

基于曲率尺度空间的 VOP 形状编码算法

石旭利 张兆扬

(上海大学通信与信息工程学院, 上海 200072)

摘要 首先对曲率尺度空间(CSS)算法进行改进,并用改进的 CSS 算法对任意形状物体的形状信息进行特征点的分层提取;其次,用自适应算术编码算法对固定长度的八分圆编码算法进行改进,并对所提取的特征点进行算术编码。实验结果表明,这种新的基于特征点提取的形状编码方法比 MPEG-4 校验模型中基于上下文的算术形状编码算法,在 D_s 相同的条件下,主观质量更好,且压缩比更高。

关键词 MPEG-4 形状编码 VOP 形状表达 多尺度空间 特征点 角

中图分类号: TN911.22 TM919.81 **文献标识码**: A **文章编号**: 1006-8961(2002)10-1071-06

CSS Based Vop Shape Coding Algorithm

SHI Xu-li, ZHANG Zhao-yang

(School of Communication and Information Engineering, Shanghai Univ., Shanghai, 200072)

Abstract In this paper, a new shape coding algorithm called curvature scale space based shape coding (CSS) is proposed. For dealing with video image, we regard a shape as a set of points that are parameterized by arc length. Then, the evolution of curve at different resolution levels in Gaussian scale space is achieved by convoluting the curve with Gaussian kernel. By calculating the curvature, the feature-points including significant information of shape contour can be found. But for shape coding, this shape representation will not very efficient when the shape consists of arcs. So, the CSS is modified, then extract the feature points of shape using the modified CSS algorithm. The octants algorithm is modified and encode the feature points by using adaptive arithmetic code algorithm. Simulation results show that our proposed algorithm has higher compress ratio, better subjective quality compared with other results from the CAE in MPEG-4.

Keywords MPEG-4, Shape coding, Shape representation of VOP, Multiscale space, Feature point, Corner

0 引言

MPEG-4 是 ISO 发表的最新视、音频标准^[1]。MPEG-4 支持基于内容或基于视频对象(VO)的任意形状编码,是其区别于传统编码标准的许多新特点中最核心的内容。目前,国外在形状编码上的研究主要有:MPEG-4 校验模型中,基于上下文的算术形状编码(CAE)^[2,3];MPEG-4 核心实验中,基于顶点的形状编码(VBSE)^[4],以及在此基础上改进的基于顶点的形状编码方法^[5,6];色度键、基线等其他形状编码方法。其中,基于顶点的形状编码方法和基

于上下文的算术形状编码方法最有竞争力,但各有缺陷。基于上下文的算术形状编码方法首先将 VOP (视频对象平面,即视频对象在给定时间的“快照”)掩模分成 16×16 宏块,根据宏块中像素亮度值为 255 的不透明像素个数选择不同的编码模式,这种基于块的技术,虽然可以和框架中的基于块的纹理编码最为兼容,但却无法避免帧内、帧间编码中存在的严重方块效应;基于顶点的形状编码方法在其压缩效率上显示出优越性,特别是在有损编码时,效率高,但其多边形顶点探测算法较为复杂。为此,提出一种基于改进的曲率尺度空间(CSS)的新的 VOP 形状编码算法。算法的贡献主要有以下几点:首先,

用于形状检索的 CSS 算法进行曲线平滑时,必然造成曲线总长度的缩减,这使得曲线上点的位置偏离其在原始曲线上的位置,从而因位置信息无效而不能用于形状编码,而采用跟踪算法对 CSS 进行改进,可获得位置信息,使之能用于形状编码;第二,由于 CSS 将局部曲率最大点作为特征点,必然使形状曲线中,由圆弧组成的部分所提取的特征点数为零,因此采用棋盘距离对 CSS 进行改进,并提取逼近点;第三,文献[5]、[6]中所用的顶点编码算法均来自于文献[4]中基于八分圆的固定长度的顶点编码方法,采用自适应算术编码算法对其进行改进,并用改进的八分圆算法对所提取出来的特征点进行熵编码,可进一步提高压缩效率。

1 用改进的 CSS 算法进行特征点的提取

1.1 CSS 算法

曲率尺度空间 CSS 算法首先由 Mokhtarian 等人提出^[7],主要用于图象的远程检索中。首先用弧长将形状曲线参数化,设所要处理的形状曲线为弧长以步长 1 递增的离散点的集合,即

$$\{f(s) = (x(s), y(s)) | 1 \leq s \leq N\} \quad (1)$$

其中, N 为总弧长,即形状曲线的总像素数。其次用高斯函数对式(1)进行平滑处理,即用高斯函数对曲线坐标进行卷积

$$\begin{aligned} x(s, t) &= x_0(s) * g(s, t) \\ y(s, t) &= y_0(s) * g(s, t) \end{aligned} \quad (2)$$

式中, t 代表尺度参数,它与高斯函数的宽度 δ 满足关系式: $\delta = \sqrt{2t}$,由高斯函数的性质以及卷积运算的性质可使式(2)中的 $x(s, t), y(s, t)$ 满足下列方程

$$\begin{aligned} \frac{\partial x(s, t)}{\partial t} &= \frac{\partial^2 x(s, t)}{\partial s^2} \\ \frac{\partial y(s, t)}{\partial t} &= \frac{\partial^2 y(s, t)}{\partial s^2} \end{aligned} \quad (3)$$

式(3)称为几何热方程。由式(2)可见,随着尺度参数 t 的改变,上述卷积过程获得一系列由原始曲线通过曲率形变演化而来的曲线。文献[7]中的证明表示,凡是满足几何热方程的一系列曲线必然具有下列性质:

- (1) 包含次序不变。若形状 A 在形状 B 的内部,则曲率形变以后,这种包含关系不变。
- (2) 在曲率形变过程中,不会引起曲线的交叉,

即所有的曲线随着曲率形变操作逐渐变为凸曲线,并最终成为一个圆点。

(3) 局部曲率绝对值最大点(角点)和局部曲率零点(拐点)的数量,在演化过程中单调减少。

(4) 此过程可并行实现。

(5) 曲线总长逐渐缩短。

因此高斯函数对曲线坐标的卷积,同样具有上述性质,性质 1、性质 2 保证了用高斯函数对曲线进行卷积后,所得到的一系列曲线仍保持各自的独立性,对于复杂形状的物体来说,仍可利用上述理论进行特征点的提取,即上述算法适用于任意形状;性质 3 表明利用高斯函数对曲线进行卷积后,可形成曲线的多尺度空间表示。原始曲线表示为最精细尺度的空间曲线,随着尺度的增加,曲线的某些细节信息被去除,而保留基本的信息。将这种层层“剥离”的操作进行下去,直到在要求的粗糙尺度上,获得需要的曲线最基本的信息。称这样形成的一系列曲线所构成的空间为原始曲线的多尺度空间。用曲率函数计算每一点的曲率,称局部曲率负最小和局部曲率正最大点为特征点“角”。以弧长 s 为参数,尺度为 t 的曲线曲率方程为

$$k(s, t) = \frac{x_s(s, t)y_{ss}(s, t) - x_{ss}(s, t)y_s(s, t)}{((x_s(s, t))^2 + (y_s(s, t))^2)^{3/2}} \quad (4)$$

式中

$$x_s(s, t) = \frac{\partial}{\partial s}(x(s) * g(s, t))$$

$$x_{ss}(s, t) = \frac{\partial^2}{\partial s^2}(x(s) * g(s, t))$$

利用式(4)在高斯函数的某一 t 下,对多尺度空间中的某一条曲线进行特征点“角”的提取,在形状剧变部分获得密集的点,而在平坦部分获得稀疏的点集,从而既可以保留重要的信息,又能用较少的比特表示图象。同时从整个多尺度空间来考虑,这种层层“剥离”的操作可以从大到小改变高斯函数的 t 值,从而可以获得由少到多的特征点,并且 $t=L+1$ 时所获得的特征点等于 $t=L$ 时的特征点与新增加的特征点之和,以此类推,可以获得对于原始曲线的分层近似,这个特点非常适合于实现 MPEG-4 视频,在网上传输时所要求的视频流的分层表达,以便于接收者根据实际条件,对视频流进行有选择的接收和解码。如图 1 所示,传输过程中,根据带宽和接收端的计算能力、存储能力等,可以选择仅对基层信息解码,即仅对流 1 进行传输和解码,也可选择对流 2

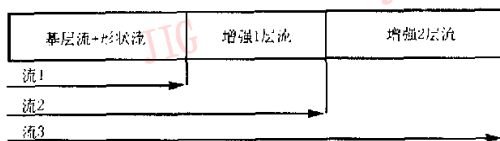


图 1 视频流的分层表示

进行传输和解码;也可选择对流 3 进行传输和解码。基于特征点提取的形状编码算法优于基于下抽样的 CAE 方法;MPEG-4 校验模型中,由下抽样实现的形状信息的码率控制,在分层上不如 CSS 灵活。基于上述原因认为,在用 CSS 进行形状编码的算法中,特征点的提取操作是可行的。很显然,这种新的 VOP 形状编码算法与基于顶点的形状编码算法的不同之处在于,它避免了顶点提取过程中,大量的循环逼近计算^[1-6]。

1.2 特征点的定位跟踪算法

CSS 算法的最大缺陷在于性质 5,即随着式(2)中尺度参数 t 的增加,所形成的曲线与原始曲线相比,越来越光滑,越来越短。这种长度的缩减,使得曲线上点的位置偏离其在原始曲线上的位置,而对形状编码来说,这种现象使得特征点的位置信息无效。在基于特征点提取的形状分层多边形表示中,特征点的准确定位是至关重要的。特征点的定位跟踪算法如下:在较粗的尺度 t 上,将特征点“角”提取出来,得到角点集合 P_t ;减小尺度的值,在较细的尺度 $t-1$ 上,将特征点“角”提取出来,得到角点集合 P_{t-1} ;根据性质 3,在准确定位后的两集合 C_{P_t} 与 $C_{P_{t-1}}$ 之间,有 $C_{P_t} \subset C_{P_{t-1}}$,其中, C_{P_t} 表示 P_t 集合中的点做准确点后形成的点集合,设 $A(X_A, Y_A) \in P_t$,在 P_{t-1} 中寻找与 A 距离最近的点,若 $B(X_B, Y_B) \in P_{t-1}$,满足 $D(A, B) = \min\{D(A, B_j); B_j \in P_{t-1}\}$,则将 A 点的位置纠正为 $A(X_B, Y_B)$,其中 $D(A, B_j)$ 为 A, B_j 两点之间的欧几里德距离。按上述方法减小尺度 $t-1$ 的值,并按由粗糙到细节的方法进行跟踪,可以得到 A 点在原始曲线上的位置 $A(X_0, Y_0)$ 。对 P_t 中的所有点进行上述跟踪运算则可以得到矫正的点集 C_{P_t} 。同样对集合 P_{t-1} 进行尺度从 $t-1$ 到 0 的跟踪,可以得到尺度上的点集 $C_{P_{t-1}}$ 。因此利用尺度值的改变,可以得到一系列尺度为 $t, t-1, \dots$ 等的以特征点“角”为顶点的多边形,这些多边形在尺度值较大时,与原始曲线的差别较大,但压缩比高,适合于低比特率的形状编码;在尺度较小时,与原始形状曲线的误差小,直至完全表示原始的形状。适合用于无损或失真

较小的形状编码。需要说明的是,尺度 t 并不一定按照整数递减。

1.3 用棋盘距离为参数对圆弧进行分层逼近

上述算法,当原始曲线由若干个圆弧组成时,由于算法的准则是提取局部曲率最大点,所以其并不能将逼近点 C 提取出来,如图 2 所示,逼近点与圆弧的弦所组成的三角形是圆弧上所有点与弦所组成的三角形中面积最大的,即 C 到弦 AB 的距离最大,所以为了将逼近点 C 提取出来,采用棋盘距离为参数, P, C 两点的棋盘距离定义为两点之间 X 坐标之差的绝对值与 Y 坐标之差的绝对值中的最大值

$$d(P, C) = \max(|X(P) - X(C)|, |Y(P) - Y(C)|)$$

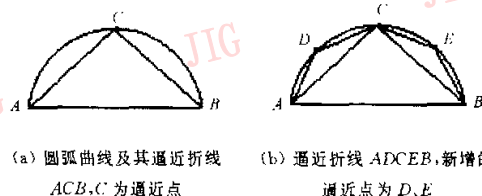


图 2

首先按照弧长步长为 1 递增的顺序来选择圆弧上的点,例如以 A 点为起点,计算圆弧上各点到弦的距离,将前一个点到弦的距离与当前点到弦的距离进行比较,直到距离由增加转为递减时,将前一个点作为逼近点,如图 2(a)所示,得到逼近折线 ACB ,并将逼近点到弦的距离作为逼近程度参数 γ_0 的值,以逼近折线 ACB 的 AC 和 BC 为弦,在弦所对应的弧上求逼近点,重复上述过程,获得逼近点 D, E ,并得到逼近参数 γ_1 的值,如图 2(b)所示,进而得到折线 $ADCEB$ 。显然,所得折线更接近于圆弧曲线;以此类推,直到获得逼近参数 $\gamma_2, \gamma_3, \dots, \gamma_N$ 和对应的逼近点,同时,逼近点也需要进行定位跟踪。上述求逼近点的过程,由于随着逼近参数的减小,逼近点数单调递减,因此可实现逼近点的分层表示。这种逼近点和上述 CSS 算法所求得特征点“角”按照在原始形状曲线上,以弧长为参考的先后顺序排列后,形成了改进 CSS 算法特征点的集合。

1.4 特征点提取算法总流程

用式(2)对形状曲线进行平滑,以高斯函数的参数 t 为尺度,从小到大改变 t 的值,就可以得到一系列细节信息逐渐减少的形状。

在以 t 值为参数的曲率尺度空间中,用式(4)计算这些不同尺度的形状中每一点的曲率。通过选择

局部曲率绝对值最大点进行特征点“角”的提取。但是由于在尺度变化过程中引起了曲线的总长度变短,因此“角”点提取过程中同时要运用定位跟踪算法进行“角”点的准确定位。

在以 t 值为参数的曲率空间中,对每一个形状曲线进行圆弧检测。如果组成形状的曲线中有圆弧存在,则应用“逼近点”搜索算法进行“逼近点”的提取。将提取的逼近点和在相同尺度 t 下提取的“角”点,按照弧长由小到大的顺序进行排序,形成特征点集合 F 。

设 N 为特征点的个数,则有

$$F = \{f_i(x, y) : i \in N\}$$

总的流程如图3所示。

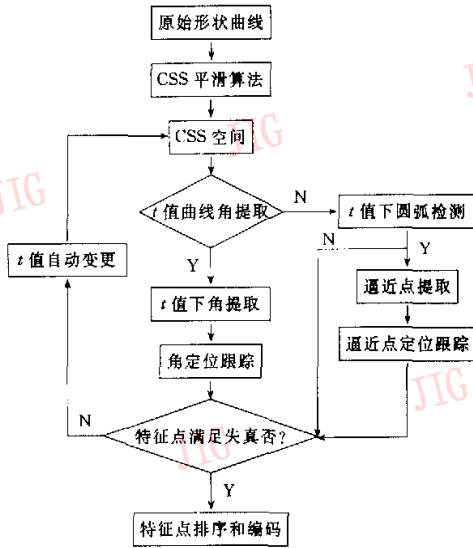


图3 特征点提取总流程

2 特征点编码算法

由于特征点的编码算法是对特征点的位置信息进行编码传输,而文献[4]中的算法是对多边形顶点的位置进行编码,所以可以文献[4]的算法为基本算法,并对其进行修改。

2.1 基本算法

首先,设编码坐标之差的绝对值为 X_F, Y_F 。分别计算第 i 个特征点和第 $i-1$ 个特征点 x 坐标及 y 坐标之差 X_i 和 Y_i ,改变 i 值,共有 $N-1$ 个 x 坐标的差值 X_i ,和 $N-1$ 个 y 坐标的差值 Y_i ,求出 $N-1$ 个 X_i 的绝对值的最大值 X_{max} 和 $N-1$ 个 Y_i 的绝对值的最大

大值 Y_{max} ,并在文献[4]中给出的范围和码字表中进行查找,选择 X_F 和 Y_F ,使得编码 X_{max}, Y_{max} 的比特数最少。对 X_F 和 Y_F 进行固定长度 3bit 的编码。

其次,对坐标之差 (X_i, Y_i) 进行编码。编码之前先对坐标之差 (X_i, Y_i) 进行预处理,将初始特征点的绝对坐标进行编码。将平面坐标系按照以下方法分为 8 个区间,形成一个八分圆。其系数定义为:当特征点坐标之差的绝对值 (X_i, Y_i) 满足 $0 \leq Y_i \leq X_i$ 时,八分圆系数为 0,以此类推,总共有 8 个八分圆系数 0,1,2,3,4,5,6,7。如图4所示。

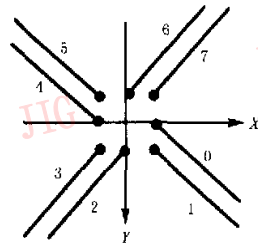


图4 八分圆系数定义

按照上述定义,确定每一个坐标之差 (X_i, Y_i) 的八分圆系数。确定每一个坐标之差 (X_i, Y_i) 中的主分量和次分量,其中,主分量为坐标之差 (X_i, Y_i) 的 x 和 y 分量中的较大分量,而次分量为坐标之差 (X_i, Y_i) 的 x 和 y 分量中的较小分量;然后按照下列步骤对坐标之差 (X_i, Y_i) 进行编码:

(1) 按照八分圆的定义,计算每一个特征点所属的八分圆系数。

(2) 对八分圆系数,按照概率分布计算码字,并进行固定长度为 3bit 的编码。

(3) 对主分量进行编码,若八分圆系数为 0,3,4,7,则相对所属八分圆起点的幅值对 X_i 进行码长为 X_F 的编码;若八分圆系数为 1,2,5,6,则相对所属八分圆起点的幅值对 Y_i 进行码长为 Y_F 的编码;

(4) 对次分量进行编码;若八分圆系数为 0,3,4,7,则相对八分圆起点的幅值,对 y 进行码长为 $\min\{Y_F+1, \log_2(|X_i|+1)\}$ 的编码;若八分圆系数为 1,2,5,6,则相对八分圆起点的幅值,对 x 进行码长为 $\min\{X_F+1, \log_2(|Y_i|+1)\}$ 的编码。

2.2 改进算法

对上述算法中的步骤 2,3,4,用自适应算术编码方法进行改进,把八分圆系数 0,1,2,3,4,5,6,7 出现的过程看作马尔可夫过程,则可以用其出现的概率来驱动算术编码器;同样可用主分量和次分量

值出现的概率来驱动算术编码器. 具体步骤如下:

首先,用区间 $[0, 32]$ 和均匀分布的概率初始化编码器,改进步骤 2 时,输入为八分圆系数;改进步骤 3 时,输入为主分量符号;改进步骤 4 时,输入为次分量符号. 这里统称为输入系数. 用下式计算算术编码器编码过程中的上下边界

$$high = low + rang \times symfre[index]$$

$$low = loe + rang \times symfre[index + 1]$$

其中 $high, low$ 分别表示上下边界, $symfre[index], symfre[index + 1]$ 分别表示与上下边界对应的概率, $index$ 为输入系数在概率表中的序号.

根据 $high, low$ 值的不同,分 3 种情况分别进行操作:

(1) 若 $high \geq 16$, 输出比特 1, 输出 $bittofollow$ 个 0, 且 $low = low - 16, high = high - 16$;

(2) 若 $high < 16$, 输出比特 0, 输出 $bittofollow$ 个 1;

(3) 若 $low \geq 8$ 且 $high < 24$, 使 $bittofollow + 1, low = low - 8, high = high - 8$;

以上步骤要进行 $low = low \times 2, high = high \times 2$ 的操作.

当 3 种情况都不满足时, 停止运算, 并对概率表进行如下修改:

For ($I = index; I \geq 0; I--$)

$symfre[I]++$

修改后的概率表作为下一次输入的输入符号进行算术编码的依据.

使用自适应算术编码算法可达到更高的压缩比.

3 实验结果

以改进的 CSS 为依据进行算法的实现. 图 5(a) 为使用 MPEG-4 标准 352×288 的 Bream 序列进行轮廓提取后的原始轮廓, 图 5(b)、(c)、(d)、(e)、(f) 为高斯函数的参数 $t=1, 1.4, 2, 3, 5$ (这里 t 参数的选取, 由所要求编码的图象的失真率来确定, 对一般的曲线, $t=8$ 可使其足够光滑) 时, 对原始序列进行平滑处理的结果, 显然, 原始图象的细节信息逐渐减少. 图 6 为使用改进的多尺度理论实现的特征点的提取结果, 图 6(a)、(b)、(c) 分别对应高斯函数的参数 $t=1, 1.4, 2$. 图 7 所示为使用改进的 CSS 算法与 MPEG-4 校验模型中的 CAE 算法的比较, 以 MPEG-4 核心实验中, 评定形状编码算法优劣的参数 D_n (解码图象不匹配像素数/VOP 中像素总数) 为标准, 在 D_n 相同的条件下, 改进的 CSS 的压缩比比 CAE 算法减少约 20%. 图 8 为在相同的失真 D_n 条件下, 两者解码后图象形状的主观质量比较, 显然 CAE 比改进的 CSS 的结果差.

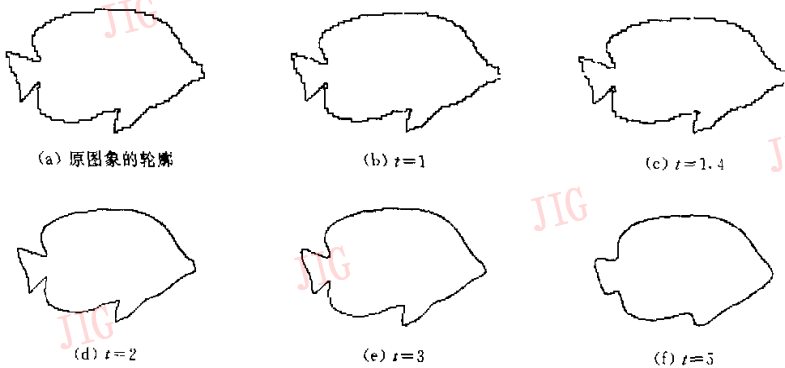


图 5 多尺度分层算法的形状分层结果

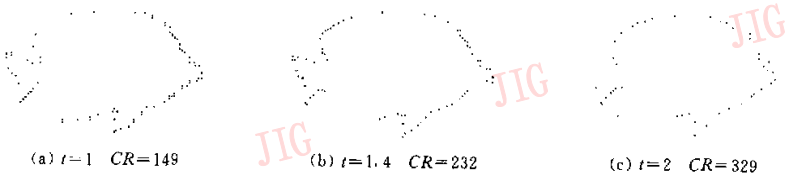


图 6 形状分层特征点的提取 (CR 为压缩比)

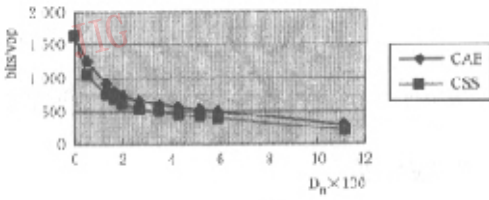


图 7 在相同失真系数 D_n 的条件下 CAE 与改进的 CSS 在编码比特数上的比较

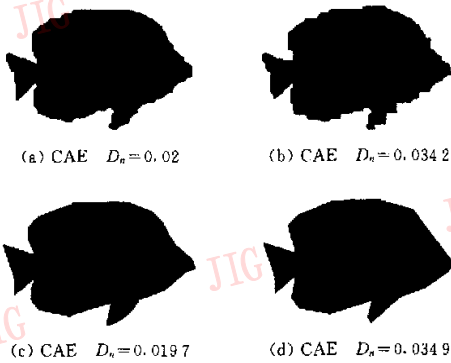


图 8 CAE 和改进的 CSS 方法在相同 D_n 条件下在主观质量的比较

4 结 论

曲率尺度空间理论最初主要应用在计算机与模式识别领域,此后应用于远程图象图形查找和检索。该方法能够很好地实现形状曲线的分层操作,经改进后可应用于形状编码算法中。与已经存在的基于多边形近似的形状编码算法相比,改进后的算法可以避免穷尽搜索运算。由实验结果可以看出,这种形状的表达和编码比 MPEG-4 校验模型中的 CAE 算法在压缩比和图象的主观质量上有很大提高。

参 考 文 献

- 1 Coding of moving pictures and audio[S]. ISO/IEC JTC1/SC29/WG11,MPEG99/4960. Melbourne, October 1999.
- 2 Brady N, Frank Bossen. Shape compression of moving objects using context-based arithmetic encoding[J]. Signal Processing: Image Communication, 2000,15:601~616.
- 3 Brady N. Context-based arithmetic encoding of 2D shape sequences [A]. In: Proc. Special Session in shape Coding, ICIP97[C], Santa Barbara, CA, 1997:1-29-32.
- 4 O'Connell K J. Object-adaptive vertex-based shape coding method[J]. IEEE Trans. Circuits Syst. Video Technol., 1997, 7(1):251~255.
- 5 Chung Jae-won, Lee Jin hak, Moon Joo hee *et al.* A new vertex-based binary shape encoder for high coding efficiency[J]. Signal Processing: Image Communication, 2000,15:665~684.
- 6 Kim Jong Il, Bovik Alan, Evans Brian. Generalized predictive binary shape coding using polygon approximation[J]. Signal Processing: Image Communication, 2000,15:643~663.
- 7 Mokhtarian F. Robust and efficient shape indexing through curvature scale space[A]. In:Proceedings of the Seventh British Machine Vision Conference, BMVC'96[C], Eedinburgh. 1996, 1:35~42.



石旭利 1970 年生,博士生,1998 年获上海大学光学专业光纤通信与光波导技术方向硕士学位,现为上海大学通信与信息工程学院通信与信息系方向博士研究生。主要从事图象编码和多媒体通信方面的研究。



张兆扬 教授,博士生导师,1962 年毕业于西安交通大学无线电工程系,现为上海大学通信与信息学科带头人,长期从事图象处理与多媒体通信研究。已完成多项国家自然科学基金及重大项目,发表论文 80 余篇。