

# MPEG-4 中的静止纹理物体编码

王琪 郭雷

(西北工业大学自动控制系, 西安 710072)

**摘要** 新一代的国际标准 MPEG-4 不仅对运动视频有着很好的压缩效果,而且定义了压缩静止图像的部分——visual texture coding (VTC)。MPEG-4 VTC 编码可以对任意形状的静止物体进行编码,而且可以产生渐进的、可扩展的、对误差鲁棒的比特流。本文根据 MPEG-4 VTC 所定义的解码过程,建立了与标准相容的编码框架;从标准要求的各个方面入手,给出了对于小波变换系数的不同扫描算法和不同的量化算法。通过与其他静止图像的压缩标准对比,可以看出实现的 VTC 编码算法不但有良好的率失真性能,而且具有许多有用的特性。

**关键词** MPEG-4 VTC 静止图像 压缩

中图分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2004)04-0477-06

## An Implement of Visual Texture Coding in MPEG-4

WANG Qi, GUO Lei

(Department of Automatic Control, Northwestern Polytechnical University, Xi'an 710072)

**Abstract** The new standard MPEG-4 not only has the excellent compression impact on the motion video, but also defines Visual Texture Coding, which is used for the compression of still image. MPEG-4 VTC can encode arbitrarily shaped still object as well as producing progressive scalable bitstream which is robust to error. An encoding frame consistent with syntax of MPEG-4 VTC is presented. This paper offers different scanning and quantization algorithm according to the syntax of MPEG-4 VTC. Comparison with other compression standard for still image shows that the VTC has better rate-distortion performance and many attractive useful characteristic.

**Keywords** MPEG-4, visual texture coding, still image, compression

## 1 引言

最近几年,数据压缩方法有了令人瞩目的发展。以运动图像数字编码为代表的信号压缩技术已走向成熟。MPEG-1<sup>[1]</sup>、MPEG-2<sup>[2]</sup>、H. 261<sup>[3]</sup>和 H. 263<sup>[4]</sup>等国际标有力促进了运动图像压缩技术的商品化进程。

随着电信业、计算机业和影视业的迅速融合,各种更高的需求不断涌现,如多媒体移动通信、视频点播、多媒体创作工具、虚拟现实、交互式家庭购物等等。与此同时,原先制定的标准如 MPEG-1/2 等其应用面不宽的缺陷日渐显露。于是从 1994 年开始, MPEG 视频专家组正式开始制定 MPEG-4<sup>[5]</sup> 标准, 1998 年发布了 MPEG-4 视频国际标准的第 1 版。

新一代视频压缩标准 MPEG-4<sup>[5]</sup> 不仅对运动视

频有着很好的压缩效果,而且突破了以往 MPEG-2<sup>[2]</sup> 和 MPEG-1<sup>[1]</sup> 只能利用帧内编码对静止图像压缩的限制,在标准中融入了直接对静止纹理物体进行编码的部分,采用性能较为优越的小波变换,代替了视频编码时的 DCT 变换,进行嵌入式的编码。相比于其他静止图像的压缩标准,例如 JPEG<sup>[6]</sup>、JPEG-LS<sup>[7]</sup>、JPEG2000<sup>[8]</sup>, VTC (visual texture coding) 可以对标准形状和任意形状的物体进行编码,并且对编码的方式提供了更大的灵活度和对各种场合的适应性,加入了更强的操纵物体的能力和更鲁棒的误差处理功能,例如误差恢复、误差隐藏。

## 2 VTC 的总体框架

VTC 编码采用性能优越的小波变换,代替了传统的 DCT 变换。整个算法是以 EZW<sup>[9]</sup> 和 ZTE<sup>[10]</sup> 算

法为基础。虽然单精度量化的 ZTE<sup>[10]</sup>算法不能提供嵌入式的特性,但是利用多精度量化的方法,对前次量化的误差进行再编码,就可以实现在应用中特别是无线环境中很有用的嵌入式编码。MPEG-4 的第 4 版对 VTC 算法加进了一些误差恢复的处理方法,使

得 VTC 算法可以更加鲁棒地应用于各种环境中。

为了给用户提供一定程度上的灵活性,标准中没有规定具体编码的细节,图 1 是结合标准的视频部分<sup>[5]</sup>和验证模型<sup>[11]</sup>部分定义的解码过程,给出的整个 VTC 编、解码的框图。

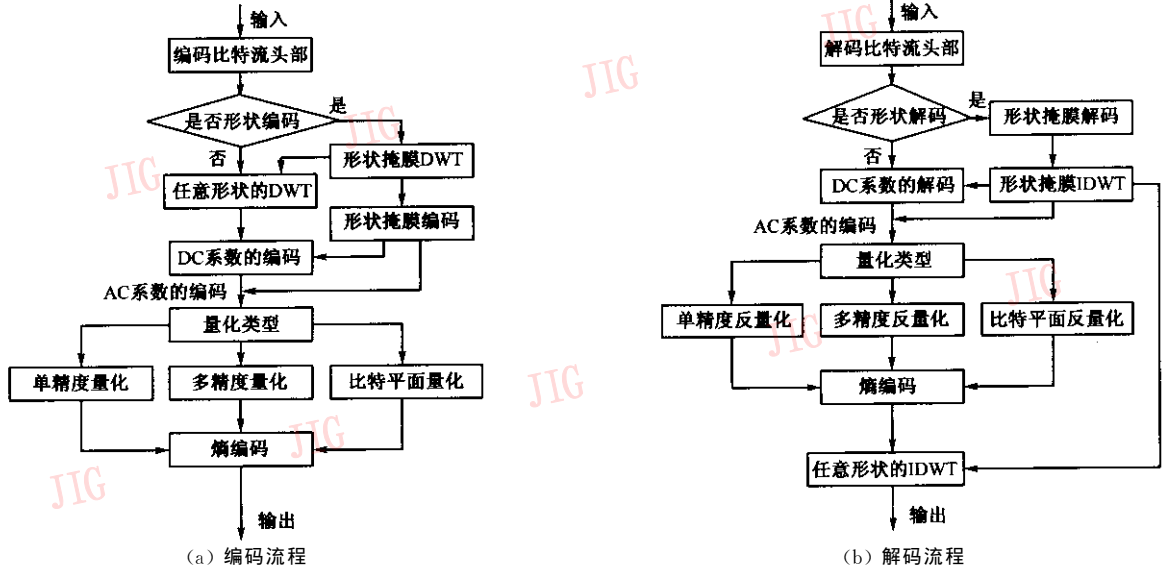


图 1 VTC 编码框图

编码算法的主要处理步骤如下:

(1) 将必要的编码参数按照标准定义的语法编进比特流中;

(2) 根据要求选择是否需要形状编码。如果需要形状编码,只有在形状掩膜内的系数才进行编码。不需要形状编码,则需要编码的图像是规则的矩形框;

(3) 进行形状自适应的 DWT,变换系数分两部分单独处理,其中小波变换的最低频系数,也就是 LL 频带内的系数,称为 DC 系数;而其他部分系数称为 AC 系数;

(4) DC 系数进行预测编码和算术编码;

(5) AC 系数根据要求的量化类型,采用不同的压缩算法;

(6) 根据标准的语法形成比特流输出。

解码过程是编码的逆过程,只是在最后对结果进行形状自适应的 IDWT。

### 3 DC 系数的编码处理方法

DC 系数只是小波分解后的最低频系数,这些系数都在一个频带内,这样对 DC 系数编码只用到

除带内的数据冗余就可以了。其算法步骤如下:

(1) 对 DC 系数进行均匀量化,量化步长以 8bit 的长度编进比特流中。

(2) 对 DC 系数进行基于上下文内容的预测,如果  $V_A, V_B, V_C, V_X$  分别表示像素 A、B、C、X 的值,则预测过程如下所示:

$$\text{if } (|V_A - V_B| < |V_B - V_C|) \quad \text{pred} = V_C;$$

$$\text{else} \quad \text{pred} = V_A.$$

$$V_X = V_X - \text{pred}$$

其中,点 A、B、C 的位置如图 2。如果点 A、B、C 中任何一个不在图像内,它的值被设定为 0。

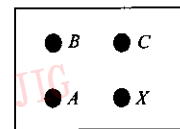


图 2 像素 X 的预测

(3) 对预测的 DC 系数进行偏移,保证编码的 DC 系数全部为正数。其偏移数也就是最小的非正的 DC 系数  $\text{band\_offset\_byte}$  以 8bit 的长度编进比特流中,而偏移后的最大值  $\text{band\_max\_byte}$  也以 8bit 的长度编进比特流中。

(4) 编码偏移后的 DC 系数。先根据  $\text{band\_}$

max\_byte 计算出共有多少个比特平面,然后为每个比特平面建立一个 2 符号的算术编码器。接下来从最有意义的比特平面开始,按平面顺序对每个 DC 系数进行自适应的算术编码。

解码过程是编码过程的逆过程。

### 4 AC 系数的编码处理方法

AC 系数是根据不同的系数扫描方法和不同的量化策略来分别处理的。其中,系数扫描方法有小波树的深度优先方式和子带接子带的方式两种。量化策略有单精度量化、多精度量化和比特平面量化 3 种。其中,单精度和多精度量化是以 ZTE<sup>[10]</sup>算法为基础的,而比特平面的量化是以 EZW<sup>[9]</sup>算法为基础的。

#### 4.1 系数扫描方法

在 AC 系数的编码过程中,标准定义了两种不同的系数扫描方法:小波树的深度优先方式和子带接子带的方式,将扫描排列的 AC 系数进行熵编码。

在深度优先方式中,要求编码下一棵树之前,这棵树的所有系数要编码完毕,并且同一棵树上位于同一层的系数,也是要先编码一个系数的所有后代之后才编码其他系数。同一层的扫描顺序是按照光栅扫描的顺序。图 3 是其扫描示意图,其中的数值代表扫描的顺序。

子带接子带的方式是从低频到高频,同一层按照 HL-LH-HH 的顺序扫描。图 4 是其扫描示意图,其中的数值代表扫描的顺序。

0	1	4	67	5	10	68	73	6	7	11	12	69	70	74	75
2	3	130	193	15	20	78	83	8	9	13	14	71	72	76	77
25	88	46	109	131	136	194	199	16	17	21	22	79	80	84	85
151	214	172	235	141	146	204	209	18	19	23	24	81	82	84	85
26	31	89	94	47	52	110	115	132	133	137	138	195	196	200	201
36	41	99	104	57	62	120	125	134	135	139	140	197	198	202	203
152	157	215	220	173	178	236	241	142	143	147	148	205	206	210	211
162	167	225	230	183	188	246	251	144	145	149	150	207	208	212	213
27	28	32	33	90	91	95	96	48	49	53	54	111	112	116	117
29	30	34	35	92	93	97	98	50	51	55	56	113	114	118	119
37	38	42	43	100	101	105	106	58	59	63	64	121	121	126	127
39	40	44	45	100	103	107	108	60	61	65	66	123	124	128	129
153	154	158	159	216	217	221	222	174	175	179	180	237	238	242	243
155	156	160	161	218	219	223	224	176	177	181	182	239	240	244	245
163	164	168	169	226	227	231	232	184	185	189	190	247	248	252	253
165	166	170	171	228	229	233	234	186	187	191	192	249	250	254	255

图 3 小波系数的深度优先扫描

0	1	4	67	16	17	28	29	64	65	68	69	112	113	116	117
2	3	10	13	18	19	30	31	66	67	70	71	114	115	118	119
5	8	6	9	40	41	52	53	72	73	76	77	120	121	124	125
11	14	12	15	42	43	54	55	74	75	78	79	122	123	126	127
20	21	32	33	24	25	36	37	160	161	164	165	208	209	212	213
22	23	34	35	26	27	38	39	162	163	166	167	210	211	214	215
44	45	56	57	48	49	60	61	168	169	172	173	216	217	220	221
46	47	58	59	50	51	62	63	170	171	174	175	218	219	222	223
80	81	84	85	128	129	132	133	96	97	100	101	144	145	148	149
82	83	86	87	130	131	134	135	98	99	102	103	146	147	150	151
88	89	92	93	136	137	140	141	104	105	108	109	152	153	156	157
90	91	94	95	138	139	142	143	106	107	110	111	154	155	158	159
176	177	180	181	224	225	228	229	192	193	196	197	240	241	244	245
178	179	182	183	226	227	230	231	194	195	198	199	242	243	246	247
184	185	188	189	232	233	236	237	200	201	204	205	248	249	252	253
186	187	190	191	234	235	238	239	202	203	206	207	250	251	254	255

图 4 小波系数的子带接子带扫描

## 4.2 单精度量化(single-quant)

单精度的量化以 ZTE<sup>[10]</sup> 算法为基础。单精度的量化没有空域可伸缩性和 SNR 伸缩性。其具体实现算法如下:

(1) 初始化 将每个节点分为两种状态: 叶子节点(LINIT)和非叶子节点(INIT)。叶子节点是没有后代的节点,而非叶子节点就是还有后代的节点。对每个节点进行均匀量化,量化步长以 8bit 的长度编进比特流中。量化后绝对值最大的 AC 系数对应的比特平面数 max\_bitplane 也以 5bit 的长度编进比特流中,指导其后的熵编码。

(2) 标记每个系数的类型 标准定义了 4 种类型,表 1 给出了这 4 种零树符号到算术编码索引的映射。

表 1 零树符号的索引

索引	符号
0	IZ
1	VAL
2	ZTR
3	VZTR

其中,IZ 表示孤立零点,自己的值为 0,但是其后代和其祖先都有非 0 值;VAL 表示自己的值非 0,而且还有非 0 子孙;ZTR 表示当前节点为零树的根,此时不需要进一步扫描零树,因为它的子孙节点值都为 0;VZTR 表示自己节点值非 0,但是 4 个子孙节点都是零树根 ZTR。

(3) 建立算术编码器 为待编码的量化后的 AC 系数建立算术编码器,根据 max\_bitplane 为每一个比特平面建立一个 2 个符号的算术编码器来编码量化后有意义的值;为需要编码的正负符号建立一个 2 个符号的算术编码器;根据节点的状态对两种不同的节点建立两个算术编码器,叶子节点由于没有后代,只会有 ZTR 和 VZTR 两种符号类型,这样为叶子节点只需要建立 2 个符号的算术编码器;非叶子节点会有上述不同的 4 种类型,建立 4 个符号的算术编码器。

(4) 熵编码 利用建立的算术编码器分别对零树符号、值、符号进行熵编码。

## 4.3 多精度量化(multi-quant)

多精度量化是单精度量化之上的进一步扩展。它是对单精度量化的误差进行编码,以更好地改进质量。与单精度的量化相比,多精度的量化可以更好

地进行可伸缩的编码,根据用户的需求,多精度的量化不但可以进行空域伸缩的编码,而且可以进行质量伸缩编码,从而适应了更多的需求。多精度量化的算法框图如图 5 所示。

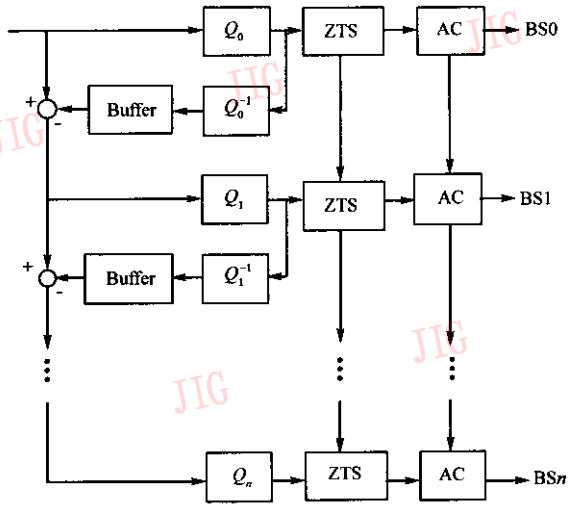


图 5 多精度量化示意图

算法步骤如下:

(1) 仍然进行单精度的量化过程。

(2) 根据编码的类型对上一次的量化状态进行更新。

(3) 根据前一次的标记和更新的系数状态重新对新的量化零树进行标记。由于小波树自身或者后代有可能对新的量化步长表现得有意义,这样可能产生 8 种状态: S\_ZTR、S\_ZTR-D、S\_IZ、S\_VZTR、S\_VAL、S\_LZTR、S\_LZTR-D、S\_LVZTR。对应每种状态可能有不同的符号类型: S\_ZTR、S\_ZTR-D 状态有 IZ、VAL、ZTR 和 VZTR 4 种类型; S\_IZ 状态有 IZ、VAL 2 种类型; S\_VZTR 有 VZTR、VAL 2 种类型; S\_VAL 只有 VAL 1 种类型; S\_LZTR、S\_LZTR-D 有 VZTR、ZTR 2 种类型; S\_LVZTR 只有 VZTR 一种类型。

如果某个系数在上次扫描的时候是 VZTR 或者 VAL,那么即使此时量化的值不是有意义的,它仍然作为有意义的值处理,并且不编码此时的类型值,而认为该系数就是 VAL 类型的,只对其此时的量化结果进行编码。

(4) 对新的标记进行算术编码。

(5) 重复第 2 步到第 4 步,直到完成要求的伸缩性为止。

注意每次量化扫描后需要对以前为 IZ 或者 ZTR 的系数进行精化处理。

#### 4.4 比特平面量化(bilevel\_quant)

4 比特平面的量化以 EZW<sup>[9]</sup>算法为基础。其算法原理是:

(1) 如果在当前比特平面中,所有系数都表现为没有意义,则直接输出一个 ZTRZ,表明零树根。

(2) 如果在当前比特平面中,所有的叶子系数都没有意义,而根节点的系数有意义,此时输出 ZTRV,表明带值的零树根。

(3) 如果在当前比特平面中,有一个叶子系数表现有意义,如果该根系数没有意义,则根系数的符号就是 IZER;如果有意义,根系数的符号就是 IVAL。同时进一步扫描其叶子系数。此时又产生了 4 个小零树,仿照上面的方法继续进行。

(4) 每进行一次比特平面的处理后,对上一个比特平面产生的有意义系数进行提精处理。

(5) 在实际扫描中需要注意的是,一个系数只有在第 1 次认为是有意义的时候才输出其符号,并且不参与当前系数的判断。

### 5 任意形状的变换和误差处理方法

与过去各种标准相比,MPEG-4 标准的突出特点是对视频数据采用基于物体的操作、存取及传输。VTC 编码可以进行任意形状的小波变换。对任意长度的序列进行小波变换主要有以下两个难点<sup>[12]</sup>:一是如何对任意长度的有限长序列进行边界扩展,提高变换的效率;二是如何在任意位置对任意长度的序列下采样,保证恢复的准确性。Li 等人<sup>[12]</sup>根据滤波器的正交性和对称性采用不同的边界扩展方式;根据应用的目的是提高零树编码的效率还是提高信号处理的效率应用不同的下采样策略,从而很好地解决上面的两个问题,并最终被 MPEG-4 专家组采用,确定为解码的反变换算法。

为了使 VTC 算法可以更加鲁棒地应用于各种环境中,尤其是低比特率容易发生误差的情况下,MPEG-4 的第 2 版在 VTC 算法中加进了一些误差恢复的处理方法。主要是利用数据分割和再同步的策略,将小波变换后 AC 和 DC 系数进行打包分割处理,在每一个数据包的前部加入标记表示数据的重新开始。这样使得整个算法可以鲁棒地应付发生的误差。

另外,MPEG-4 第 2 版中为了减少大图像运算的复杂度,引入了分块处理(tiling)的办法,将大图像分成许多同样大小的小图像,用同样的方法进行处理。

## 6 与其他标准的比较以及性能分析

表 2 比较了 JPEG、JPEG2000 和 VTC 编码可以提供的各项功能。

表 2 各标准功能列表

	JPEG2000	VTC	JPEG
无损压缩性能	+++	+++	+
有损压缩性能	+++++	++++	+++
渐进的比特流	+++++	++++	++
感兴趣区域编码	+++	+	-
任意形状的编码	-	+++	-
随机获取	++	+	-
低复杂度	+	++++	+++++
误差恢复	+++	+++	++
位率控制	+++	++	-

注: + 表示具有该项功能, + 越多表示效果越好, - 表示不支持该功能

可以看出 JPEG2000 和 VTC 由于面向应用的不同而各有千秋,而制定较早的 JPEG 无论在哪个方面都很落后。在有损压缩性能、感兴趣区域编码、随机获取以及码率控制方面,JPEG2000 较为优越一些,这是因为 VTC 中没有计算复杂度很高的位率-偏差估计,所以最终性能稍为逊色一些,但是带来的整体实现复杂度较低。在较低比特率下,VTC 的性能与 JPEG2000 相差不是很多,结果非常接近。另外 VTC 可以对任意形状的物体进行编码,并且利用这个特性可以对感兴趣区域进行重点编码。如果 JPEG2000 和 VTC 的量化步长取为 1,就都可以很容易地实现无损压缩。而采用多精度量化和比特平面量化都可以产生渐进的比特流,比特平面的量化也可以带来方便的位率控制。另外在第 2 版的 MPEG-4 标准中注重了误差处理能力,加入了很多的误差处理方法,以致于 VTC 编码有着很好的误差恢复性能。这样,从各个方面比较,VTC 的性能都是比较令人满意的。

图 6 是各种标准的压缩比特率和 PSNR 示意图,其中 JPEG2000 R 表示采用无损压缩,JPEG2000 NR 表示采用有损压缩。可以看出,虽然有损压缩的 JPEG2000 的性能是最好的,其 PSNR 是最高的,但这主要表现在超过 0.5bpp 的较高比特率的情况;在较低比特率下,MPEG-4 VTC 的性能与 JPEG2000 NR 是很接近的,这主要是因为 JPEG2000 采用的是寻找最佳的位率/偏差比,以致于有了最佳的性能,并且符合于各种情况,但是

MPEG-4 VTC 更注重的则是低比特率的场合。另外,在任何比特率下 MPEG-4 VTC 的性能都优于 JPEG2000 R。而 JPEG 的效果与任何一个标准算法都要逊色不少。

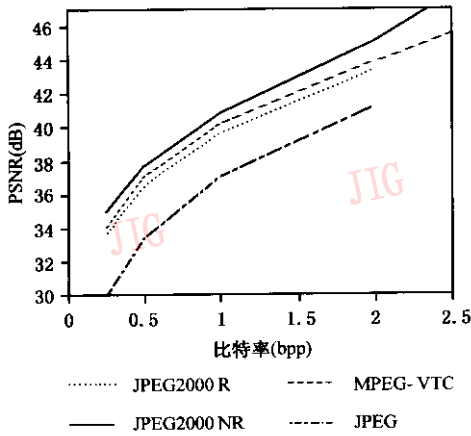


图 6 MPEG-4 VTC 与其他标准的对比

## 7 结 论

根据 MPEG-4 标准所定义的解码过程,建立了与标准相容的编码算法。另外从标准要求的各个方面入手,给出了对小波变换系数的不同扫描方法和不同量化方法,以及实验结果。实验结果证明实现的 VTC 编码算法不但具有良好的位率/偏差性能,而且具有许多优良的特性。例如对任意形状的物体进行压缩,有着很好的误差处理机制,可以渐进传输等等。这样 MPEG-4 VTC 可以独立地作为一个静止图像的压缩工具,具体应用到实际工业当中。

### 参 考 文 献

- 1 ISO/IEC CD 11172, Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to 1.5Mbits/sec [S]. Nov., 1991.
- 2 ISO/IEC CD 13818, Generic Coding of Moving Pictures and Associated Audio Information[S]. Nov., 1994.

- 3 CCITT SG XV Recommendation H. 261. Video Codec for Audiovisual Services at  $p \approx 64$  kbit/s[S]. Aug., 1990.
- 4 Draft ITU-T Recommendation H. 263. Video Coding for Low Bitrate Communication[S]. 1995.
- 5 ISO/IEC 14496-2, Final Draft International Standard, Information Technology-Generic Coding of Audio-Visual Objects, Part 2: Visual[S], Dec. 1998.
- 6 Wallace G K. The JPEG still picture compression standard[J]. Communication of the ACM, 1991, **34**(4):30~44.
- 7 ISO/IEC 14495-1:1999, Information Technology-Lossless and Near-Lossless Compression of Continuous-Tone Still Images: Baseline[S]. Dec., 1999.
- 8 ISO/IEC FCD 15444-1, JPEG2000 Part I Final Draft International Standard[S]. Aug., 2000.
- 9 Shapiro J M. Embedded image coding using zerotrees of wavelet coefficients[J]. IEEE Transactions on Signal Processing, 1993, **41**(12):3445~3462.
- 10 Martucci S A, Sodagar I, Chiang T, et al. A zerotree wavelet video coder[J]. IEEE Transactions on Circuits and Systems for Video Technology[J], 1997, **7**(1):109~118.
- 11 ISO/IEC JTC1/SC29/WG11 N3312, MPEG-4 Video Verification Model Version 16.0[S]. March 2000.
- 12 Li S, li W. Shape adaptive discrete wavelet transform for arbitrarily shaped visual object coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2000, **10**(8):725~743.



王 琪 1977 年生,1998 年毕业于西北工业大学航海工程学院,2001 年于西北工业大学获模式识别与智能系统硕士学位,现攻读西北工业大学获模式识别与智能系统博士学位。主要研究兴趣为静止图像和视频序列的压缩、视频分割、运动估计以及视频传输。



郭 雷 1958 年生,教授,博士生导师,1996 年国家杰出青年科学基金获得者。长期从事视觉、图像处理 and 神经计算等智能信息处理方面的基础研究工作。目前已出版专著 2 部,发表论文 60 篇。