2000 年上海大学博士学位论文 9

面向对象的低比特率活动图象压缩编码新算法研究

作者：张　颖
专业：通信与信息系统
导师：张兆扬

上海大学出版社

2000 年上海大学博士学位论文

面向对象的低比特率活动图象压缩编码新算法研究

作　者：张 颖
专　业：通信与信息系统
导　师：张兆扬

上海大学出版社
·上 海·

Shanghai University Doctoral Dissertation (2000)

The Study on Novel Algorithms for Object-Oriented Motion Images Compression Encoding at Low Bit Rate

Candidate: Zhang Ying
Major: Communication and Information System
Supervisor: Prof. Zhang Zhaoyang

Shanghai University Press

• Shanghai •

上 海 大 学

本论文经答辩委员会全体委员审查，确认符合上海大学博士学位论文质量要求。

答辩委员会名单：

主任：	宋文涛	教授，上海交大电子信息工程学院	200030
委员：	郑志航	教授，上海交大电子工程系	200030
	莫玉龙	教授，上海大学通信工程系	200072
	王朔中	教授，上海大学电子信息工程系	200072
	王治钢	研究员，上海航天局 809 研究所	200031
导师：	张兆扬	教授，上海大学	200072

评阅人名单：

宋文涛	教授，上海交大电子信息工程学院	200030
莫玉龙	教授，上海大学通信工程系	200072
翁默颖	教授，华东师大电子科学系	200062

评议人名单：

朱维乐	教授，电子科技大学	610054
王朔中	教授，上海大学电子信息工程系	200072
郑志航	教授，上海交大电子工程系	200030
王绍霖	教授，同济大学电气工程系	200080
符锡金	研究员，上海通用卫星导航公司	200040
王治钢	研究员，上海航天局 809 研究所	200031

答辩委员会对论文的评语

面向对象的低比特率活动图象编码是与 MPEG-4 密切相关的、当前国际上的研究热点。故该选题具有重要的理论价值和潜在实际意义。

该论文的研究成果在以下几点上有所创新：

(1) 对于基于二次空间变换的运动补偿方法提出了改进方案，有效地提高了运动估计精度；

(2) 提出了快速时空分割对象生成算法，显著地降低了运算负载；

(3) 提出了基于六领域系网格以及堆数据结构的最优轮廓编码方案。基于上述算法构成的编解码系统，克服了 H.263 标准的方块效应，信噪比有所提高。

论文反映出作者已较全面地掌握了与本课题相关的国内外发展动态，且具有扎实的基础理论及系统深入的专门知识，显示了该生已具有相当强的独立科研能力。论文立论正确，论据充分，条理性好，层次分明、实验数据可靠。在答辩过程中，回答问题思路清晰，能予以正确回答。

答辩委员会表决结果

经答辩委员会表决，全票同意通过张颖同学的博士学位论文答辩，建议授予工学博士学位。

答辩委员会主席：宋文涛

2000 年 8 月 7 日

摘　要

本文首先以信源模型为准则，对低比特率数字视频压缩编码法进行了分类和简要的分析，指出其中的面向对象的低比特率压缩编码方法是一种很有前途的压缩算法．不仅在高压缩比时可有效克服经典图象编码方法中出现的伪轮廓，而且其基于对象的功能，属于最新的 MPEG-4 标准的重要标志之一．

面向对象的视频压缩算法突破了原有图象编码的信源模型理论框架，基于三维实景模型将图象分割成若干个具有任意形状的对象．每一个对象由三个参数集表示，即运动、形状以及纹理信息，其中运动估计与形状编码为两个最重要的部分．同经典的基于块的匹配技术不同的是，面向对象编码技术中的运动估计应该与景物的对象表示相一致．而且，应该提供对对象运动的可靠估计而不仅仅是一个时间预测．故需引入较复杂的运动模型；而形状编码不仅应该在不明显降低解码图象质量的前提下，实现对象形状的有效编码，而且需为将来的基于对象的可操作性提供便利．因此，面向对象编码中的运动估计模块与形状编码模块值得深入研究，当然，以上两个模块都建立在景物对象的快速及有效生成的基础上．本文正是针对上述这些问题进行了探索和尝试，提出了一些新的算法．

论文的主要成果包括：

提出一种改进的基于二次空间变换的运动补偿，对由空域分割生成的每一个区域进行后向运动估计．用进退法及数值法确定局部最小值等方法，克服原估计在没有足够的纹理信息的

小区域上收敛缓慢或发散的弱点，提高运动估计精度.

提出一种基于二次空间变换运动模型的快速时空分割算法. 在上述图象空域分割及区域二次空间变换运动模型参数估计的基础上，通过将非线性的运动模型参数估计问题线性化、Householder 变换以及求区域相应的 R、$\bar{z}$ 矩阵的快速算法等方法，有效地降低运算负载，实现快速时空分割.

提出一种基于六边形网格邻域系以及堆数据结构的最优轮廓图编码方法. 以轮廓链为优化编码单位，以最大绝对值距离失真为约束条件，利用与时空分割结果对应的轮廓链单源有向图的稀疏特性，采用基于堆的数据结构，利用堆排序提取 T 标号点集中的最小值，有效减少在求单源有向图最短路径问题中通常采用的 Dijkstra 算法的时间复杂度，获得使轮廓图编码码率最低的最优分段线性表示的顶点位置. 而且，由于以轮廓链为优化单位，不会破坏原分割图的区域连通性.

对于轮廓链的起始结点，提出了在深度优先遍历搜索过程中，通过在编解码器两端，同时建立动态生成树的方法，使解码端对任何一个轮廓链编码数据进行解码时，其起始结点可从已解码的轮廓链起始结点中选取，如此可有效降低结点编码代价.

提出了基于六边形邻域系的分段线性近似顶点的有效编码方法，使基于六边形邻域系的分段线性近似顶点的编码比特数与通常采用的基于八边形邻域系所需比特数至少相当，证实了能消除轮廓变厚和轮廓点集聚现象的基于六边形邻域系的轮廓表示的有效性. 编码的有效性基于分析得出的下述两个结论：① 在具有六边形邻域系的轮廓网中，在已知最优分段线性表示的当前顶点的条件下，下一个顶点相对于当前顶点只有两个可

能的走向. ② 就同一段折线轮廓而言、如果对其游程长度β采用$(\beta-1)$个 0 加上最后一个 1 的编码方式，则基于六边形邻域系的编码方式只比基于八边形邻域系的方式增加一个标志的比特数(1 比特)

为解决在前向运动预测中存在的由非均匀间距样本求均匀间距样本这样一个困难的问题，我们引入信号处理中的由非均匀采样恢复连续信号的思想，导出非线性内插算法，使前向运动预测的图象质量与通常采用的后向运动预测的图象质量相当.

提出一种属于面向对象的活动图象压缩编码的方案，在时空分割初始化时，采用后向运动补偿，即以I^1为参考帧，对由空域分割生成的I^0中的每一个区域进行后向运动估计，生成初始的均匀时空区域，实现从I^0帧开始的面向对象的编码，而通常的面向对象的活动图象编码皆从I^1帧开始. 在后继帧的编码中，以初始时空分割为基础，采用前向运动补偿、非线性内插等算法，跟踪均匀时空区域的时域变化，维持帧间时空分割的一致性.

关键词 图象压缩. 面向对象的编码. 信源编码. 快速时空分割. 轮廓编码

Abstract

After reviewing the digital video compression encoding approaches based on the source model employed, we point out that the object-oriented encoding at low bit rate is a very promising compression approach which is not only capable of overcoming the blocking artifacts produced in classical translational block-based methods but also of a object-based characteristic belonging to one of the important marks of the latest MPEG-4 standard.

The object-oriented video compression algorithm has exceeded the traditional framework of source model theory. The basic notion behind the technique is to partition a image into several objects with arbitrary shape based on 3-D scene model. Each object is specified by three sets of parameters, namely, motion, shape as well as texture information. Among them, both motion estimation and shape encoding are particularly critical. In contrast to the traditional block based matching technique, the motion estimation in the object-oriented compression approach should be consistent with a representation of a scene in terms of objets. Furthermore, it should aim at reliably estimating the motion of the considered objects rather than providing a mere temporal prediction, and therefore more complex motion models should be used. The shape encoding should not only assure efficiently coding for objects boundary but also provide the advantage permitted manipulating objects easily in the

receiver. Therefore, both the motion estimation module and the shape encoding module are worth further investigation in the framework of the object-oriented motion images encoding. Furthermore, it should be noticed that the modules motioned above are both on the basis of the fast and efficient generation of objects in the scene. In this dissertation , we focus on the issues motioned above and several novel algorithms are presented.

The achievements of the dissertation include:

In this dissertation , a modified quadratic spatial transformation motion estimation is proposed. After performing a image intensity segmentation, in order to overcome the weaknesses of the original algorithm which converge tardily or even diverge in small region, the original algorithm is modified by forward-backward method as well as numerical approach to improve the accuracy and stability of the motion estimation.

A fast temporal-spatio segmentation algorithm using quadratic spatial transformations is presented. Based on the image spatial segmentation and the estimation of the parameters of quadratic spatial transformations motion model, the amount of computation for temporal-spatio segmentation is efficiently decreased through the linearization the estimating for model parameters, householder transformation and the fast approach for $\boldsymbol{R}$, $\boldsymbol{Z}$ matrix.

A optimal contour graph encoding scheme based on both hexagonal neighborhood system and heap data structure is proposed. In this method, we employ maximum absolute distance as distortion measure and divide the contour graph into several contour chains on

which the contour graph is optimally encoded. Since the single-source directed graph corresponding to the contour chain is sparse, we can efficiently reduce the time complexity in Dijkstra algorithm commonly uesd for the shortest-path problem by using a heap data structure. So that the vertex positions in piecewise linear approximation can be efficiently obtained. Furthermore, the region connectivity could not be lost because the optimization is targeted for each contour chain.

In general, the encoding of the nodes is costly, therefore, in this dissertation we proposed a nodes encoding method which set up a tree dynamically in the encoder and decoder simultaneously in the process of deep-first traversing. Therefore, the cost for encoding nodes could be reduced efficiently because when any contour chain is decoded, its starting node could be chosen from the starting nodes of decoded contour chains in the receiver.

A efficient encoding for piecewise linear approximation vertices based on hexagonal neighborhood system approach is proposed. The bit number for optimal piecewise linear approximation vertices in hexagonal neighborhood system is at least comparable to that in octagon neighborhood system commonly used , which assure the validity of accurate contour representation based on the hexagonal neighborhood system. The efficiency is relying on the following conclusions: ①In the contour network with hexagonal neighborhood system, if the current vertex of the optimal piecewise linear approximation is located, then the next optimal vertex only has two possible position relative to the current vertex. ②As for the

same contour chain, if its run length is β and we use the following code word assignment: $(\beta-1)$ zero followed by a final 1, then the bit number needed in hexagonal neighborhood system is at most one bit more than that in octagon neighborhood system.

To solve the difficult problem in reconstructing regularly sampled images from irregularly spaced samples which result from the use of forward motion estimation, the idea for resuming continuous signal from irregularly samples in signal processing is introduced. Based on the idea above, the forward can be performed by a nonlinear interpolation method. Based on the algorithm, the performance of forward motion prediction is very close to that of backward motion prediction.

In the dissertation, we propose a object-oriented motion images compression encoding scheme. In order to produce the original temporal-spatio coherent region in the I^0 frame, we use I^1 frame as reference frame and backward motion estimation for regions resulted from a image intensity segmentation in I^0. Based on this idea, instead of starting the object-oriented encoding from the I^1 frame in the common object-oriented scheme, we realize the object-oriented encoding from the I^0 frame. In the encoding of subsequent frames, on the basis of the original temporal-spatio segmentation, the temporal tracing of coherent regions is enabled through both forward motion prediction and nonlinear interpolation, and therefore the region correspondence is maintained.

Key words image compression, object-oriented encoding, source encoding, fast temporal-spatio segmentation, contour encoding

目　录