以保持边缘连续性, 文中  $T1 = 40$ 。

② 减小弱边缘的干扰 在  $3 \times 3$  邻域内, 若最大值超过最小值的两倍, 则将最小值置零。

③ 在 8 连通区域内, 采用外剥法, 取区域极大值将边缘逐步细化为一个像素宽。

④ 在细化后边缘上, 如果四个相联结的像素中最大值超过最小值的两倍, 则将最小值置零, 并消除只有一个像素的分枝点使之尽量光滑。

#### (2) 纹理区边缘消除

在图象压缩编码应用中, 人的视觉能容忍较大的纹理区边缘噪声, 因此纹理区边缘应尽量消除。一

般来说这种边缘梯度值不很大,且容易相互交错连接,规则性较差。将边缘位置看成树形结构,计算每一个树的总长度  $L$ 、均值  $M$ 、方差  $V$  和节点数  $N$ ,如果  $f_1(L, M, V, N) < T_2$ ,则该树舍去。其中:

$$f_1(L, M, V, N) = \frac{L}{N+2} \times \frac{M}{V+20}$$

$$V = \text{sqrt}\left(\frac{1}{L} \sum_{i=0}^{L-1} (g(j+1) - g(j))^2\right)$$

$g(j)$ 表示树上的象素值。常数项为减弱  $f_1$  对  $N, V$  的灵敏度,取  $T_2 = 15$ 。

(3) 弱边缘消除

① 对每个树搜索,在其每一节点,依据区域特性拆开其一支,亦即在节点  $3 \times 3$  邻域内,将极小值点置零。这样整个  $G$  图象由互不交叉的线段构成。

② 由于细化时不考虑边缘的坡度走向,两个不同斜坡方向的边缘可能相连一起,应根据边缘下沿的位置变化将其断开。设  $g(j)$  表示边缘上的象素,如果  $g(j+1)$  的下沿与  $g(j)$  的下沿位置不在同一边,则令  $g(j) = 0$ 。

③ 自然图象中许多边缘的强度逐渐减小,使得边缘两端强度较弱,对这种情况也需整理。令  $x_u(j), x_l(j)$  分别表示某边缘上、下沿上的原始象素值,则从端点  $j = 0$  开始。如果:

$$\sum_{l=0}^2 x_u(j+l) < \frac{4}{3} \sum_{l=0}^2 x_l(j+l),$$

则  $g(j), g(j+1), g(j+2)$  置零,

$$j = j + 3, \text{ 返回上式;}$$

否则转入另一边缘处理。

④ 对平均相对强度较小、长度较短、不太均匀的弱边缘进行消除。即当  $f_2(L, C, V_1, V_2) < T_3$  时,边缘删除。其中:

$$f_2(L, C, V_1, V_2)$$

$$= 100 \times (L + 8) \times C \times \frac{1}{V_1 + V_2 + 100}$$

$$C = \frac{M_u - M_l}{M_u}, \quad M_u = \frac{1}{L} \sum_{j=0}^{L-1} x_u(j),$$

$$V_1 = \text{sqrt}\left(\frac{1}{L} \sum_{j=0}^{L-1} (x_u(j+1) - x_u(j))^2\right)$$

以上参数均为原始图象  $I$  中的数值,  $L$  为边缘长度,  $M_u, M_l$  表示上、下边沿轮廓上的均值,  $C$  反映了边缘的相对变化强度(亦可认作局部对比度),  $V_1, V_2$  表示上、下边沿上的梯度方差,反映了边缘本身的均匀程度,见示意图 2。

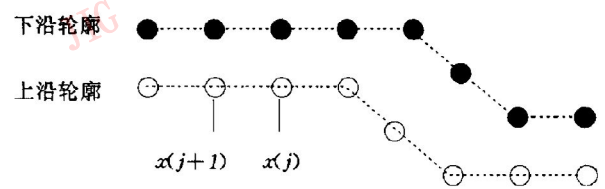


图 2 强边缘上、下沿轮廓示意图。

Fig. 2 The contours of upper and lower brimsqd of strong edge

(4) 阈值处理

经过以上过程,再根据边缘的(绝对)平均变化强度进行分段线性阈值处理。令  $L_1=6, L_2=12, L_3=18, L_4=30; D_1=140, D_2=120, D_3=50, D_4=30$ 。对长度为  $L (< L_4)$  的边缘,令  $D = M_u - M_l, L_i \leq L < L_{i+1}, i=1, \dots, 3$ , 有:

$$D_i = D_i + \frac{(L-L_i) \cdot (D_{i+1}-D_i)}{(L_{i+1}-L_i)}, \text{ 若 } D < D_i, \text{ 或}$$

$L < L_1$ , 则边缘消除。

(5) 连接

将相邻近且平均变化强度相近的边缘连接,以提高编码效率。如果两斜坡方向一致的边缘,端点在 6 个象素距离之内,且  $|D^{(1)} - D^{(2)}| < 20$ , 则将两者相连。

3.2 VPC 模型与编码

上节所得到的强边缘轮廓,其意义表示沿边缘两象素宽度上有平均最大的灰度变化。上沿轮廓位置确定了下沿的位置,以及它们的象素平均差值:  $D = M_u - M_l$  这样选择上沿轮廓起点,即沿上沿轮廓行走时,下沿在右手边。以上沿为中心和  $0.8D$  为初值向两边以指数方式扩展,这样得到的分量,由于其视觉重要性,称之为视觉主分量(Visual Primary Component, VPC)如图 3。图 4 中显示了 VPC 从原

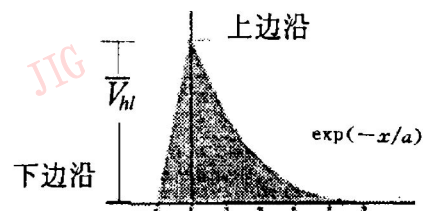


图 3 VPC 截面示意图。

Fig. 3 The profile of VPC

边缘分离状况。VPC 模型可表示为:

$$M_{VPC} = f(D, m, (x_0, y_0), s^0, r^1, r^2, \dots, r^{m-1}).$$

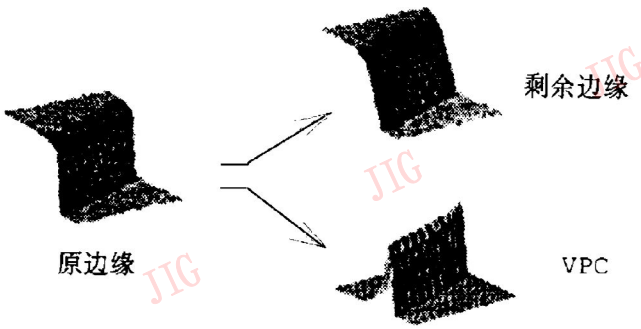


图 4 原边缘分解为 VPC 和剩余边缘。  
Fig. 4 Original edge is decomposed into VPC and the residual

#### 4 实验结果

$B$  条小波与 Marr 提出的拉普拉斯—高斯 (Laplacian of Gaussian, LoG)<sup>[5]</sup>算子相比,其高斯函数实际上被  $B$  样条函数所代替,并存在有近似关系:  $\frac{n+1}{12}m^2 \approx \sigma^2$ 。其中  $m, n$  为尺度和阶数,  $\sigma$  为高斯函数参数。同 LoG 算子的  $\sigma$  一样,尺度和阶数的选择十分重要,大的尺度和阶数虽有利于压抑纹理边缘,但对强边缘易产生较大定位误差。由于强边缘通常变化较剧烈,为了准确定位和计算方便,本文选择零次  $B$  样条小波,尺度  $m$  为 3。

为检验本文 VPC 提取算法的有效及鲁棒性,对六幅不同特性的图象进行了计算机模拟实验,并给出其中 3 幅的结果。图象大小为  $512 \times 512 \times 8$  bits,“Peppers”的边缘变化十分丰富,“Dock”边缘较规则但强度变化大小不一,“Girl”则含较多纹理性边缘。图象处理中由于图象内容的千变万化,阈值的选取是棘手的问题。所以在提取算法中,尽量协调各步骤间阈值关系,且相对固定下来,只保留一个可调门限  $T_3$  以适应不同性质的图象。 $T_3$  主要体现了与区域对比度的关系,比如在“Dock”中,尽管有的边缘对比度不是很大,但由于很规则并较孤立,其视觉效果仍较重要,因此阈值会低一些。3 幅图的  $T_3$  顺序为 10, 6, 8。图 5 显示了实验结果。

其中参数顺序为:上下沿轮廓上的像素平均差值,轮廓像素点数,轮廓起始位置,起始点链编码方向值,第二个轮廓点相对第一个轮廓点的链编码方向值,...

由于轮廓上链编码相对方向值规律性强,可以将所有的 VPC 链编码相对方向值放在一起用自适应代数编码压缩。而前 5 个参数可用固定比特表示。

此模型有效的保留了强边缘中的首要的视觉信息,并有利于图象剩余分量的量化编码。与[2]相比, VPC 编码比特减少一半。因为在[2]中,强边缘的上下沿需分别编码,尤其在低编码率情况下,这种码字节省是十分显著的。

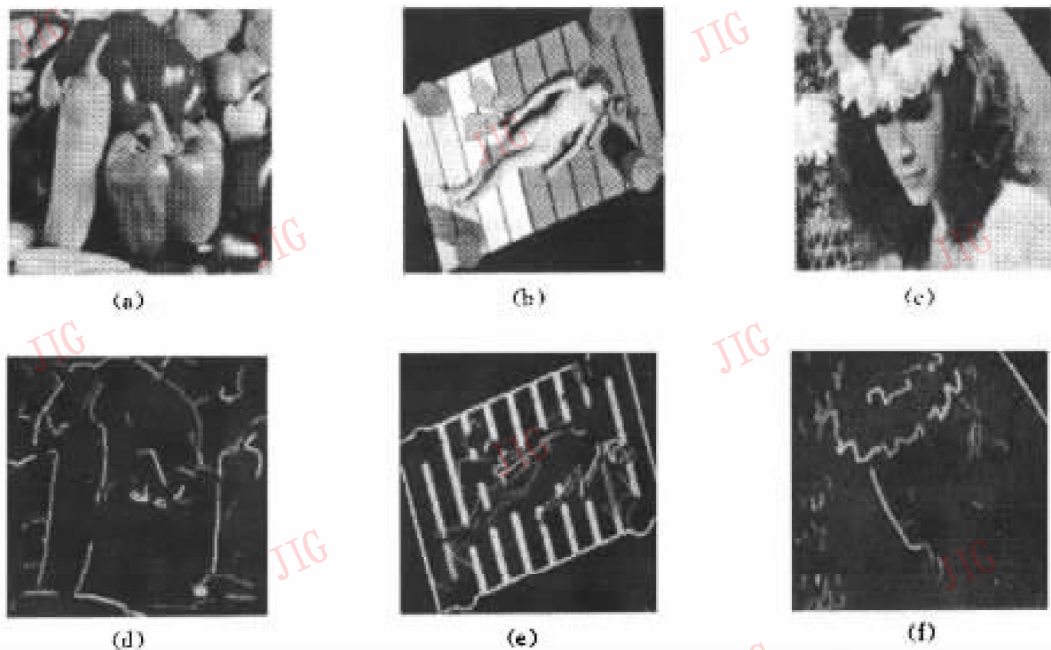


图 5 (a)“Peppers”, (b)“Dock”, (c)“Girl”, (d), (e), (f) 分别为(a), (b), (c)之 VPC 图象。  
Fig. 5 (a) “Peppers”, (b) “Dock”, (c) “Girl”; (d), (e), (f) are the VPC of (a), (b), (c) respectively.

## 5 结 论

本文引入了新的视觉主分量模型,以及结合具有快速算法  $B$  样条小波变换的  $VPC$  提取算法。 $VPC$  有效地保留了图象中首要的视觉信息,同时也改善了原图象信号的非平稳性。实验结果表明了这种模型及算法的有效性。

### 参 考 文 献

1 Ran Xiaonong, Farvardin N. A Perceptually Motivated Three-

Component Image Model—Part I: Description of the model. IEEE Trans. IP, 1995, 4(4):401~415.

2 Xiaonong Ran, Farvardin N. A Perceptually Motivated Three-Component Image Model—Part II: Applications to Image Compression. IEEE Trans. IP, 1995, 4(4):430~447.

3 Egger O, Li Wei. High Compression Image Coding Using an Adaptive Morphological Subband Decomposition. Proc. of IEEE, 1995, 83(2):272~287.

4 Yuan Jiehui, Cai Yuanlong, Wei Zhenggang, Low-rate Image Coding Using Visual Primary Component and Wavelet Transform, ICSP' 96, 1996. 10.

5 Canny J. A Computational Approach to Edge detection. IEEE Trans. PAMI, 1986, PAMI-8, (6):679~698.



袁杰辉, 博士研究生。1988 年获武汉大学无线电电子学硕士学位, 其后一直从事图象方面的技术研究和开发应用。感兴趣的领域为数据压缩, 图象处理, 数字图象通信等。

# The Modeling and Extraction of Visual Primary Component in Images

Yuan Jiehui, Wang Yuping, Cai Yuanlong

( Xian JiaoTong University, Xian 710049 )

( National Machinery Perception Lab. of Beijing Univ., Beijing 100871 )

**Abstract** Edges are of very visual importance. In this paper a novel visual primary component (VPC) model is proposed. By using the B-spline wavelet transform which has the advantage of fast algorithm, and according to the edge information such as strength, length and integrity, an effective algorithm for the extraction and coding of VPC is developed in the paper. Computer simulation results are given as well.

**Keywords** Visual perception, Visual primary component, Edge detection, Multiscale B-spline wavelet transform, Image compression