



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所  
中国图象图形学学会  
北京应用物理与计算数学研究所

# 中国图象图形学报

2016  
01  
VOL.21

ISSN1006-8961  
CN11-3758/TB



医学图像处理 P69

# 中国图象图形学报

刊名题字：宋健 | 月刊（1996年创刊）



第21卷第1期（总第237期）  
2016年1月16日

中国精品科技期刊  
中国国际影响力优秀学术期刊  
中国科技核心期刊  
中文核心期刊

## 版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

## Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院  
主办单位 中国科学院遥感与数字地球研究所  
中国图象图形学学会  
北京应用物理与计算数学研究所

主 编 李小文  
编辑出版 《中国图象图形学报》编辑出版委员会  
邮政信箱 北京9718信箱  
邮 编 100101  
电子信箱 jig@radi.ac.cn  
电 话 010-64807995  
网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号  
总 发 行 北京报刊发行局  
订 购 全国各地邮局  
海外发行 中国国际图书贸易集团有限公司  
(邮政信箱: 北京399信箱 邮编: 100048)  
印刷装订 北京科信印刷有限公司

## Journal of Image and Graphics

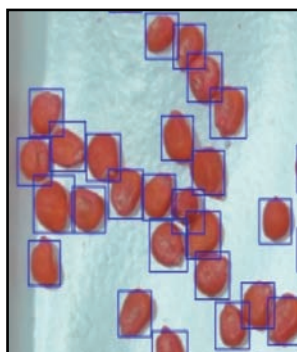
Title inscription: Song Jian | Monthly, Started in 1996

Superintended by Chinese Academy of Sciences  
Sponsored by Institute of Remote Sensing and Digital Earth, CAS  
China Society of Image and Graphics  
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief LI Xiaowen  
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics  
P.O.Box 9718, Beijing, P.R.China  
Zip code 100101  
E-mail jig@radi.ac.cn  
Telephone 010-64807995  
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals  
Domestic All Local Post Offices in China  
Overseas China International Book Trading Corporation  
(P.O.Box 399, Beijing 100048, P.R.China)  
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB  
ISSN 1006-8961  
CODEN ZTTXFZ  
国外发行代号 M1406  
国内邮发代号 82-831  
国内定价 60.00元



多尺度特征融合与极限学习机的玉米种子检测  
(第0024页)



深度卷积神经网络的显著性检测(第0053页)



双树复小波域的MRI图像去噪(第0104页)

## 图像处理和编码

空时形状预测与高效编码

朱仲杰, 王玉儿, 蒋刚毅..... 1

第二小方向导数信息熵的兴趣点检测

卢健, 黄杰, 潘峰..... 8

采用面积采样的反走样字符旋转算法

郑凯文, 刘文波..... 17

## 图像分析和识别

多尺度特征融合与极限学习机的玉米种子检测

柯道, 杜明智..... 24

## 图像理解和计算机视觉

结合NSCT和压缩感知的红外与可见光图像融合

陈木生..... 39

Plücker直线描述的空间后方交会

盛庆红, 陈姝文, 肖晖, 张斌, 王青, 费利佳..... 45

深度卷积神经网络的显著性检测

李岳云, 许悦雷, 马时平, 史鹤欢..... 53

## 计算机图形学

以优先点为中心的Delaunay三角网生长算法

尤磊, 唐守正, 宋新宇..... 60

## 医学图像处理

快速数字影像重建的2维/3维医学图像配准

刘坤, 吕晓琪, 谷宇, 于荷峰, 任国印, 张明..... 69

小波-Lagrange方法进行医学图像层间插值

武士想, 尚鹏, 王立功..... 78

## 遥感图像处理

多主体框架下结合最大期望值和遗传算法的SAR图像分割

张金静, 李玉, 赵泉华..... 86

## 第十届图像图形技术与应用学术会议

视差互信息引导下的立体航空影像与LiDAR点云自动配准

吴军, 胡彦君, 饶云, 彭智勇..... 95

双树复小波域的MRI图像去噪

黄学优, 张长江..... 104

快速定位图像尺度和区域的3维跟踪算法

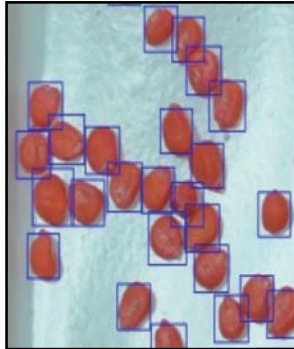
赵奇可, 孙延奎..... 114

局部不变特征点的精度指标

滕日, 周进, 蒋平, 崔雄文..... 122

# CONTENTS

## JOURNAL OF IMAGE AND GRAPHICS



Detection of maize seeds based on multi-scale feature fusion and extreme learning machine(P0024)



Saliency detection based on deep convolutional neural network(P0053)



MRI denoising based on dual-tree complex wavelet transform(P0104)

### Image Processing and Coding

- Spatio-temporal shape prediction and efficient coding  
Zhu Zhongjie, Wang Yuer, Jiang Gangyi ..... 1
- Interest point detection by using information entropy of the second small direction derivative  
Lu Jian, Huang Jie, Pan Feng ..... 8
- Anti-aliasing algorithm for character rotation based on area sampling  
Zheng Kaiwen, Liu Wenbo ..... 17

### Image Analysis and Recognition

- Detection of maize seeds based on multi-scale feature fusion and extreme learning machine  
Ke Xiao, Du Mingzhi ..... 24

### Image Understanding and Computer Vision

- Image fusion of visual and infrared image based on NSCT and compressed sensing  
Chen Musheng ..... 39
- Space resection method based on Plücker line  
Sheng Qinghong, Chen Shuwen, Xiao Hui, Zhang Bin, Wang Qing, Fei Lijiao ..... 45
- Saliency detection based on deep convolutional neural network  
Li Yueyun, Xu Yuelei, Ma Shiping, Shi Hehuan ..... 53

### Computer Graphics

- Growth algorithm centered on priority point for constructing the Delaunay triangulation  
You Lei, Tang Shouzheng, Song Xinyu ..... 60

### Medical Image Processing

- The 2D/3D medical image registration algorithm based on rapid digital image reconstruction  
Liu Kun, Lyu Xiaoqi, Gu Yu, Yu Hefeng, Ren Guoyin, Zhang Ming ..... 69
- Inter-slice interpolation for medical images by using the wavelet-lagrange method  
Wu Shixiang, Shang Peng, Wang Ligong ..... 78

### Remote Sensing Image Processing

- Combining the expectation maximization and genetic algorithms for SAR image segmentation within a MAS framework  
Zhang Jinjing, Li Yu, Zhao Quanhua ..... 86

### Column of IGIA' 2015

- Automatic registration of aerial stereo imagery to LiDAR data through iterative parallax mutual information computation  
Wu Jun, Hu Yanjun, Rao Yun, Peng Zhiyong ..... 95
- MRI denoising based on dual-tree complex wavelet transform  
Huang Xueyou, Zhang Changjiang ..... 104
- Three dimensional tracking with fast locating of image scale and area  
Zhao Qike, Sun Yankui ..... 114
- Extraction precision of local invariant feature points  
Teng Ri, Zhou Jin, Jiang Ping, Cui Xiongwen ..... 122



中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2016)01-0001-07

论文引用格式: Zhu Z J, Wang Y E, Jiang G Y. Spatio-temporal shape prediction and efficient coding[J]. Journal of Image and Graphics, 2016, 21(1): 0001-0007. [朱仲杰, 王玉儿, 蒋刚毅. 空时形状预测与高效编码[J]. 中国图象图形学报, 2016, 21(1): 0001-0007. ] [DOI: 10.11834/jig.20160101]

## 空时形状预测与高效编码

朱仲杰<sup>1</sup>, 王玉儿<sup>1</sup>, 蒋刚毅<sup>2</sup>

1. 浙江万里学院宁波市 DSP 重点实验室, 宁波 315100;

2. 宁波大学电路与系统研究所, 宁波 315211

**摘要:** 目的 形状是视觉对象的关键特征, 形状编码是对象基图像和视频处理中的关键技术, 但现有无损形状编码方法压缩效率普遍不高。为此, 提出一种基于链码表示和空时预测的高效无损形状编码新算法。方法 首先逐帧提取视觉对象的形状轮廓并转化为链码表示; 然后基于对象轮廓的帧间活动性将形状视频序列分成帧内预测编码帧和帧间预测编码帧, 并基于轮廓链码的空域相关性和时域相关性对二者分别进行空域和时域补偿与预测; 最后基于链码的方向约束特性对预测后的位移矢量和预测残差进行高效编码压缩。结果 为了检验所提算法的性能, 基于 MPEG-4 标准形状测试序列进行了编码实验测试。与现有主要方法相比本文算法能提高压缩效率 6% 到 71.6% 不等。结论 本文算法可广泛应用于对象基编码、基于内容的图像检索、图像分析与理解等领域。

**关键词:** 视觉对象; 形状编码; 预测编码; 高效编码

## Spatio-temporal shape prediction and efficient coding

Zhu Zhongjie<sup>1</sup>, Wang Yuer<sup>1</sup>, Jiang Gangyi<sup>2</sup>

1. Ningbo Key Lab. of DSP, Zhejiang Wanli University, Ningbo 315100, China;

2. Institute of circuits and system, Ningbo University, Ningbo 315211, China

**Abstract:** **Objective** The use of a shape is a popular way to define objects, and efficient shape coding is a key technique in object-based applications. Shape coding is also a hot research topic in the field of image and video signal processing, and many shape-coding techniques have been proposed. Among these methods, chain-coding is a popular technique that can be used for lossless shape coding. However, most existing chain-based shape-coding methods have not exploited the spatio-temporal redundancy contained within shape image sequences. Similar to the existence of strong spatio-temporal redundancy within and among video textures, a strong redundancy also exists within and between object contours. This redundancy can be exploited to improve coding efficiency. Hence, in this paper, a novel chain-based lossless shape-coding scheme is proposed by exploiting the spatio-temporal correlations among object contours to acquire high coding efficiency. **Method** First, for a given shape image sequence, the contours of visual objects are extracted, thinned to perfect single-pixel width, and transformed into chain-based representation frame by frame. Second, the activity of object contours in each frame is detected.

收稿日期: 2015-02-04; 修回日期: 2015-09-15

**基金项目:** 国家自然科学基金项目 (60902066); 浙江省自然科学基金项目 (LY14F010006); 宁波市自然科学基金项目 (2015A610136); 人社部留学科技人员择优资助项目 (2013-277); 教育部留学回国人员科研启动基金项目 (2014-1685)

**第一作者简介:** 朱仲杰 (1976—), 男, 教授, 2004 年于浙江大学获电子科学与技术专业博士学位, 主要研究方向为 2 维与 3 维视频编码与传输, 近年来在国内外学术期刊上发表学术论文 30 余篇。E-mail: zhongjiezhu@hotmail.com

**Supported by:** National Natural Science Foundation of China (60902066); Natural Science Foundation of Zhejiang Province, China (LY14F010006); Natural Science Foundation of Ningbo, China (2015A610136)

ted and evaluated. The shape frames are classified into two coding categories on the basis of this activity: intra-coding frames and inter-coding frames. If the contour activity in a frame is larger than a preset threshold, the activity will be encoded as an inter-coding frame; otherwise, it will be encoded as an intra-coding frame. For an intra-coding frame, the spatial correlations within object contours are exploited on the basis of chain-based spatial prediction and compensation. For an inter-coding frame, the temporal correlations among object contours are exploited on the basis of chain-based temporal prediction and compensation. Finally, a new method is introduced to efficiently encode the prediction residuals and motion displacements by analyzing the constraints among chain links. **Result** To evaluate the performance of the proposed scheme, experiments are conducted and a partial comparison is performed against some well-known existing methods, including the lossless coding scheme proposed by the Joint Bi-level Image Experts Group (JBIG), the improved lossless coding scheme proposed by JBIG (JBIG2), the Context-based Arithmetic Encoding with Intra-mode (CAE Intra) of MPEG-4, the Context-based Arithmetic Encoding with Inter-mode (CAE Inter) of MPEG-4, the Digital Straight Line Segments-based Coding with Intra-mode (DSLSC Intra) and the Digital Straight Line Segments-based Coding with Inter-mode (DSLSC Inter), are also presented. The experimental results show that the average code length of our scheme is only 28.4% of JBIG, 32.3% of JBIG2, 39.9% of CAE Intra, 78.1% of CAE Inter, 48.4% of DSLSC Intra, and 94.0% of DSLSC Inter. **Conclusion** As a whole, the proposed scheme outperforms all existing techniques and is considerably more efficient than other methods. As far as we know, the DSLSC Inter is the most efficient lossless shape-coding approach. However, compared with the DSLSC Inter, the proposed scheme has an average code length that can be reduced by 6%. The proposed scheme has wide prospects in many object-based images and video applications, such as object-based coding, object-based editing, and object-based retrieval.

**Key words:** visual objects; shape coding; predictive coding; high efficient coding

## 0 引言

形状是视觉对象的重要特征,形状信息是进行视觉对象定义、表示与处理的关键信息。因此,形状编码是对象基图像与视频应用领域中的关键技术,也是当前国内外图像与视频处理领域的一个研究热点<sup>[1-3]</sup>。形状编码方法一般可以分为两类:基于位图的方法和基于轮廓的方法。基于位图的方法将对象形状表示为一个二值图像,然后采用二值图像编码方法进行编码,如JBIG、JBIG2、MPEG-4的基于上下文的算术编码(CAE)等<sup>[4-7]</sup>。基于轮廓的方法首先提取视觉对象的轮廓曲线,然后对轮廓曲线进行编码,常用的方法有基于链码的方法和基于多边形和样条等曲线拟合的方法<sup>[8-11]</sup>。一般来说,相比于基于位图的方法,基于曲线拟合的形状编码方法能获得更高的压缩效率,但其一般用于有损编码而不能用于无损编码。由于形状信息的重要性,许多实际应用领域,如高精度对象基视频检索与识别、高效对象基视频编码等领域,往往要求视觉对象的形状信息尽可能准确,对形状信息的编码要求低失真甚至是无损编码。因此,研究更为高效的形状无损编

码方法具有重要意义。

近年来,一些新的形状编码方法不断被提出。例如,文献[12]提出了一种基于算术编码的二值图像形状编码方法。它利用目标轮廓中存在的局部线性边缘来增强算术编码上下文建模的准确性,从而提高编码效率。文献[13]提出了一种四分树结构的基于上下文算术编码的形状编码方案,与MPEG-4 CAE相比可以有效提高编码效率。Lai等人<sup>[4]</sup>研究了率失真优化形状编码中的边缘选择与优化编码问题,提出了基于8分区和16分区的边缘编码方案,可以降低待编码定点数从而提高编码效率。文献[15]提出了一种基于图像相关性的高效形状编码方法,利用图像内容和对象形状之间的相关性来提高编码效率。然而,该方法也是一种有损编码方法,不能应用于无损编码领域。

基于链码的轮廓基形状编码由于简单、快速受到广泛关注,可以用于形状无损编码。但现有的诸多算法主要基于链码本身的方向约束特性进行帧内编码,而没有利用对象轮廓链码间的空域和时域冗余。因此,其压缩效率相对不高。为此,本文在分析对象轮廓链码间的帧内空域相关性和和帧间时域相关性的基础上,提出一种面向形状视频的高效无损

编码算法,整体框图如图1所示。首先逐帧提取形状视频的对象轮廓并转化为链码表示;然后基于对象轮廓的帧间活动性将视频分成帧内预测编码帧和帧间预测编码帧,并基于轮廓链码的空域相关性和时域相关性对二者分别进行空域和时域预测与补偿;最后基于链码的方向约束特性对预测后的位移矢量和预测残差进行高效编码压缩。与最新无损形状编码方法相比所提算法能获得更好的压缩性能。

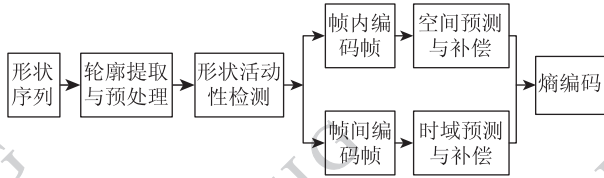


图1 算法总体框图

Fig. 1 Diagram of the proposed scheme

## 1 轮廓提取与预处理

对于给定的二值形状序列,首先逐帧提取对象轮廓并将其细化成单像素宽度。要求除了起始点和轮廓交叉点外,对象轮廓上的每个边缘点在其8邻域内存在且只存在2个相邻的轮廓边缘点。然后基于交叉点将一帧内的轮廓分成若干段,使得每段轮廓内的每个边缘点在其8邻域内都只存在2个相邻的轮廓边缘点。最后对每个轮廓段进一步将其转换成链码表示。轮廓提取和预处理的主要步骤如下:

1) 设  $\{m_i(x, y)\} (i=0, \dots, N-1)$  表示给定的二值形状图像序列,  $N$  表示序列的长度,  $m_i(x, y)$  表示第  $i$  帧形状图像。令  $m_i(x, y)=0$  的区域表示前景视觉对象,  $m_i(x, y)=255$  的区域表示背景。

2) 对每个  $m_i(x, y)$ , 依次检测所有值为0的像素, 如果某个像素在其8邻域内存在值为255的像素, 则将其提取作为目标轮廓点。

3) 检测出  $m_i(x, y)$  中所有的目录轮廓点, 其构成的集合表示该帧中的对象轮廓, 用  $S_i$  表示。

4) 将  $S_i$  细化成单像素宽度并基于轮廓交叉点将其分成若干轮廓子段  $S_{ik} (k=0, \dots, M_i-1)$ , 使得  $S_{ik}$  中的每个轮廓点, 在其8邻域内存在且只存在2个相邻的轮廓点。其中,  $M_i$  表示轮廓子段的数量。

5) 基于链码将  $S_{ik}$  转换成相应的链码形式  $L_{ik}$ 。本文使用经典的8方向Freeman链码, 其对应的8

个方向如图2所示。

6) 重复步骤2), 直到完成所有形状序列的轮廓提取和预处理。

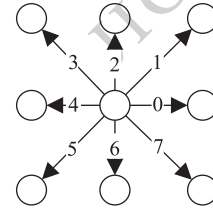


图2 8方向Freeman链码

Fig. 2 Eight 8-connected Freeman chain code

## 2 形状活动性检测与时域预测

### 2.1 形状活动性检测

对进行轮廓提取和预处理后的形状图像序列按照轮廓活动性分为帧内编码帧和帧间编码帧, 对帧内编码帧图像进行帧内空间预测与补偿编码, 对帧间编码图像进行帧间时域预测与补偿编码。

对于给定的第  $i$  帧形状图像  $m_i(x, y)$ , 设  $\{L_{ik}\} (k=0, \dots, M_i-1)$  表示其所有轮廓子段的链码集合,  $M_i$  表示轮廓子段的数量。令  $N_{ik}$  表示  $L_{ik}$  的长度, 则图像  $m_i(x, y)$  的轮廓活动性  $c_i$  定义为

$$c_i = \frac{\left| \sum_{k=0}^{M_i} N_{ik} - \sum_{k=0}^{M_{i-1}} N_{(i-1)k} \right|}{\sum_{k=0}^{M_i} N_{ik}} \times \frac{|M_i - M_{i-1}|}{M_i} \quad (1)$$

式中,  $M_{i-1}$  表示第  $i-1$  帧  $m_{i-1}(x, y)$  中轮廓子段的数量。如果  $c_i$  大于给定的阈值  $T$ , 则对  $m_i(x, y)$  进行帧内空域预测编码, 否则进行帧间时域预测编码。本文中, 阈值  $T$  设为 0.05。

### 2.2 空时预测与补偿

由于每帧图像的对象轮廓被分成了若干子段, 因此在进行空时预测与补偿编码时是以轮廓子段为基本单位。空域预测和时域预测的基本方法类似, 不同之处主要在于参考轮廓子段的选择上。对进行空域预测与补偿编码的轮廓子段, 其参考轮廓子段选取于同一帧中已编码的轮廓子段。对于进行时域预测与补偿编码的轮廓子段, 其参考轮廓子段选取于前一帧中已编码的轮廓子段。对当前编码的轮廓

子段  $L_{ik}$ , 其链码序列用  $\{l_n\} (n=0, \dots, N_{ik}-1)$  表示, 设  $L'$  表示其选取的参考轮廓子段, 相应链码序列为  $\{l'_n\} (n=0, \dots, N^r-1)$ , 则对  $L_{ik}$  进行空时预测与补偿的主要步骤如下:

1) 首先将  $\{l_n\}$  分割成一定长度的预测单元  $U_j$  ( $j=0, \dots, N^u$ ),  $N^u$  是分割后的预测单元的数量, 即

$$N^u = \lfloor N_{ik}/N^L \rfloor \quad (2)$$

式中,  $N^L$  是事先给定的预测单元的长度。

2) 对每个  $U_j$  在给定的搜索区间  $[a, b]$  搜索其运动位移

$$V_j = \arg \max_{p \in [a, b]} \sum_{q=0}^{N^L-1} e_{jq}(p) \quad (3)$$

$$e_{jq}(p) = \begin{cases} 1 & l_{j \cdot N^L + q} = l'_{j \cdot \lceil N^r/N^u \rceil + q + p} \\ 0 & l_{j \cdot N^L + q} \neq l'_{j \cdot \lceil N^r/N^u \rceil + q + p} \end{cases} \quad (4)$$

3) 计算对  $L_{ik}$  进行运动补偿后的预测误差

$$L_{ik}^e = \{l_n^e, n=0, \dots, N_{ik}-1\} \quad (5)$$

$$l_{j \cdot N^L + q}^e = \begin{cases} 0 & l_{j \cdot N^L + q} = l'_{j \cdot \lceil N^r/N^u \rceil + q + MV_j} \\ l_{j \cdot N^L + q} & l_{j \cdot N^L + q} \neq l'_{j \cdot \lceil N^r/N^u \rceil + q + MV_j} \end{cases} \quad (6)$$

$$0 \leq j \leq N^u - 1, 0 \leq q \leq N^L - 1$$

一般来说,  $L_{ik}^e$  中多数链路的预测误差都为 0。因此相对于直接编码  $L_{ik}$ , 编码  $L_{ik}^e$  会获得更高的编码效率。同时, 从式(6)可以看出, 当  $L_{ik}$  中的某个链路  $l_{j \cdot N^L + q}$  等于其参考轮廓子段中的对应链路  $l'_{j \cdot \lceil N^r/N^u \rceil + q + MV_j}$  时,  $L_{ik}^e$  中的对应的残差链路  $l_{j \cdot N^L + q}^e$  值为 0, 否则其值等于原来  $l_{j \cdot N^L + q}$  的值。这样处理的优点是对非零  $l_{j \cdot N^L + q}$  进行编码时可以利用前后链路之间的空域相关性, 从而提高编码效率。

### 3 熵编码

为了进行高效熵编码, 定义了两种编码模式: 直接编码模式和预测编码模式。直接编码模式不进行空时预测与补偿, 直接利用轮廓子段间的空域相关性进行编码。由于轮廓细化对前后链路间的方向进行了限制和约束, 对于给定的某个方向码, 其可能出现的下一个方向码的类型会大幅度减少。例如, 如果当前链路所对应的方向码为 0, 则下一链路只可能是 0, 1 或 7。类似地, 链码中所有链路之间的方向约束特性如表 1 所示。同时, 由于实际对象轮廓中前后两个链码具有很强的相关性, 其方向变化一

般是缓慢和渐进的。因此对于给定的某个方向码其后续每一个可能方向码出现的概率是不一样的。一般来说, 与当前链路方向相差越小的链路出现的概率越大。反之, 方向相差越大出现的概率越小。因此为了提高编码效率, 对表 1 中所示的当前链路的所有可能的下一个链路采用 VLC 编码, 对出现概率较大的链路分配较短的码字, 对出现概率较小的链路分配较长的码字。

表 1 链码中前后链路之间的约束特性及 VLC 码

Table 1 Constraints between two consecutive links and the VLC codes for the next possible candidates

当前链路	下一链路	VLC 码	当前链路	下一链路	VLC 码
0	7	10	4	3	10
0	0	0	4	4	0
0	1	11	4	5	11
1	0	01	5	4	01
1	1	00	5	5	00
1	2	10	5	6	10
1	7	110	5	3	110
1	3	111	5	7	111
2	1	10	6	5	10
2	2	0	6	6	0
2	3	11	6	7	11
3	2	01	7	6	01
3	3	00	7	7	00
3	4	10	7	0	10
3	1	110	7	5	110
3	5	111	7	1	111

预测编码模式首先对轮廓子段进行空域或时域补偿预测, 然后对预测后的位移矢量和残差编码。对某个预测编码模式轮廓子段  $L_{ik}$ , 由于和其参考轮廓子段之间存在的空时相关性, 其空时预测与补偿后的残差  $L_{ik}^e$  中多数链路的值为 0, 只有少数链路的值是非 0。0 链路的行程往往较大, 我们对其进行行程编码。对于非 0 链路, 采用直接编码模式中相同的编码策略, 基于前后链路之间的约束特性和空域相关性进行变长编码。

在本文算法中, 对帧内编码帧的第 1 个轮廓子段始终采用直接编码模式进行编码, 对其他轮廓子段首先分别采用直接编码和空域补偿预测编码进行预编码, 然后选取能获得更高压缩比的模式作为其最终编码模式。类似地, 对于帧间编码帧中的每个



轮廓子段首先分别采用直接编码和时域补偿预测编码进行预编码,然后选取能获得更高压缩比的模式作为其最终编码模式。

#### 4 实验结果

为了检验所提算法的性能,进行了实验仿真。实验中使用了 news 和 weather 两个 MPEG-4 标准形状视频测试序列。其中 news 为 cif 格式,300 帧,weather 为 sif 格式,100 帧。测试序列的原始图像是二值形状掩模图像。为此,首先逐帧提取形状视频帧的对象轮廓并转化为链码表示。然后基于对象轮廓的帧间活动性将视频分成帧内预测编码帧和帧间预测编码帧,并基于轮廓链码的空域相关性和时域相关性对二者分别进行空域和时域预测与补偿。最后基于链码的方向约束特性对预测后的位置矢量和预测残差进行熵编码。

图3给出了对象轮廓的帧间运动预测补偿示意图。其中  $L$  表示待编码轮廓子段,  $L'$  表示参考帧中的参考子段,  $L^e$  是经过运动补偿后的残差。可以看出,由于  $L$  和  $L'$  中存在的相关性,残差  $L^e$  中的多数比特0。因此,相对于直接编码  $L$ ,编码  $L^e$  可以获得更高的编码效率。

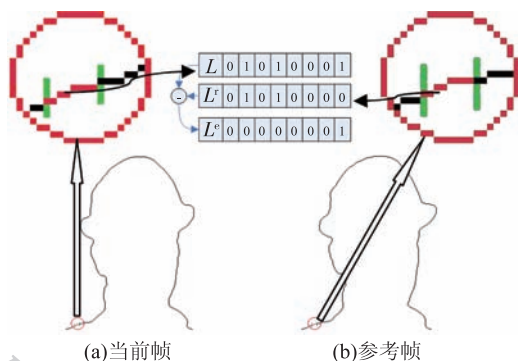


图3 帧间预测示意图(预测单元长度为8)

Fig. 3 Example of inter-prediction where the length of prediction unit is set to 8 ( (a) the current frame; (b) the reference frame)

图4和表2给出了本文算法在不同预测单元长度和不同搜索区间大小下的编辑结果。可以看出预测单元的长度和搜索区间的大小直接影响编码后的平均码长。一般来说搜索区别越大,编码后的平均码长越越段,编码效率越高。但编码效率与预测单

元的长度并不是单调性函数。一般来说,选取较长的预测单元可以减少预测单元的数量,从而减少运动矢量的数据量。但长度过长可能会导致预测残差增大反而增加残差的编码比特。

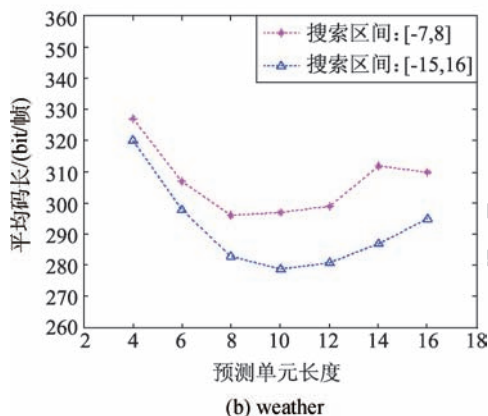
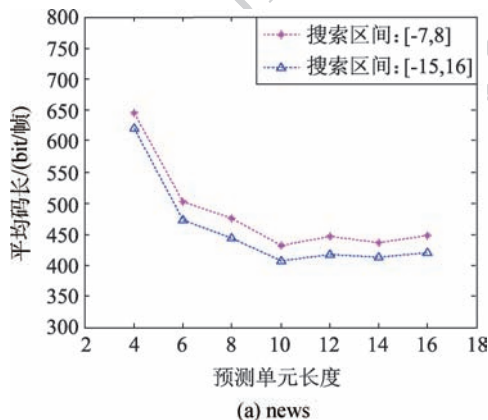


图4 news 和 weather 测试序列在不同预测单元长度和搜索范围下的编码结果

Fig. 4 Coding results of news and weather sequences under different prediction units with different lengths and under different search scopes ((a) news; (b) weather)

表2 不同预测单元长度和搜索范围下的编码结果

Table 2 Coding results under different prediction units with different lengths and under different search scopes

序列	搜索范围	预测单元长度						
		4	6	8	10	12	14	16
news	$[-7,8]$	646	503	476	433	447	436	448
	$[-15,16]$	621	473	444	408	418	413	420
weather	$[-7,8]$	327	307	296	297	299	312	310
	$[-15,16]$	320	298	283	279	281	287	295

表3给出了本文算法与现有主要无损形状编码方法的对比实验结果。主要包括 JBIG, JBIG2, MPEG-4 CAE Intra, MPEG-4 CAE Inter 以及 DSLSC Intra, DSLSC Inter 等。其中,据笔者所知 DSLSC Inter 算法是目前已有文献中压缩性能最好的无损形状编码方法。从表3实验结果可以看出,与现有方法相比,所提算法能获得更高的压缩性能。为了进一步定量考察所提算法与已有方法的相对压缩性能,定义相对压缩比特率为

$$\gamma = \frac{C_p}{C_i}, i \in \{JBIG, JBIG2, CAE, DSLSC\} \quad (7)$$

式中,  $\gamma$  表示相对压缩比特率,  $C_p$  和  $C_i$  分别表示所提算法与现有方法编码压缩后的比特率,表4给出了本文算法对比上述4种方法的  $\gamma$  值。可以看出,与现有方法相比,本文算法能大幅降低形状图像编码压缩后的平均比特率,分别是 JBIG 的 28.4%, JBIG2 32.3%、CAE Intra 的 39.9%、CAE Inter 的 78.1%、DSLSC Intra 的 48.4%、DSLSC Inter 的 94.0%。

表3 在预测单元长度为10 搜索区间为[-15,16]时与现有方法的对比结果

Table 3 Comparison results against the existing methods where the length of prediction unit is set to 10 and the search scope is set to [-15, 16]

序列	JBIG	JBIG2	/(bit/帧)				本文 算法
			CAE Intra	CAE Inter	DSLSC Intra	DSLSC Inter	
news	1 672	1 528	1 568	592	1056	424	408
weather	864	736	520	320	480	304	279

表4 在预测单元长度为10 搜索区间为[-15,16]时与现有方法的相对压缩比特率  $\gamma$

Table 4 The relative compression ratio  $\gamma$  comparing with the existing methods where the length of prediction unit is set to 10 and the search scope is set to [-15, 16]

序列	JBIG	JBIG2	/%			
			CAE Intra	CAE Inter	DSLSC Intra	DSLSC Inter
news	24.4	26.7	26.0	68.9	38.6	96.2
weather	32.3	37.9	53.7	87.2	58.1	91.8
平均值	28.4	32.3	39.9	78.1	48.4	94.0

## 5 结 论

形状是视觉对象的重要特征,高效形状编码是对象基图像与视频处理中的关键技术。本文提出了一种基于链码表示和空时预测的高效无损的形状编码方法。算法根据视觉对象轮廓的帧间活动性将视频分成帧内预测编码帧和帧间预测编码帧,并基于轮廓链码的空域相关性和时域相关性对二者分别进行空域和时域预测与补偿,可以充分去除对象轮廓链的帧内空域冗余和帧间时域冗余。实验结果显示,与现有方法相比,可以获得更高的编码效率。所提算法能广泛应用于对象基编码、对象基检索以及对象交互等图像和视频应用领域。下一步拟重点研究将所提方法进行拓展以应用于3 维视频和多视点视频的 shape 编码。

## 参考文献 (References)

- [1] Zhu Z J, Wang Y E, Jiang G Y. On multi-view video segmentation for object-based coding [J]. Digital Signal Processing, 2012, 22(6):954-960. [DOI:10.1016/j.dsp.2012.05.006]
- [2] Aghito S M, Forchhammer S. Efficient coding of shape and transparency for video objects [J]. IEEE Trans. on Image Processing, 2007, 16 (9): 2234-2244. [DOI: 10.1109/TIP.2007.903902]
- [3] Liu Q, Ngan K N. Arbitrarily shaped object coding based on H.264/AVC [C]//International Symposium on Intelligent Signal Processing and Communication Systems. Kanazawa, Japan: IEEE, 2009: 343-346. [DOI:10.1109/ISPACS.2009.5383832]
- [4] ISO/IEC JTC1/SC29. ISO/IEC-11544 Coded representation of Picture and Audio Information-progressive Bi-level Image Compression[S]. Japan: ISO/IEC,1993.
- [5] ISO/IEC JTC1/SC29. ISO/IEC-14492 Coded Representation of Picture and Audio Information-lossy /Lossless Coding of Bi-Level Images (JBIG2)[S]. Japan: ISO/IEC, 2000.
- [6] ISO/IEC JTC1/SC29. ISO/IEC-14496-2 Information Technology-coding of Audio-visual Objects-part 2: Visual[S]. Japan: ISO/IEC, 1999.
- [7] Martin K, Lukac R, Plataniotis K N. SPIHT-based coding of the shape and texture of arbitrarily shaped visual objects [J]. IEEE Trans. on Circuits and Systems for Video Technology, 2006, 16(10):196-1208. [DOI:10.1109/TCSVT.2006.882388]
- [8] Bandyopadhyay S K, Kondi L P. Optimal bit allocation for joint texture-aware contour-based shape coding and shape-adaptive texture coding [J]. IEEE Trans. on Circuits and Systems for Video

- Technology, 2008, 18 ( 6 ): 840-844. [ DOI: 10.1109/TCSVT.2008.918784 ]
- [ 9 ] Kim K J, Suh J Y, Kang M G. Generalized inter-frame vertex-based shape encoding scheme for video sequences [ J ]. IEEE Trans. on Image Process. , 2000, 9 ( 10 ): 1667-1676. [ DOI: 10.1109/83.869178 ]
- [ 10 ] Nunes P, Marques F, Pereira F, et al. A contour-based approach to binary shape coding using a multiple grid chain code [ J ]. Signal Processing: Image Communication, 2000, 15 ( 7 ): 585-599. [ DOI:10.1016/S0923-5965(99)00041-7 ]
- [ 11 ] Wulandhari L A, Haron H. Characteristic of rectangular vertex chain code for shapes with hole [ C ] // Proceedings of the International Conference on Information Management and Engineering. Kuala Lumpur, MYS: IEEE, 2009: 648-650. [ DOI:10.1109/ICIME.2009.32 ]
- [ 12 ] Aghito S M, Forchhammer S. Context-based coding of bi-level images enhanced by digital straight line analysis [ J ]. IEEE Trans. on Circuits Systems for Video Technology, 2006, 15 ( 8 ): 2120-2130. [ DOI: 10.1109/TIP.2006.875168 ]
- [ 13 ] Shen Z L, Frater M R, Arnold J-F. Quad-tree block-based binary shape coding [ J ]. IEEE Trans. on Circuits and Systems for Video Technology, 2008, 18 ( 6 ): 845-850. [ DOI: 10.1109/TCSVT.2008.919086 ]
- [ 14 ] Lai Z Y, Zhu J H, Ren Z, et al. Arbitrary directional edge encoding schemes for the operational rate-distortion optimal shape coding framework [ C ] // Proceedings of the Data Compression Conference. Snowbird, UT: IEEE, 2010: 20-29. [ DOI: 10.1109/DCC.2010.10 ]
- [ 15 ] Luo H T. Image-dependent shape coding and representation [ J ]. IEEE Trans. on Circuits and Systems for Video Technology, 2005, 15 ( 3 ): 345-354. [ DOI:10.1109/TCSVT.2004.842596 ]