

面向对象分配带宽的极低比特率视频编码器

翁南钊 蔡德钧 刘李杰

(华中理工大学电子与信息工程系, 武汉 430074)

摘要 在低比特率环境下,如何提高图象编码质量是许多人感兴趣的课题,针对此问题,提出了一种面向对象分配带宽(OOBA)的视频编码器,用以提高低比特率环境下编码图象的主观质量,在充分考虑人的主观视觉特性的基础上,该编码器把图象序列分成不同的视频对象,而每个对象以宏块为单位构成,并且对每个对象单独分配大小不同的带宽,以把更多的比特分配给主观感兴趣的对象,来达到在低比特率环境下提高编码图象主观视觉质量的目的.虽然每个对象独立编码,但在比特流的组织上却能与H.263兼容.为了合理地多个对象间分配带宽,还给出了基于速率-失真(R-D)模型、基于序列分析和比值加权法等3种带宽分配方法.实验结果表明,该编码器与传统的基于帧的编码器相比,在极低的速率下,峰值信噪比虽略有降低,但图象的主观视觉质量却得到较大的提高.从而表明了该编码器的有效性,也说明在低比特率环境下,仅以客观准则作为评价编码器好坏的标准是不够的.

关键词 视频编码 低比特率视频编码 视频对象 带宽分配

中图分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2001)01-0070-07

Very Low Bit-Rate Video Coder with Object-Oriented Bandwidth Allocation Technology

WENG Nan-shan, CAI De-jun, LIU Li-jie

(E&I Department, Huazhong Univ of S&T, Wuhan 430074)

Abstract How to improve the subjective visual quality of the video coded at low bit-rate is an interested subject of many people. This paper proposes a video coding system based on object-oriented bandwidth allocation(OOBA) technology to improve the subject quality of coded images. Exploiting human subjective visual characteristic sufficiently, this coder divides video sequence into multi video objects(VO) and allocates more bandwidth for the VO with more subjective significance to improve subjective quality of images coded at low-bit-rate. Each object is composed of macroblocks and coded separately. In bit stream context, the OOBA coder is compatible with H.263. For reasonable bandwidth allocation, present three schemes to allocate bandwidth for multi objects: rate-distortion (R-D) model based scheme, sequence analysis based scheme and ratio weighting scheme. The result of experiments shows that the PSNR of OOBA coder is a little lower than that of frame-based coder, but the subjective quality is much better. So, it proves the OOBA coder effective and indicates it is not enough to evaluate a video coder only with objective criterion such as PSNR.

Keywords Video coding, Very low bit-rate video coding, Video object, Bandwidth allocation

0 引言

在低比特率环境下,如何提高图象编码质量是许

多人感兴趣的课题,因为在极低的速率下,不应只追求图象的客观评价标准(峰值信噪比)的高低,而应更多地考虑人的主观视觉质量,即应充分利用人的主观视觉特性,把更多的比特分配给主观感兴趣的对象.

本文针对可视电话的应用,提出了一种面向对象分配带宽(OOBA)的视频编码器,其在低比特率环境下可以提高编码图象的主观质量,同时可以提供基于对象的图象质量可扩缩性(scalability),而在比特流结构上还能与 H. 263 兼容. 本文还提出基于速率失真(R-D)模型的方法、基于视频序列分析的方法和简单的比值加权法等 3 种带宽分配方法. 这种视频编码器也可以用在主观上希望突出某部分图象区域的应用(如视频监控)场合中. 本文在带宽分配问题初步研究^[1]的基础上,又进一步进行了补充和完善. 通过对可视电话视频序列进行测试的结果表明,这种低比特率视频编码器解码图象的 PSNR 虽略有下降,但主观视觉质量却有了较大提高.

1 编码系统概述

通过考察人的主观视觉特性可以发现:人的注意力更多地集中在运动剧烈和主观感兴趣的部分,如在可视电话的场景中,人的脸部是运动最剧烈,也

是观众最感兴趣的部分,肩部则稍次,而背景则经常被忽略,所以,可先把场景分为头、肩和背景 3 个视频对象(VO)来进行编码,而每个对象以宏块为单位构成,再根据重要性程度的不同,分配不同的带宽. 一般对象区域的定义每隔一段时间更新一次,与之相应,带宽分配也每隔同样的时间进行一次. 虽然每个 VO 独立地进行编码和速率控制,但传输时,各个 VO 同一时刻的视频对象面 VOP(即 VO 的某一帧)又组合成原来的矩形帧,再以 H. 263 的比特流结构进行打包传输,解码也是以矩形帧为单位进行的,也就是说,在比特流结构上,与 H. 263 保持了兼容性. 这种方法的优点在于提供了基于对象的图象质量可扩缩性,因而可以在更大程度上把比特分配给感兴趣的对象,并可以更加灵活和更大范围地控制不同对象的图象质量. 简言之,即传输以帧为单位,编码和速率控制以对象为单位. 本文提出的一种 OOBA 编码器模型可参见图 1.

OOBA 编码器与 MPEG-4 基于对象的视频编码器相比,其存在的异同点可参见图 1 和图 2.

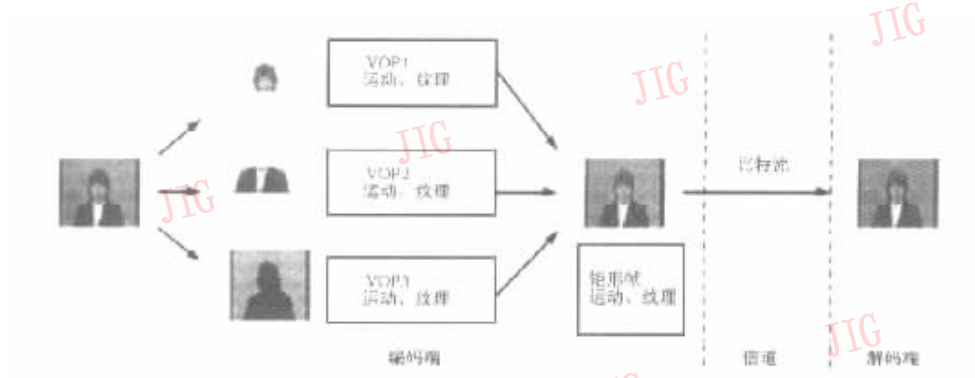


图 1 OOBA 编码器模型

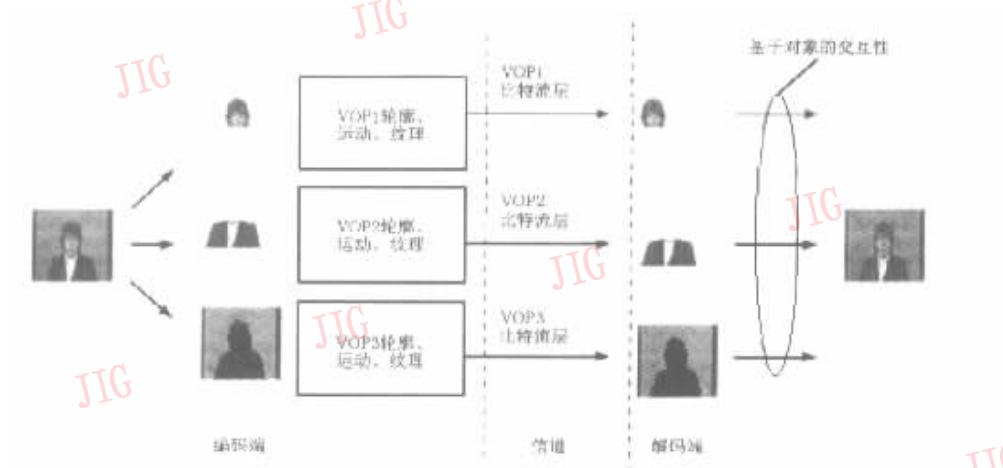


图 2 MPEG-4 编码器模型

MPEG-4^[2] 视频编码器也是把场景分为几个 VO, 分别进行编码, 这一点与 OOBA 编码器是相同的, 但不同的是, MPEG-4 编码器编码的每个 VO 是形成各自独立的比特流层, 而各个比特流层独立地进行传输, 解码时也是对各个 VO 分别进行解码, 以提供基于对象的交互性; 而 OOBA 编码器编码后的各个 VO, 则在编码端又组合成矩形帧, 以与 H. 263 兼容的比特流语法进行传输, 因此无法提供基于对象的交互性能(在有些应用场合并不需要这种性能). 由于 MPEG-4 编码器编码的 VO 可以是任意形状的, 因此形状信息也要进行编码传输; 而本文提出的 OOBA 编码器在编码端就把 VO 组合成矩形帧进行传输, 因此无需编码和传输形状信息.

OOBA 编码器之所以采用与 MPEG-4 编码器不同的做法, 主要是因为 OOBA 划分 VO 的目的是为了提高低比特率编码图象的质量, 而 MPEG-4 划分 VO 的目的是为了提供基于对象的交互性. 这样两者相比, 前者由于不需编码 VO 的形状信息, 因而节省了比特开销, 同时提供了与 H. 263 兼容的比特流语法, 以适用于一些传统的应用.

2 对象定义

总的来说, 对象定义的方法可以分为手动定义和自动定义两种. 由于在某些应用场合(如视频监控)中, 摄像机的镜头可能是固定的, 场景也是基本上不变的, 因此这时候就可以人为地划定对象区域,

以突出感兴趣区域的图象细节, 在这种场合下, 对象的手动定义是适合的, 但需要注意的是, 手动定义对象的形状在编码的全过程中是不变的, 因为在编码过程中, 如果再利用人的交互来不时地改变对象形状是不现实的. 而自动定义法则无需人的交互, 它可以每隔一帧或一段时间更新对象形状一次. 文献[1]针对可视电话的应用场景提出的一种对象的自动定义方法, 其过程分为对象分割、头肩分离和宏块覆盖等 3 个步骤.

2.1 对象分割

该步骤是首先采用四阶矩检测器^[3]来检测两幅图象的帧间变化, 以得到一个二值图(模板), 如图 3(a)所示, 图中黑色的部分表示静止部分, 白色的部分表示运动部分. 由于两幅图象间的变化可能不是太剧烈, 所以这个模板可能还不能覆盖整个头肩. 为了分割出有意义的头肩部, 在一段时间的图象序列中, 通过多次对不同的两幅图象进行检测, 即可得到一组模板, 然后把这些模板叠加起来, 就可得到一个能完全覆盖头肩部的模板(图 3b), 但是, 这个模板还不能精确地表示对象的轮廓, 所以还要接着用数学形态学的腐蚀算子^[4]从模板的边缘向内腐蚀, 直到对象的边缘, 以得到最后的二值分割图(图 3c). 根据二值分割图即可分割出头肩部(图 3d). 这种方法对可视电话序列图象的分割效果较好, 但也适用于其它类型的序列图象, 更详细的讨论可参见文献[1].

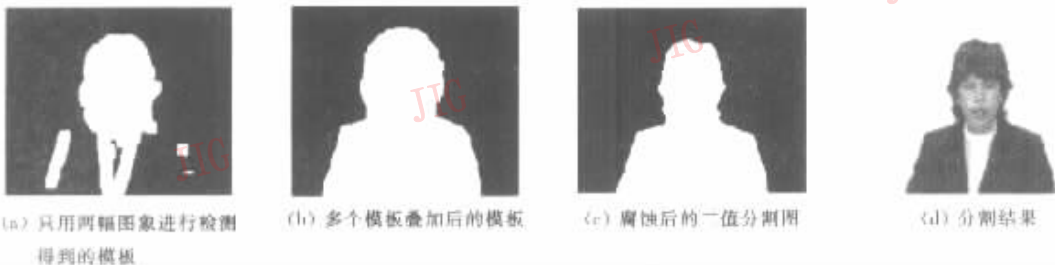


图 3 对象分割过程示意

2.2 头肩分离

头肩分离的目的在于把头肩部分为头部和肩部两个对象, 该步骤可采用一种简单然而高效的基于查找头肩轮廓最大转折点的头肩分离算法^[1]. 由于头肩部具有对称性, 所以只要查找头部与肩部的一条水平分界线(称为分离线)即可. 具体做法是, 首先, 用多边形近似法^[5]来近似对象的轮廓, 以避免轮

廓细节对分离结果的影响; 然后, 以头肩部的高度为自变量 x , 以宽度为因变量 y , 建立函数 $y=f(x)$. 这样要分离头肩部即相当于先找到 $f(x)$ 上具有最大转折率的一点 (x_0, y_0) , 然后用与该点对应的水平线 $x=x_0$ 来分离头肩. 这里, 定义转折率为 $f(x)$ 上一点 $+x$ 方向的导数(正导数)与 $-x$ 方向的导数(负导数)的差的绝对值. 本文对导数的计算采用了插值

型的 3 点求导公式^[6]. 图 4 是形状近似和头肩分离后的结果.



图 4 形状近似和头肩分离的结果

2.3 轮廓覆盖

为了减少编码的复杂度,可以宏块为单位构成对象,并用最少数目的宏块覆盖头肩像,然后把位于分离线以上(包括分离线上)的宏块定义为头部对象,分离线以下的宏块定义为肩部对象,从而把场景分为图 5 所示的 3 个对象.由于每个对象以宏块为单位构成,所以对象之间的区域不互相覆盖.



图 5 对象区域定义图

(图中白线把图象分为头部、肩部和背景 3 个对象)

在这个编码器中,头肩像的分割要考察一段时间内的图象序列,但实际上,对象分割、头肩分离和轮廓覆盖都只需对这段时间内的第一帧图象进行,因为在这段时间内对象的形状和大小是不变的,分配的带宽也是不变的.

3 带宽分配

OOBA 视频编码器最核心的特征在于针对不同的对象分配不同的带宽,也就是说,带宽分配是这个视频编码器中的关键问题.近来,由于 MPEG-4 视频编码器的兴起,多对象的速率控制问题已经引起人们的重视^[7,8],本文提出 3 种基于对象的带宽分配算法,即基于速率失真(R-D)模型的分配方法、基于视频序列分析的方法和简单然而有效的比值加权法.文献[7]、[8]中研究的比特分配方法着重考察如何对不同的 VO 在同一时刻的 VOP 之间分配比

特,而本文研究的重点则在于考察如何对不同的 VO,在同一时间段内的 VOP 序列之间分配比特,由此可见,后者更为宏观,当时间段取得很短,只包含一个 VOP 帧时,后者研究的问题就与前者相同.下面以一个矩形图象序列分为 3 个 VO 为例,来说明带宽分配方法.

3.1 基于速率失真的分配方法(简称 RD 法)

该方法首先是建立视频编码的速率失真模型,可采用文献[9]中提出的二次速率失真模型

$$R = aQ^{-1} + bQ^{-2} \quad (1)$$

其中 R 为速率, Q 为量化参数, a 和 b 分别为一阶和二阶模型参数.在这个模型中是采用量化参数作为失真的度量.

接着,估计各个 $VO_i (i=1,2,3)$ 的模型参数 a_i, b_i . 这里模型参数的估计可以采用文献[9]中介绍的方法,也可以采用最佳最小二乘解^[10]的方法.值得一提的是,这种带宽分配方法并不受限于如何找到模型参数.假设 B 为总带宽, $B_i (i=1,2,3)$ 为分配给 VO_i 的带宽,则有下列等式

$$B = B_1 + B_2 + B_3 \quad (2)$$

同时有

$$B_1 = a_1 \bar{Q}_1^{-1} + b_1 \bar{Q}_1^{-2} \quad (3)$$

$$B_2 = a_2 \bar{Q}_2^{-1} + b_2 \bar{Q}_2^{-2} \quad (4)$$

$$B_3 = a_3 \bar{Q}_3^{-1} + b_3 \bar{Q}_3^{-2} \quad (5)$$

其中, \bar{Q}_i 为 VO_i 的所有 VOP 量化参数的平均值.由于 OOBA 视频编码器考虑到人的主观特性,并希望主观感兴趣的对象具有较好的图象质量,所以各个对象达到的失真度也应是不同的,即 \bar{Q}_i 不同,各个对象之间的失真比可以用下式来表示

$$\bar{Q}_1 : \bar{Q}_2 : \bar{Q}_3 = k_1 : k_2 : k_3 \quad (6)$$

$k_i (i=1,2,3)$ 取值的大小将影响 VO_i 的失真度 \bar{Q}_i , 即 k_i 取值越小, \bar{Q}_i 就越小, \bar{Q}_i 越小,则分配给 VO_i 的带宽 B_i 就越大.由此可见, k_i 取值的大小与 VO_i 希望分得的带宽大小成反比,即主观上希望提高图象质量的 VO_i 应具有较小的 k_i 值,这样即可通过调整 k_i 值来达到调整带宽分配的目的.根据式(1)~式(6)可以计算出 B_1, B_2, B_3 .

3.2 基于序列分析的分配方法(简称分析法)

合理的带宽分配至少应该考虑如下 4 个因素:

(1)对象面积大小;(2)对象空域细节的丰富程度;(3)对象时域上的变化程度;(4)希望编码对象达到的图象质量.其中对第 1 个因素,即对象面积大小采用 VO 所包含的宏块数来表征,记为 SIZE;对第 2

个因素,即空域细节丰富程度,采用 VO 的 I 帧图象 DCT 系数绝对值的平均来表征,记为 MACDCT;对第 3 个因素,即时域变化程度用 MOT 和 VAR 来表征.其中,MOT 为对象内所有宏块的运动矢量中各分量的绝对值和,而 VAR 的定义如下

$$VAR = MAD^2 \quad (7)$$

式中的 MAD 为 VO 的 P 或 B 帧图象经运动补偿后,其帧间差值绝对值的平均;对第 4 个因素,即希望编码图象达到的图象质量,是采用主观权重因子 SUBJ 来表示, SUBJ 是一个小于 1 的数,它表示主观上希望某个对象的图象质量与其它对象相比的好坏程度.在定义上述几个表征量后,即可写出带宽分配公式

$$B_i = B(w_1 SIZE_i + w_2 MACDCT_i + w_3 VAR_i + w_4 MOT_i + w_5 SUBJ_i) \quad (8)$$

其中, $SIZE_i$ 、 $MACDCT_i$ 、 VAR_i 、 MOT_i 分别为 VO_i 的 $SIZE$ 、 $MACDCT$ 、 VAR 、 MOT 对所有对象的归一值. $SUBJ_i$ 为 VO_i 的主观权重因子,它满足约束条件

$$\sum_i SUBJ_i = 1, \quad SUBJ_i \in [0,1] \quad (9)$$

$w_1 \sim w_5$ 为加权因子,它满足条件

$$\sum_{i=1}^5 w_i = 1, \quad \{w_1, w_2, w_3, w_4, w_5\} \in [0,1] \quad (10)$$

3.3 比值加权法(简称比值法)

虽然上面的两种方法都可以较为合理地分配带宽,但缺点在于花费的时间较多.这里再介绍一种简单的比值加权法.用这种方法来分配带宽几乎不花费时间,但较前两种方法更依赖于人的主观判断.

由于对象的面积大小是不同的,所以用“对象单位面积内分得的带宽(B/S)”作为分配的准则.假设用 S_i ($i=1,2,3$) 表示每个对象的面积(即所包含的像素数目或宏块数目),则对象间的 B/S 比如下:

$$(B_1/S_1) : (B_2/S_2) : (B_3/S_3) = W_1 : W_2 : W_3 \quad (11)$$

其中, W_i 为加权因子,由于可通过改变加权因子的取值,来改变带宽的分配,因此希望具有更佳图象质量的对象应该赋予更大的加权重.由式(11)可得各个对象的带宽比为

$$B_1 : B_2 : B_3 = W_1 S_1 : W_2 S_2 : W_3 S_3 \quad (12)$$

在给定了编码总带宽 B (速率)的情况下,即可得到分配给每个对象的带宽为

$$B_i = B \left(W_i S_i / \sum_j W_j S_j \right) \quad (13)$$

4 实验与讨论

本文采用 QCIF (176×144) 格式的多个序列,分别用上面介绍的 3 种带宽分配方法进行实验,这里仅给出 Claire 和 Miss America 序列的实验结果.实验是用 VO_1 、 VO_2 、 VO_3 表示头部、肩部和背景 3 个对象.在基于速率失真模型的带宽分配方法中,取 $k_1 : k_2 : k_3 = 1 : 1.3 : 1.4$,因为希望看到分配更多比特的对象,所以取了较小的比值;而在基于序列分析的带宽分配方法中,取 $w_1 = 0.1$, $w_2 = 0.05$, $w_3 = 0.2$, $w_4 = 0.05$, $w_5 = 0.6$, $SUBJ_1 = 0.5$, $SUBJ_2 = 0.35$, $SUBJ_3 = 0.15$,为了充分利用人的主观视觉特性,故给主观因子 $SUBJ$ 这一项赋予了较大的加权系数(w_5),同时为了突出主观感兴趣的对象,头部对象的 $SUBJ$,即 $SUBJ_1$ 取值也较大;在比值加权法中则给重要性不同的对象赋予了大小不同的比值,即取 $W_1 : W_2 : W_3 = 5 : 3 : 1$.基于帧的编码(未采用 Ooba 技术)和 Ooba 编码后的重建图象可见图 6、图 7,其平均 PSNR 可参见表 1、表 2.由表 1、表 2 可见,虽然 Ooba 编码图象 PSNR 比未采用 Ooba 编码图象略低,但是由图 6、图 7 可见,图象的主观视觉质量却得到了提高,且 PSNR 下降和图象质量提高的幅度因不同的序列而略有不同.这也说明了在低比特环境下,客观评价准则并不能作为评价图象编码效果的唯一标准.

表 1 不同带宽分配方法重建的 Claire 图象中 Y、Cb、Cr 3 个对象分量的平均 PSNR(8kbit/s, 8fps)

编码重建图象 带宽分配方法	Y	Cb	Cr
基于帧	31.08	23.20	24.02
RD 法	30.68	22.92	23.81
分析法	30.55	22.97	23.89
比值法	30.65	22.99	23.86

表 2 不同带宽分配方法重建的 Miss America 图象中 Y、Cb、Cr 3 个对象分量的平均 PSNR(10kbit/s, 8fps)

编码重建图象 带宽分配方法	Y	Cb	Cr
基于帧	31.20	27.65	27.39
RD 法	31.12	27.77	27.46
分析法	31.17	27.14	27.78
比值法	31.18	27.77	27.56



图 6 解码图象及脸部放大图(Claire 142 帧)

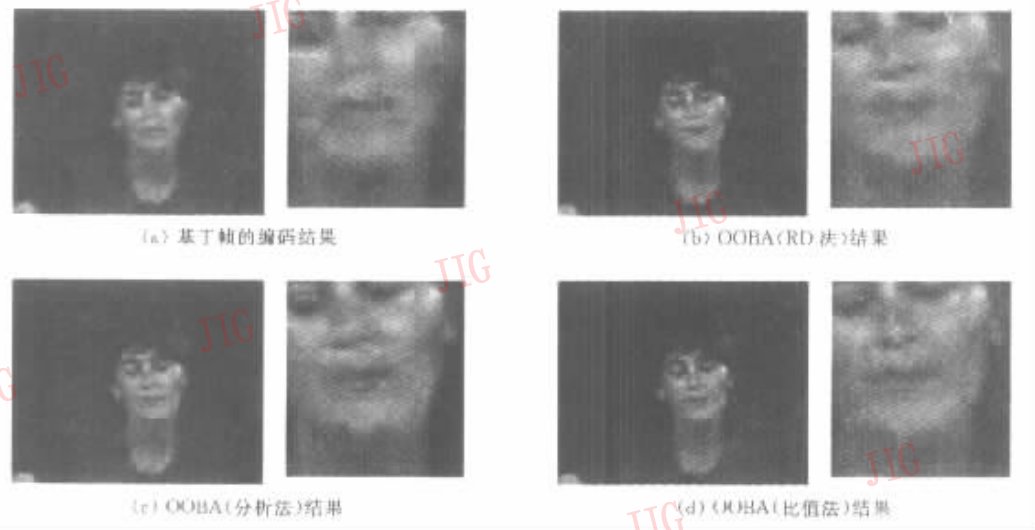


图 7 解码图象及脸部放大图(Miss America 120 帧)

5 结束语

本文提出了一种面向对象分配带宽的视频编码器,并给出了三种多对象带宽分配算法.实验结果表明,这种编码器在低比特率环境下,可以提高编码图象的主观质量.同时,由于该编码器提供了基于对象的图象质量可扩缩性,因此也可以用在诸如视频监控等应用场合,并用来突出感兴趣图象区域的细节.

参 考 文 献

1 翁南钊,蔡德钧. 视频对象分割和两种面向对象的视频编码器.

电子学报,2000,28(10):106~110.

2 Sikora T. The MPEG-4 video standard verification model. IEEE Trans CSVT,1997,7(1):19~31.

3 Neri A *et al.* Automatic moving object and background separation. Singal Processing, 1998,66(2):219~232.

4 Salembier P. Morphological multiscale segment for image coding. Singal Processing, 1994,38(3):359~386.

5 Ballard D H, Brown C M 著. 计算机视觉. 王东泉,徐心平,赵经伦译. 北京:科学出版社,1987.

6 李庆扬,王能超,易大义编. 数值分析(第3版). 武汉:华中理工大学出版社,1986.

7 Vetro A, Sun H, Wang Y. MPEG-4 rate control for multiple video objects. IEEE Trans CSVT,1999,9(1):186~199.

8 谢韬等. 基于物体的编码方法中的多物体速率控制. 通信学报,

1999,20(8):1~6.

- 9 Chiang T, Zhang Y Q. A new rate control scheme using quadratic rate distortion model. IEEE Trans CSVT,1997,7(1): 246~249.
- 10 余鄂西. 矩阵论. 北京:高等教育出版社,1995.

翁南钊 1975年生,1997年获福州大学无线电工程系学士学位,2000年获华中理工大学电子与信息工程系硕士学位. 主要研究领域为视频编码、图象处理、多媒体通信、无线通信等.

蔡德钧 1938年生,华中理工大学电子与信息工程系教授,中国电子学会高级会员. 主要研究领域为随机信号处理、图象压缩及多媒体通信.

刘李杰 1980年生,1999年获华中理工大学信息工程系学士学位,现为华中理工大学电子与信息工程系硕士研究生. 目前主要研究兴趣为多媒体通信、图象、视频编码及视频处理.