

## 摘要

随着视频通信技术的发展,数字视频在军事应用和人们的日常生活中都起着越来越重要的作用。然而,数字视频信号在传输过程中由于信道干扰等原因会出现突发性误码,使得解码终端出现差错块或丢失块,最终导致视频质量下降。因此,除了采用纠错编码、遇错重传等方法外,往往还可以在解码端利用正确接收的信息恢复丢失的信息,这就是基于解码的差错恢复方法即差错掩盖技术。H.26L 和 MPEG-X 等视频编码标准中均采用了差错掩盖技术。对于帧内编码的 I 帧,它不能利用时间方向上的冗余信息进行恢复,只能利用当前帧内的空间信息进行错误恢复。I 帧的误码会导致差错传播到 GOP(图像组)的其他帧。因此寻找恢复准确度高、可实现的帧内差错掩盖方法至关重要。

本论文首先研究了各种基于经典统计数学的帧内差错掩盖算法,包括基于像素插值的方法和基于块匹配的方法,并分析这些算法的优势和局限性。

然后,在研究模糊集理论的基础上,提出两种基于模糊集理论的视频帧内差错掩盖算法,包括基于模糊推理的方法和模糊聚类的方法,详细介绍算法的基本思想和掩盖流程,并重点讨论算法的关键问题。在基于模糊推理的视频帧内差错掩盖方法中,将模糊推理应用于是否接受初步差错掩盖结果的判决上,合理修正掩盖结果,使得边缘连续。在基于模糊聚类的方法中,首先利用帧内差错块周围相邻块的平滑程度和纹理方向等约束条件,对差错块周围的相邻块进行模糊分类,然后用相似块的像素插值得到差错块的像素值。

最后在 H.264 参考软件 JM8.6 的基础上,实现各类算法,并利用不同特性和误码率的序列进行了测试。实验结果表明:文中提出的基于模糊集理论的帧内差错掩盖算法优于基于经典数学的方法,改善了图像的边缘细节和纹理信息,提高了图像的主观质量和 PSNR(峰值信噪比)值。

**关键词:** 视频解码 差错掩盖 I 帧 模糊集 模糊推理和聚类

## Abstract

As the development of video communication technology, digital video is widely used in military application and daily life. However, bit errors or packets lost occur during transmission due to noises in the physical channel and finally lead to the degradation of video quality. Thus besides FEC, ARQ and other error resilience strategies, error concealment (EC) is needed at the decoder side to reduce the degradation and recover the damaged images. The existing international video encoding standards such as H.26L and MPEG-X employ error concealment strategies. As the Intra-frame is the reference of subsequent P or B frames, it cannot be reconstructed by the temporal information. Furthermore, the imperfect recovery of Intra-frame will lead to error propagation to subsequent frames in the same GOP. As a result, looking for an accurate and feasible spatial error concealment algorithm for Intra-frame becomes crucial.

First, this paper briefly introduces various spatial error concealment algorithms based on classical statistics mathematics, including the pixel interpolation methods and the block matching methods. And the limitations and advantages of these algorithms are researched deeply.

Then the algorithms based on fuzzy sets are proposed, including the algorithm based on fuzzy reasoning and the algorithm based on fuzzy clustering. The theory basis, concealment steps and the key problems are discussed in details. Fuzzy reasoning is applied to judge the initial error concealment. If the initial concealment is not accurate, pixel modification is employed to make the edges continuous. Fuzzy clustering is applied to classify the neighbor blocks around the lost block according to the smoothness, direction and other restrictive conditions. Then the lost block is interpolated by the similar blocks.

Finally the experiments are carried out based on the reference software JM8.6 of H.264 and different video sequences are tested. The experimental results show that the algorithms proposed in this paper surpass the classical spatial algorithms, can improve the edge and texture details of the image, and increase the PSNR value as well.

**Key Words:** Video Decoding, Error Concealment, Intra-frame, Fuzzy Sets, Fuzzy Reasoning and Fuzzy Clustering

## 南京邮电大学学位论文独创性声明

本人声明所呈交的学位论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知,除了文中特别加以标注和致谢的地方外,论文中不包含其他人已经发表或撰写过的研究成果,也不包含为获得南京邮电大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

研究生签名: 沈隽 日期: 2008.4.15

## 南京邮电大学学位论文使用授权声明

南京邮电大学、中国科学技术信息研究所、国家图书馆有权保留本人所送交学位论文的复印件和电子文档,可以采用影印、缩印或其他复制手段保存论文。本人电子文档的内容和纸质论文的内容相一致。除在保密期内的保密论文外,允许论文被查阅和借阅,可以公布(包括刊登)论文的全部或部分内容。论文的公布(包括刊登)授权南京邮电大学研究生部办理。

研究生签名: 沈隽 导师签名: 朱俊勇 日期: 2008.4.18

# 第一章 绪论

## 1.1 课题的提出

音频和视频是人类进行信息通信中最直接有效的方式。随着现代通信理论、计算机技术、微电子技术和网络技术的快速发展,信息社会对于多媒体通信有迫切的需求,人们实现 any time, any where, with any person 的多媒体通信也变得越来越接近现实。而图像和视频的传输通信都是以海量数据、超高带宽的需求为特征。IPTV、NGN、3G、NGBW、IP over DVB、三网融合等新技术也都是以高速度、高安全性、高可靠性的数据传输为基础,对大量存储空间和高传输带宽的需求促进了视频编解码技术的飞速发展。因此,作为视频业务及存储应用核心技术的高效率视频数字压缩编解码技术,越来越引起人们的关注,成为目前广播、视频与多媒体通信领域中的热点与亮点。

压缩效率的提高与数据传输的鲁棒性是相互矛盾的,压缩率提高的同时导致了视频数据的抗差错能力较为脆弱。压缩编码去除了图像信息中大量的空间和时间上的冗余信息,导致在信道中传输的视频信息相关性降低<sup>[1]</sup>。因此,抗误码性能急剧降低,误码在解压缩以后又会产生差错扩散,最终导致视频质量下降。在许多有传输效率、实时性等要求的应用场合,纠错编码、自动请求重传等差错控制方法会增加传输的数据量,从而加剧网络的传输拥塞,因而必须寻找更适于视频传输的差错控制方法。由于视频数据在时间和空间上具有高度相关性,传输过程中发生错误的视频数据有可能利用其时间和空间上邻近的数据来掩盖,即差错掩盖技术。

目前的视频编解码技术大多是建立在图像信号的统计特性基础上的,利用图像的空间或(和)时间上的相关性来进行,它要对所解码的图像进行统计分析、计算和处理,属于精确计算方法,与之相应的常见差错掩盖技术也是基于这样的“硬”计算环境的。图像信息不同于数据信息或文本信息,它具有一定的模糊性(Fuzzy),尤其是常见的观赏视频图像更是如此。因此,在允许一定的失真条件下,可以充分利用人眼视觉的这一特性,对所接收的有损图像进行模糊分类及聚合,利用属于同一模糊类的图像部分,通过一定的算法来掩盖因误码而丢失的部分,或对已经掩盖的结果进行模糊推理,修正掩盖的像素值,可以在尽量减少差错对人的视觉效果影响的前提下,达到提高重建图像质量的目的。

## 1.2 视频差错掩盖技术

### 1.2.1 差错掩盖的研究意义和约束条件

图 1-1 是一般视频通信的基本框架，原始的音视频信号（视频一般为 YUV 格式）经过信源编码、压缩、码流复用，将多路音视频信号合成一路信号，经过信道编码、传输、信道解码、解复用、信源解码等一系列的传输变换重新得到接收的视频信号，送到显示终端显示。

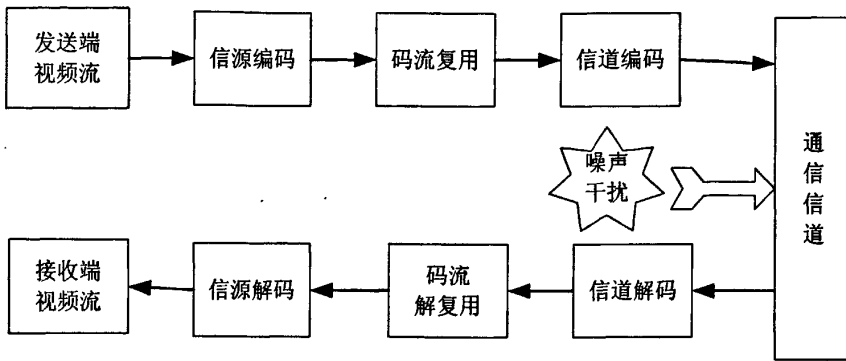


图1-1 视频通信的一般框架

数字视频信号在传输过程中由于线路干扰等原因会出现传输 QoS 下降，尤其是在无线信道和基于 IP 的信道中表现更加严重。无线信道固有的多径效应、瑞利衰落损耗和抖动等特点使得误码率高，稳定性差。视频信号在基于 IP 的网络中传输，和其他业务流量共用网络容易导致网络拥塞，此时会发生丢包现象，同时受到路由选择、信道容量和节电处理能力影响，视频信号还会出现延迟。如果这些误码和丢包在信道编解码过程中得不到有效的控制和纠正，显示终端就会出现马赛克、停顿甚至黑屏现象，严重影响视觉效果和多媒体通信的质量。

因此，为了减小传输错误的影响，抗误码技术成为视频压缩标准中的重要组成部分，如 H.26X 和 MPEG<sup>[2]</sup>。其中，差错掩盖技术不对编码器作任何要求，不增加编码冗余度，因此不增加通信的传输负担；这种方法是属于视频信息的后处理范畴，原则上对任何图像编解码格式和标准都适用，因此其应用范围很广。但是，视频误码的差错掩盖技术是基于不同的假设前提进行的，这些假设前提正是进行差错掩盖的先验约束条件<sup>[3]</sup>。

#### 1) 一致性 (Consistency):

一致性是一个基础的先验约束条件，这是所有差错掩盖方法必须遵守的。它要求差错掩盖算法不能改变已经正确重建的像素值，用差错掩盖算法掩盖后的像素值必须在所能接

受的范围内（如亮度信号为 $[0, 255]$ 之内的整数）。

### 2) 光滑性 (Smoothness):

光滑性也是一个基础的差错掩盖的假设前提。这种光滑性约束条件可应用在时间和空间两个方向上。在空间上，差错掩盖假设出现差错的块和其周围的块是光滑连接的，可以利用周围块的像素值恢复差错块，恢复后的差错块与周围相邻块要像素值线性变化，但是这种光滑性的假设限制了差错块内部的边缘和纹理信息的恢复。

在时间域上，光滑性可解释为当前帧内的某一区域与前一帧中的某一区域相似或相同。基于这种假设，时域掩盖的一种方法是当前帧内的差错块用前一帧中对应位置的块来代替，或者用前一帧的运动矢量估计差错块的运动矢量，然后用它所指示的前一帧中的相应块来代替当前的差错块。这种假设对于快速运动的区域所得到的结果是很不理想的。

### 3) 统计相关性 (Statistical correlation):

统计相关性是视频差错掩盖的另外一种先验性假设约束条件，这种先验性约束条件假设视频中的像素可由某种统计模型生成。源视频中的图像和损坏视频中的图像被建模为随机分布模型，由这种先验模型，再通过最大后验概率 MAP (Maximum a Posteriori Probability) 的方法从正确解码的像素来估计差错的像素。最常用的模型为马尔可夫随机场模型 MRF (Markov Random Fields) 和哈伯-马尔可夫随机场模型 (Huber-Markov Random Field)。这种基于模型的方法时间复杂度相当高，而且在掩盖差错块的边缘方面也有问题。

### 4) 边缘连续性 (Edge continuity):

边缘连续性的假设是指图像中一个区域的边缘（一般指线性边缘）是连续的，这个假设前提暗示：如果一个差错块周围的块内有边缘，而且其方向指示通过有差错的块，那么，这条边缘线一定通过差错的块，需要进行掩盖。如基于凸集投影 POCS (Projection Onto Convex Sets) 的差错掩盖方法就使用该假设进行掩盖。

### 5) 运动矢量的连续性 (Motion continuity)

运动矢量的连续性假设是指差错块的运动矢量与空间和时间上相邻块的运动矢量是相似或相同的。基于这种假设，用平滑滤波或边缘匹配来选择运动矢量，然后用候选运动矢量在参考帧中找到相应的块来代替差错块，最后根据差错块与它周围块的边缘连续性进行修正。

## 1.2.2 视频帧内差错掩盖的重要意义

国际上的视频编码标准(如H.26x和MPEG)普遍采用了基于块的混合编码方式，因此，差错掩盖主要是利用相邻块的空域及时域冗余信息来尽量恢复丢失块的信息<sup>[4][5][6]</sup>。时间域

的差错掩盖技术就是解码器利用与丢失区域在时间上相邻(如前一帧)的图像帧的数据来恢复受损区域;而空间域差错掩盖技术则是利用同一图像帧中的相邻宏块对丢失宏块进行恢复的技术。这种方法也可用在掩盖运动矢量不存在时,对帧内编码的宏块进行差错掩盖。

时域方式主要是利用相邻帧的信息来对丢失块的信息进行恢复,然而视频序列或图像组GOP(Group Of Pictures)的第一帧是帧内编码的I帧,与前后帧的相关性不大。在一般情况下,它不能利用时间方向上的冗余信息进行恢复,只能利用当前帧内的相邻块的信息进行错误恢复。I帧掩盖的不好会导致掩盖误差传播到GOP的其他帧,最终导致视频序列质量的下降。

传统的I帧误码掩盖方法主要是基于像素内插<sup>[7]</sup>或者块匹配<sup>[8]</sup>的方法。基于相邻块像素内插的方法是利用相邻块相应位置的像素值进行加权平均得到一个重建的像素值,用重建的像素值恢复丢失的块。这种方法在平滑区域效果还好,在细节丰富和边缘处会使得图像模糊,主观质量严重降质;在基于块匹配准则的帧内差错掩盖方法中,首先找到与丢失块相似的块,然后用拷贝或内插的方法进行恢复,这种方法容易出现块效应,它使用单一的特征量如MAD(Mean of Absolute Difference)寻找相似块,找出的相似块往往不是最佳的,而且需要多次匹配。因此寻找更佳的I帧的掩盖技术日益受到关注。

### 1.2.3 国内外视频帧内差错掩盖的研究现状

近年来对视频差错掩盖技术的研究一直是国内外的研究热点领域之一,发表了不少有关算法的论文,有些算法已经付诸实施,取得了长足的进展。空间域的方法主要使用了块拷贝和插值方法,如MPEG-2中的块匹配方法和H.264中使用的加权平均插值方法,这些方法具有简单实用的特点,但只适用于平滑区域,对有边缘的区域恢复的图像会产生模糊。后续的改进方法有方向插值算法<sup>[9]</sup>、基于凸集投影(POCS)的算法<sup>[10]</sup>和基于马尔可夫随机场(MRF)的方法<sup>[11]</sup>等,这些方法可以恢复出图像的边缘信息。对于比较规则或者对称的图像,还可以利用图像的远端相关性,类似于帧间的差错掩盖方法,利用包含误码块的本地窗口寻找与它匹配的远端窗口,利用该窗口中和误码块对应的块代替。

但这些算法基本上都是基于图像统计特性的,仍然存在差错掩盖效果不佳、算法复杂度较高的缺陷。因此,采用何种办法,如何进一步提高掩盖图像质量、降低算法的耗时,仍然是一个急需研究和解决的技术难题。

## 1.3 论文结构和主要研究内容

本论文研究的重点是基于模糊数学的视频帧内差错掩盖技术,主要是根据视频当前帧

内的图像像素相关的模糊性质，利用模糊数学的方法进行分析和研究，获得性能更好的视频差错掩盖算法，使得用它掩盖后的视频主客观质量比用传统方法掩盖后的视频质量要好。本论文的主要工作是：

1. 研究和实现多个典型的帧内差错掩盖算法；
2. 研究模糊数学的经典理论，并分析图像中存在的模糊性；
3. 提出新的基于模糊数学的帧内差错掩盖算法；
4. 实现基于模糊聚类和模糊推理的帧内差错掩盖程序，并与经典算法比较。

本文的结构如下：第一章主要介绍本文的研究背景。第二章给出了模糊数学（模糊集）的基本理论，重点阐述了模糊聚类和模糊推理两大分支的基本理论，对其在信号处理中的应用进行了总结，探讨了本论文研究的数学基础和研究方法。第三章介绍几种经典的帧内差错掩盖算法及其改进算法，并分析其优势及存在的局限性。第四章介绍基于模糊推理的帧内差错掩盖算法，并讨论隶属度函数和推理规则的选取。第五章提出基于模糊聚类的帧内差错掩盖算法，讨论聚类特征量及相似度的选取。第六章使用不同特性和误码率的序列进行实验，并进行了分析。最后对全文做出总结，并提出了下一步工作的方向。



## 第二章 模糊集 (Fuzzy Sets) 理论

扎德 (L. A. Zadeh) 在 1965 年提出了模糊集的概念, 在经典数学之外开拓了模糊数学的先河。模糊数学是建立在模糊集合之上的, 它是描述和处理人类语言中所特有的模糊信息的理论。它是用精确的数学方法处理模糊对象。其中模糊推理和模糊聚类是模糊集理论的两大分支, 在控制论和图像处理中获得广泛的应用。本论文主要研究了基于模糊数学的方法处理视频的差错掩盖, 以提高差错视频的掩盖质量。

### 2.1 模糊集理论的提出

#### 2.1.1 经典集合及其特征函数

在数学上, 概念的外延可以通过“集合”来表达, 即Cantor集合理论<sup>[12]</sup>。集合即“把一定的并且彼此可以明确识别的东西—东西可以是直观的对象, 也可以是思维的对象—放在一起”。例如全体自然数、全体有理数、全体实数都是集合, 集合中的每一个具体的对象 (成员) 称为该集合的元素。事实上, 集合与一个概念在人脑中的形成密切相关。一个概念的形成大致需要包含两个方面: 一方面是从内在条件把握各个有关因素对这个概念所作的规定, 即这个概念的内在涵义, 称之为概念的“内涵”; 另外一方面就是这个概念所包含的内容, 换句话说就是符合此概念的事物全体, 我们称之为概念的“外延”。内涵和外延是描述概念的两个方面, 它们是相辅相成的。

设  $A$  为集合, 若具体对象  $a$  是集合  $A$  的元素, 则称  $a \in A$ , 即  $a$  属于  $A$ , 否则  $a \notin A$ , 即  $a$  不属于  $A$ 。 $a$  与  $A$  之间的隶属关系可以用  $a$  的特征函数刻画, 即

$$\mu(a) = \begin{cases} 1, & a \in A \\ 0, & a \notin A \end{cases} \quad (2-1)$$

其中  $\mu(a)$  为元素  $a$  相对集合  $A$  的特征函数。

从Cantor集合论可以看出, 在经典数学中, 任何一个元素和一个集合只有一种关系, 即  $a$  属于  $A$  或者  $a$  不属于  $A$ , 它是建立“二值逻辑”的基础上的, 具有“非此即彼”的关系。

#### 2.1.2 模糊集与隶属度函数

一个经典集合的“内涵”和“外延”都必须明确的, 所以对论域中的任何元素, 或

者属于某一集合，或者不属于这个集合，两者必居其一且仅居其一。然而在现实世界中，有许多概念并没有明确的“外延”，如“天气好”、“水温很低”、“个子很高”等都是模糊的概念。这样就使得经典集合论对于这样的概念显得力不从心了，因为模糊概念很难简单地用“属于”或“不属于”来表示，而只能通过属于的程度来描述。换句话说，就是论域中的元素符合某一概念的程度不能仅仅用0或1来表示，而需要借助于介于0到1之间的实数来表示。

论域  $X$  上的“模糊集合”  $A$  [13][14] 定义为：

$$A = \{(x, A(x)) | x \in X\} \quad (2-2)$$

或者

$$A = \{(x, \mu_A(x)) | x \in X\} \quad (2-3)$$

其中  $A(x)$  或者  $\mu_A(x)$  称为“隶属度函数”，它满足： $A: X \rightarrow M$ ，这里  $M$  称为“隶属空间”，最常见的隶属空间为区间  $[0, 1]$ 。隶属度函数  $\mu_A(x)$  用于刻画元素  $x$  对模糊集合  $A$  的隶属程度，即“隶属度”，所以模糊集合  $A$  的每个元素  $(x, \mu_A(x))$  都能明确地表现出  $x$  的隶属等级。 $\mu_A(x)$  的值越大， $x$  的隶属程度就越高。当  $\mu_A(x) = 1$  时，说明  $x$  完全属于  $A$ ；而  $\mu_A(x) = 0$  时，说明  $x$  不属于  $A$ ；而  $\mu_A(x)$  值介于0和1之间时，说明隶属于  $A$  的程度也介于“属于”与“不属于”之间，是模糊的。与经典集合类似，在模糊集合的表示中，对于隶属度为0的元素可以不列出。

由定义可以看出，模糊集合  $A$  是由隶属度函数  $\mu_A(x)$  唯一确定的，可以把模糊集合  $A$  与隶属度函数  $\mu_A(x)$  看成是等同的。隶属程度的思想是模糊数学的基本思想。当  $\mu_A(x)$  仅取值0和1时，模糊集合退化为经典集合，可见经典集合是模糊集合的特殊情形。 $X$  上的所有模糊集合组成的集合称为  $X$  的模糊幂集，记为  $\Gamma(X)$ 。

## 2.2 模糊推理 (Fuzzy Reasoning)

推理就是根据已知的一些命题按照一定的法则去推断一个新的命题的思维过程和思维方式，简言之从已知条件求未知结果的思维过程就是推理。模糊逻辑推理是一种不确定性的推理方法，其基础是模糊逻辑，它是根据给定的行为规则集来得到输入信息对应的隶属度值的过程。

模糊推理是对模糊信息进行处理的方法,也是模糊数学中除了隶属度函数之外的又一个重要的组成部分<sup>[15][16]</sup>。由于它在诸多工业生产领域上的成功应用,使它在近20多年来模糊系统理论的研究中始终占有重要的地位。Zadeh于1973年第一次提出了模糊分离规则FMP(Fuzzy Modus Ponens),并被Mamdani等人所发展,形成了如今被广泛使用的CRI(Compositional Rule of Inference)等方法<sup>[17]</sup>。

模糊推理作为近似推理的一个分支,以数值计算而不是以符号推演为特征。正因如此,它与经典逻辑有明显的不同。它不注重基于公理的形式推演,甚至也没有基于赋值的语义运算。这也是模糊推理区别于人工智能方法的特征所在。

### 2.2.1 模糊推理的基本思想

一个系统的输出与预定的标准之间总是有误差的,以 $e$ 记此误差。这个误差又是随着时间而变化的,以 $e^0$ 记误差变化率。所谓推理就是要根据 $e$ 与 $e^0$ (或许还要 $e$ 的多阶导数)对系统的输入进行调整。经典控制理论中这个调整量是 $e$ 与 $e^0$ 的函数 $f(e, e^0)$ 。在某些情况下,这个 $f$ 是不易建立或不需要建立的,取而代之的是已知一组专家经验,即已知当误差为 $e'$ 且变化率为 $e^{0'}$ 时应当采取的调整量为 $c_i(i=1,2,\dots,n)$ ,并要在这一组经典情况的基础上针对随时测得的 $e'$ 与 $e^{0'}$ 计算出相应的控制量 $c^*$ 来。模糊推理的基本原理是<sup>[18]</sup>:

第一步,把 $e'$ ,  $e^{0'}$ ,  $c_i$ ,  $e'$ 与 $e^{0'}$ 分别模糊化为 $X$ ,  $Y$ ,  $Z$ ,  $X$ 与 $Y$ 上的模糊集 $A_i$ ,  $B_i$ ,  $C_i$ ,  $A^*$ 与 $B^*$ ;

第二步,列出算式:

$$\begin{array}{ll}
 \text{已知} & A_1 \text{ 且 } B_1 \rightarrow C_1 \\
 & \dots\dots \\
 & A_n \text{ 且 } B_n \rightarrow C_n \\
 \text{且给定} & A^* \text{ 且 } B^* \\
 \hline
 \text{求} & C^*
 \end{array} \quad (2-4)$$

第三步,将 $c^*$ 去模糊(Defuzzify)后就得到最终的数值控制量 $c$ 。以上的第二步就是模糊推理。由于已知条件 $A_i$ 且 $B_i$ 可用乘积 $X \times Y$ 上的一个模糊集去取代,且 $n$ 条规则可以通过聚合(Aggregate)而成为一条超规则,或者可以分别使用这 $n$ 条规则单独推理后将所得 $n$ 个中间结果以某种方式合成为最终的 $C^*$ ,所以模糊推理可归结为以下最基本的形式:

$$\begin{array}{l}
 \text{已知} \quad A \rightarrow B \\
 \text{且给定} \quad A' \\
 \hline
 \text{求} \quad B'
 \end{array} \tag{2-5}$$

这里  $A, A'$  是  $X$  上的模糊集,  $B, B'$  是  $Y$  上的模糊集.

求解方程 (2-5) 的传统方式即 Zadeh 于 1975 年提出的 CRI 算法. Zadeh 提出了一种蕴涵算子  $R_z : [0,1]^2 \rightarrow [0,1]$  如下:

$$R_z(a,b) = a' \vee (a \wedge b) \tag{2-6}$$

这里  $a'$  表示  $1-a$ ,  $\vee$  和  $\wedge$  分别表示取上、下确界运算. CRI 算法的第一步是利用蕴涵算子  $R_z$  把已知条件  $A \rightarrow B$  转化为  $X \times Y$  上的一个模糊关系  $R(x,y)$  如下:

$$R(x,y) = R_z(A(x),B(y)), (x,y) \in X \times Y \tag{2-7}$$

然后在第二步用  $A'$  与  $R$  复合就得出  $B'$ , 即  $B' = A' \circ R$ . 这里复合运算 “ $\circ$ ” 的具体表达式为

$$\begin{aligned}
 B'(y) &= \sup_{x \in X} [A'(x) \wedge R(x,y)] \\
 &= \sup_{x \in X} [A'(x) \wedge R_z(A(x),B(y))], \quad y \in Y
 \end{aligned} \tag{2-8}$$

这样就求出了最终答案  $B'$ .

### 2.2.2 模糊推理系统

一个模糊推理系统的原理框图如图 2-1 所示<sup>[19]</sup>:

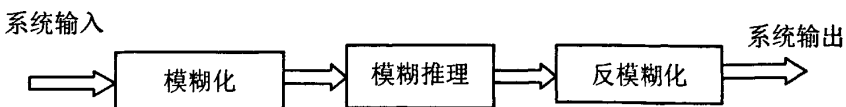


图 2-1 模糊系统

模糊推理系统通常分为 3 个部分: 模糊化、模糊推理和反模糊化. 在模糊化中首先确定输入模糊集合和对应的模糊隶属度函数. 传统的集合是一种“非此即彼”的思想, 一个元素是否属于该集合是确定的, 可以用 true 或 false 表示; 而在模糊集合中, 一个元素属于该集合利用隶属度来表示, 即该元素在多大程度上属于该模糊集, 隶属度取值在 0 到 1 之间. 模糊隶属度函数定义了输入空间的某点与隶属度之间的映射关系. 整个模糊化的过程就是将输入的确切值用各个模糊集合的隶属度来表示.

模糊推理过程利用模糊逻辑将由模糊化过程得到的输入模糊集的隶属度映射到输出模糊集的隶属度. 经常用到的模糊逻辑有 AND 和 OR 等, 对于模糊集中的 AND 逻辑来

说, 取隶属度最小, 而 OR 取隶属度最大。在模糊推理过程中要利用模糊规则, 常用的是 IF-THEN 结构, 如 “IF(x is A)AND(y is B) THEN z is C”, 其中, A/B 为输入模糊集, C 为输出模糊集, 利用模糊逻辑可以得到条件成立的隶属度, 那么结论的隶属度与之相同。在一般应用中多个模糊规则, 每个模糊规则确定了输出在某个输出模糊集的隶属度, 在各个模糊集的重合部分取隶属度最大值, 这样得到最终输出的隶属度图形, 这种推理方法称为最大-最小法。

在模糊过程中, 得到了输出的隶属度图形, 最后的反模糊化过程是要将不确定性映射到确定的输出, 如取隶属度图形的重心、二分法等。

## 2.3 模糊聚类 (Fuzzy Clustering)

聚类就是按照一定的要求和规律对事物进行区分和分类的过程, 即 “物以类聚”。其目的是使得属于同一类别的个体之间的距离尽可能小, 而不同类别的个体之间的距离尽可能的大。模糊聚类在这一过程中没有任何关于类分的先验知识, 仅靠事物之间的相似性作为类属划分的准则, 因此属于无监督分类的范畴<sup>[20][21]</sup>。

聚类分析则是指用数学的方法研究和处理给定对象的分类。它是数理统计中多元统计分析的一种, 也是非监督模式识别的一个重要分支。它把一个没有类别标记的样本集按照某种准则划分成若干个子集 (类), 使相似的样本尽可能归为一类, 而不相似的样本尽可能划分到不同的类中。聚类分析算法除了广泛应用于模式识别、图像分割、特征匹配等领域外, 还在心理学、生理学、医学和地理学等领域中有着广泛的应用。

### 2.3.1 传统聚类分析

传统的划分方法是一种硬划分, 是把每个待处理的对象严格地划分到某个类中。硬划分方法的典型代表是C-均值算法<sup>[22]</sup>。

设  $X = \{x_1, x_2, \dots, x_n\} \subset R^s$  是一个数据集,  $u = \{u_{ik}\}_{c \times n} \in M_{fcn}$  是一个隶属度矩阵,  $v = \{v_1, v_2, \dots, v_c\}$  是  $c$  个聚类中心,  $v_i \in R^s$ ,  $2 \leq c < n$ 。

C-均值算法把  $n$  个向量  $x_i$  ( $i=1, 2, \dots, n$ ) 分成  $c$  个簇  $G_i$  ( $i=1, 2, \dots, c$ ), 并求得每个簇的聚类中心, 使得簇内方差的和达到最小  $J(u, v) = \sum_{k=1}^n \sum_{i=1}^c u_{ik} \|x_k - v_i\|^2$ , 其中  $\sum_{i=1}^c u_{ik} = 1, u_{ik} \in \{0, 1\}$ 。C-均值聚类算法的基本步骤如下:

(1) 初始化: 给出初始聚类中心  $v^{(0)} = \{v_1^{(0)}, v_2^{(0)}, \dots, v_c^{(0)}\}, l=0$ ,  $l$  为迭代次数, 最大迭代次数为

$T$ , 阈值为  $e$ 。

(2) 用下列公式更新  $u_{ik}^{(l+1)}$ ：

$$u_{ik}^{(l+1)} = \begin{cases} 1, & \text{if } i = \arg \min \{\|x_k - v_i^{(l)}\|\} \\ 0, & \text{otherwise} \end{cases} \quad (2-9)$$

(3) 用下列公式更新  $v_i^{(l+1)}$ ：

$$v_i^{(l+1)} = \frac{\sum_{k=1}^n u_{ik}^{(l+1)} x_k}{\sum_{k=1}^n u_{ik}^{(l+1)}} \quad (2-10)$$

如果  $\max_i \|v_i^{(l+1)} - v_i^{(l)}\| < \varepsilon$  或者  $l > T$ , 则停止；否则,  $l = l + 1$ , 转至步骤(2)。

C-均值算法思想简单, 实现容易, 收敛快, 运行速度快, 内存消耗小, 能有效地处理大数据集, 是目前最常用的聚类算法之一。但是, C-均值算法也有不少缺点, 如采用的是硬划分, 每类由类中心代表, 每个数据点的影响一样, 没有考虑噪声数据的影响, 也缺少对类中心的约束等。

### 2.3.2 模糊聚类分析

传统的聚类分析是一种硬划分, 它把每个待识别的样本对象严格地划分到某类中, 具有非此即彼的性质, 因此这种类别划分的界限是鲜明的。而实际上大多数对象并没有严格的属性, 它们在形态和类属方面存在着中介性, 具有亦此亦彼的性质, 适合于软划分。模糊集理论为这种软划分提供了有力的分析工具, 人们开始用模糊的方法来处理聚类问题, 即模糊聚类分析。由于模糊聚类得到了属于各个类别的不确定性程度, 表达了样本类属的中介性, 即建立起样本属于各个类别的不确定性的程度的描述。

给定数据集  $X = \{x_1, x_2, \dots, x_n\} \subset R^r$  为模式空间的一组有限观测样本集。

$x_k = (x_{k1}, x_{k2}, \dots, x_{kr})^T \in R^r$  为观测样本  $x_k$  的特征矢量或模式矢量。由该数据集得到  $C$  个子集  $X_i (i = 1, 2, \dots, c)$  满足<sup>[23]</sup>：

$$\begin{aligned} \bigcup_{i=1}^c X_i &= X \\ X_i \cap X_k &= \emptyset, 1 \leq i \neq k \leq c \end{aligned}$$

$$\begin{aligned}
 & X_i \neq \emptyset, X_i \neq X_j, 1 \leq i \neq j \leq c \\
 & M_{hc} = \{U = [u_{ik}]_{c \times n} \mid u_{ik} \in [0, 1], \forall i, k; \sum_{i=1}^c u_{ik} = 1, \forall k; 0 < \sum_{k=1}^n u_{ik} \leq n, \forall i\} \quad (2-11)
 \end{aligned}$$

其中,  $u_{ik}$  为样本  $x_k$  隶属于子集  $X_i$  的程度。

C-均值算法的隶属度要么是1, 要么是0, 这不能反映数据点与类中心的实际关系。为了处理这个问题, 模糊聚类算法引入每个样本属于每个子集的隶属度, 其主要思想是将经典划分的定义模糊化, 聚类的过程与经典C-均值聚类相似。

### 2.3.3 基于模糊等价关系的模糊聚类分析

从实现方法上分, 聚类分析方法可大致分为四种类型: 谱系聚类法、基于等价关系的聚类方法、图论聚类法和基于目标函数的聚类方法<sup>[24][25]</sup>。

谱系聚类法也叫分层聚类分析法, 将模式样本按距离准则逐步聚类, 类别由多到少, 直到满足合适的分类要求时为止。这种聚类方法是由不同层次的分割聚类组成, 层次之间的分割具有嵌套关系。图论聚类方法最早由Zahn提出, 又称为最大(小)支撑树聚类算法, 其基本思想是, 画出以被分类元素为顶点, 以模糊相似矩阵 $R$ 的元素 $r_{ij}$ 为权重的一棵最大的树, 取定 $\lambda \in [0, 1]$ , 砍断权重小于 $\lambda$ 的枝, 得到一棵不连通的图, 各个连通的分支便构成了在 $\lambda$ 水平上的分类。基于目标函数的聚类方法是把聚类分析归结为一个带约束的非线性规划的问题。

基于等价关系的聚类分析较为常用, 它的思想来源于等价布尔矩阵, 构造等价关系完成划分<sup>[26]</sup>。在一般的情况下, 利用初始样本构造的模糊关系仅仅能满足自反性和对称性, 而不满足传递性, 所以生成的只是一个模糊相似矩阵 $R$ , 而不是模糊等价矩阵, 所以为了进行分类, 还要在这个模糊相似矩阵的基础之上生成一个模糊等价矩阵, 最自然的方法就是去求该模糊相似矩阵 $R$ 的传递闭包 $t(R)$ , 这样便可以得到一个模糊等价矩阵。当生成模糊等价矩阵后, 取某一实数 $\lambda \in [0, 1]$ , 计算出 $t(R)$ 上的布尔矩阵 $P$ 便得到论域 $X$ 的一个等价划分: 当 $p_{ij} = 1$ 时说明 $x_i$ 与 $x_j$ 在同一个等价类中; 否则不在同一个等价类中。如果依次将 $\lambda$ 值从1变小至0时, 便可以得到 $X$ 的一个逐渐由细变粗的动态分类。在这个方法中, 由于模糊等价矩阵是采用传递闭包的方法得到的, 故此方法也称为传递闭包法。

基于等价关系进行模糊聚类的过程, 可以归纳为以下的两个步骤:

(1) 生成模糊等价矩阵: 由一个模糊相似矩阵通过求闭包生成一个模糊等价矩阵。给

定数据集  $X = \{x_1, x_2, \dots, x_n\} \subset R^r$  为模式空间的一组有限观测样本集。  $x_k = (x_{k1}, x_{k2}, \dots, x_{kr})^T \in R^r$  为观测样本  $x_k$  的特征矢量或模式矢量。计算  $X$  中样本  $x_i$  和样本  $x_j$  相似的程度  $r_{ij}$  ( $0 \leq r_{ij} \leq 1, (i, j = 1, 2, \dots, n)$ ), 由此得到论域  $X$  上的一个模糊相似关系  $R$ , 通过合成运算得到  $R$  的等价矩阵  $\bar{R}$ 。

(2) 划分: 由大到小, 依次取实数  $\lambda \in [0, 1]$ , 计算  $\bar{R}_\lambda$ , 再根据  $\bar{R}_\lambda$  对  $X$  进行划分。最后便得到  $X$  在不同的水平下对事物的划分。

## 2.4 模糊集理论的应用

模糊集合的理论和方法是一种能够有效针对那些非精确、不完全类信息的有效处理手段, 从单纯的“硬”计算环境拓展至更为广阔的“软”计算 SC (Software Computation) 领域, 它与人脑的功能相对应。自从 1965 年 Zadeh 提出模糊集理论之后, 模糊集的理论和方法在各方面取得了广泛的应用, 特别是在控制和图像处理方面, 其应用已经比较成熟, 如图像的模糊形态学处理、模糊增强、模糊分类和模糊边缘检测等<sup>[27][28][29][30]</sup>。最近几年来, 将模糊集合理论和方法应用到对视频信号的处理和压缩方面也取得了进展<sup>[31][32]</sup>。

在模式识别中, 两个最主要的分支为有监督分类和无监督分类两个方面, 其中无监督分类与聚类分析相对应。在没有训练样本的情况下, 模糊聚类可根据数据集的内在结构通过机器学习自动划分特征空间, 达到自动分类的目的。基于模糊规则的推理是用人类推理的语言来描述如何处理的指示, 即用普通语言形式描述处理的规则, 并可以转化为数学算子的结构, 可以广泛应用于控制理论和数字视频的后处理中。

在图像处理中, 由于人眼视觉的主观性使图像适合于模糊手段处理, 训练样本图像的匮乏又需要无监督分析, 而模糊聚类正好满足这两方面的要求。目前, 模糊集理论被广泛应用于图像分割、边缘检测、图像增强、图像压缩、图像平滑、图像匹配等众多方面, 取得了丰硕的成果。但是将模糊集技术应用到图像传输领域, 特别是应用到误码处理, 如差错掩盖等方面, 国内外这方面的理论研究刚刚开始。基于统计学的差错掩盖算法都是满足某单一约束条件的最优化来实现的, 而实际上由于视频信息的丰富性, 基于这种单一约束条件的掩盖算法是不可取的。而图像信息的模糊性可以将模糊集理论引入差错掩盖中, 充分利用视频信息。



## 2.5 本章小结

本章主要总结了模糊数学的基本理论及其在信号处理中的应用。模糊推理和模糊聚类分析是模糊数学研究的两大主要分支，应用范围比较广，尤其是在图像处理方面的研究较为深入，但把模糊集理论应用于视频差错掩盖方面的研究，国内外还鲜有报道。本章第一小节给出了模糊集合的背景和定义，重点介绍隶属度函数。第二小节和第三小节分别介绍了模糊推理和模糊聚类的基础知识，重点阐述了模糊推理的基本思想和基于等价类的模糊聚类方法，第四小节总结了模糊集理论在信号处理方面的应用，重点指出了把模糊数学应用于视频差错掩盖研究的基础和意义。

## 第三章 基于经典数学的视频帧内差错掩盖

尽管在通信底层及系统层采用了各种纠错手段,由网络引起的传输差错无法彻底被去除,仍有残留差错存在。在此情况下,解码器可以利用差错掩盖技术尽可能减小残留差错对重建视频质量的影响,即解码器试图用主观可以接受的、近似原始质量的视频数据来遮盖差错受损数据,而不需要从编码器得到额外的信息。

现有的视频编码标准(如 H.26x 和 MPEG)普遍采用了基于块的混合编码方式,受损的宏块 MB(MacroBlock)中有三种类型的信息需要估算和恢复:纹理信息、运动向量以及编码模式(帧内/帧间: Intra/ Inter)<sup>[33]</sup>。恢复的基础是自然景物的低频特性,即空间和时间上相邻的像素具有平滑性,因此,差错掩盖主要是利用相邻块的空域及时域冗余信息来尽量恢复丢失块的信息。现有的视频帧内差错掩盖主要是利用经典数学,由当前帧内的像素值估计丢失的像素值。

### 3.1 简单加权插值算法

#### 3.1.1 算法描述

当受损块前面没有参考帧时(如I帧内的宏块),就不能用基于时域掩盖的方法来掩盖。但是图像内部空间相邻的像素是高度相关的,可以用受损块的四周邻块的像素值来估计受损块的像素值,这就是基于空域掩盖的空间插值法。简单加权插值算法就是利用受损块周围的一圈像素点来进行插值<sup>[34]</sup>。

在确定受损块的可用相邻块之后,即进行逐像素的恢复。首先根据可用相邻块的起始坐标获取可用相邻块与受损块相邻的边界像素值。对受损块内的每一像素点,用与它最邻近的上下左右四个可用边界像素点进行加权平均。每个像素点的权重与它到当前丢失像素点的距离成反比。公式如下:

$$p(i, j) = \frac{(p_T * d_B + p_B * d_T + p_L * d_R + p_R * d_L)}{(d_T + d_B + d_L + d_R)} \quad (3-1)$$

其中  $p_T = p(i, N+1)$ ,  $p_B = p(i, 0)$ ,  $p_L = p(0, j)$ ,  $p_R = p(N+1, j)$ ,  $d_T$ ,  $d_B$ ,  $d_L$ ,  $d_R$  分别是四个可用边界像素点与待恢复像素点之间的距离,  $N$  为恢复块的大小,如图 3-1 所示。

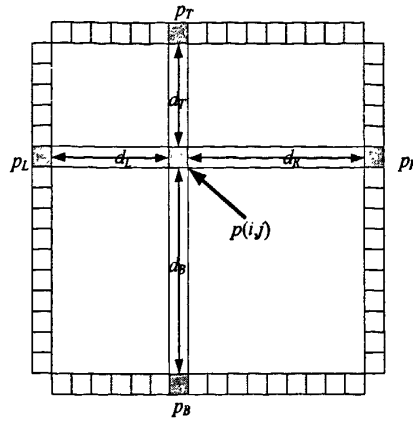


图 3-1 简单加权插值

### 3.1.2 H.264 帧内差错掩盖流程

在 H.264 参考软件 JM 中，没有完整性或比特差错检查，所有出错或不完整的 Slice 视为丢失，在解码前丢弃，不会被解码。在所有正确接收的 Slice 被解码后，开始对丢失的块进行差错掩盖。I 帧的重建采用简单加权空间插值的掩盖方法。掩盖的过程基于宏块的状态图，如图 3-2 所示。在此状态图中，若可对 Slice 中的 MB 进行解码，则将此 MB 的状态标记为“已正确接收的”，反之，则将 MB 标记为“丢失的”，掩盖完毕的 MB 标记为“已掩盖的”。

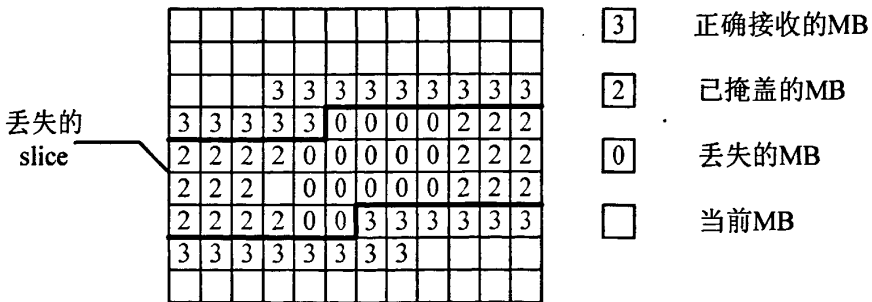


图 3-2 解码器中的宏块状态

I 帧的掩盖很重要，因为如果掩盖的不好会导致掩盖误差传播到 GOP 的其他帧，最终导致视频序列质量的下降。因此掩盖顺序也很重要，在 H.264 中，掩盖过程以宏块为单位进行，从边缘开始，逐列进行，一列一列向内推进，同一列向帧内部移动，以避免掩盖差错。

在参考算法中，如果有两个以上(考虑边角块的情况下，至少需要四个块)相邻的“已正确接收的”宏块存在时，仅用这些“已正确接收的”宏块进行掩盖差错，否则，相邻的“已掩盖的”宏块也参与像素加权平均。

### 3.1.3 算法优势及存在缺陷

简单加权空间内插的基本思想是：离受损点越近的边界像素点被赋予的权重越大，其权重的大小与受损点到边界的距离成反比。这种方法是利用相邻块的像素内插的方法得到丢失块的像素值，该算法对视频帧内比较平滑的区域有比较好的效果，但是对具有很多细节和纹理的区域则会造成图像模糊。H.264 所建议的边界单像素空间内插方法，虽然注重了相邻图像块之间的相关性与平滑性问题，但是并没有涉及邻近块的纹理边缘信息与边缘方向，只采用了丢失宏块在垂直与水平两个方向上邻近边界像素值作插值。而且，H.264 参考算法是水平与垂直方向加权插值空域掩盖算法，当水平方向或垂直方向的相邻宏块也属于丢失区域而不能用作空间插值时，那么二维插值滤波器会降为一维滤波器，从而降低了空域掩盖的有效性。

## 3.2 块匹配差错掩盖算法

### 3.2.1 基本的块匹配差错掩盖算法

基于块匹配的差错掩盖方法 BMA(Block Matching Algorithm)<sup>[35]</sup>也是一种常用的方法，基本的块匹配算法如 3-3 所示，设丢失块大小  $M \times M$ ，以丢失块为中心确定一个  $N \times N$  的操作窗  $W(p, q)$ ， $p, q = 1, 2, \dots, N$ ，在这个范围内进行像素匹配。将丢失块及其周围一圈的像素构成一个  $(M+1) \times (M+1)$  的本地窗  $y(i, j)$ ， $i, j = 1, 2, \dots, M+1$ ，用这个本地窗去和操作窗口中不包含受损块的任意一个  $(M+1) \times (M+1)$  的窗（称为匹配窗） $y_0(i, j)$ ， $i, j = 1, 2, \dots, M+1$ ，进行匹配，计算它们周围一圈像素均方误差 MSE(Mean Square Error)或者绝对误差总和 SAD(Sum of Absolute Difference)，并将其定义为匹配差异度。匹配差异度最小的窗就称为最佳匹配窗，最佳匹配窗中除外面一圈像素外中间的  $M \times M$  块就称为最佳匹配块  $x(m, n)$ ， $m, n = 1, 2, \dots, M$ 。

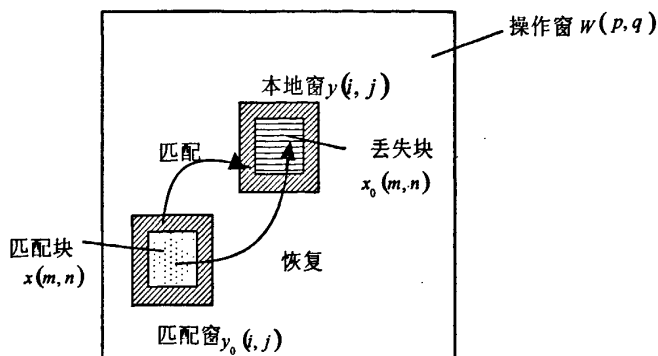


图 3-3 像素匹配过程

周围一圈像素均方误差为:

$$MSE = \sum_{m,n \in \Psi} [x(m,n) - x_0(m,n)]^2 \quad (3-2)$$

周围一圈像素的绝对误差总和为:

$$SAD = \sum_{m,n \in \Psi} |x(m,n) - x_0(m,n)| \quad (3-3)$$

其中:  $\Psi$  为匹配窗(本地窗)周围一圈像素的集合。 $DIF = \min\{MSE\}$  或者  $DIF = \min\{SAD\}$  为匹配差异度。

### 3.2.2 块划分及改进的块匹配算法

基本的块匹配算法利用了图像像素间的远程相关性,但是匹配块的选取是逐像素移动并计算 MSE 或者 SAD,同时块尺寸较大,计算量比较大;而且匹配准则是基于单一匹配特征量 MSE 或者 SAD,找到的相似块不一定准确。后续的改进算法也是针对这两个问题:减小块的尺寸或者增加匹配量。

分裂式块匹配方法<sup>[36]</sup>主要是解决块尺寸较大的问题,对匹配块进行划分,迭代恢复丢失块。这种方法主要由四步组成,它所使用的匹配准则是通常的 MAD (Mean Absolute Difference),如公式 (3-4) 所示

$$MAD = \frac{1}{b_x \times b_y} \sum_{x=0}^{b_x} \sum_{y=0}^{b_y} |B_T(x,y) - B_B(x,y)| \quad (3-4)$$

其中,  $B_T$  和  $B_B$  是两个  $b_x \times b_y$  大小的块。当 MAD 最小时,为最佳匹配块。

以  $16 \times 16$  丢失块为例,分裂式块匹配算法的主要步骤为四步:

(1) 初始化,在初始化时试图匹配垂直相邻的两个  $16 \times 8$  的块,如果这两个块被认为是相似的,则通过拷贝操作对差错的块进行掩盖,否则进入第二步。

(2) 第一次分裂,上一步的两个块被分裂为四个  $8 \times 8$  小块  $b_1, b_2$  和  $t_1, t_2$ ,然后分别对  $b_1$  与  $t_1$  和  $b_2$  与  $t_2$  进行匹配,如果有匹配的块,则差错块在两块之间的部分通过拷贝方法掩盖,如果至少有一对不能匹配,则进入第三步。

(3) 第二次分裂。把第二步中的块分裂成更小的  $4 \times 4$  的块,为  $b_i$  和  $t_i$ ,然后对下面的块和上面的块进行两两的匹配,对找到的匹配块之间的部分用拷贝的方法进行掩盖,对这种方法连续的进行,直到找不到匹配块或整个 MB 都已做完掩盖。

(4) 如果第三步还有未掩盖的像素,则扩大匹配块的范围,重复第三步的方法。

基于统计量的块匹配算法是解决单一匹配特征量的问题,采用双限制条件提高匹配的

准确度。假设两个要匹配的块的均值和方差分别为  $m_T$ 、 $m_B$  和  $\sigma_T^2$ 、 $\sigma_B^2$ ，则匹配公式如下：

$$(|m_T - m_B| < Q^* \sigma_T) \quad \text{和} \quad (|m_T - m_B| < Q^* \sigma_B) \quad (3-5)$$

### 3.2.3 块匹配差错掩盖算法分析

块匹配算法利用丢失块周围的像素值在当前帧内寻找相似块，并拷贝相似块的像素值来代替丢失块的像素值，为了提高效率和增加匹配的准确度，后续的改进块匹配算法将待匹配块划分成小的尺寸或者增加匹配限制条件。这种方法也经常用于时域掩盖中估计丢失块的运动矢量 MV (Motion Vector) [37][38]。

这种方法在运动平滑的区域能够取得较好的效果，但是在运动比较剧烈或者场景切换的情况下，这种方法会产生明显的块效应。同时寻找相块的过程中，它采用较少的相似特征量或统计量（均值和方差）作为判决准则，这种硬判决方法很容易使找到的不是最佳相似块，而且寻找相似块的过程需要多次匹配。

## 3.3 基于方向插值的差错掩盖算法

边缘意味着像素值剧烈变化或不连续，它在人眼的感觉中扮演着重要的作用。在人类视觉系统中，边缘连续而且完整很重要。通常观察者认为主观上边缘连续的图像的质量要优于按照客观标准（如 MSE）评价的客观平滑图像的质量。简单加权插值算法和基于块匹配的差错掩盖算法都是对于平滑区域能够取得较好的掩盖效果，而对于纹理或边缘细节丰富的图像产生模糊或者方块效应。为了对存在其它方向边缘的区域获得较好的掩盖效果，人们又提出了各种基于边缘检测的方向插值算法 DIABED (Directional Interpolation Algorithm based on Edge Detection)。

### 3.3.1 基于单一纹理方向的方向插值算法

基于单一纹理方向的方向插值算法<sup>[39]</sup>主要考虑图像中边缘的完整性，假设丢失块中只存在一个纹理方向，并预测该丢失块的纹理方向，把该纹理方向近似为一条直线，利用已知像素沿该方向对丢失块进行插值以恢复该纹理方向，如图 3-4 所示，用平行于边缘方向的斜向直线上的相邻宏块的边界像素对丢失宏块中的像素进行加权线性插值，而不像简单加权插值算法那样简单地使用水平和垂直方向的边界像素进行插值。

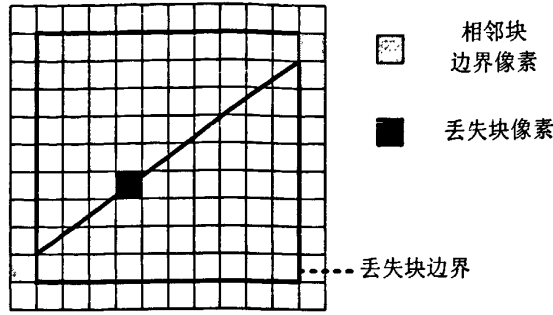


图 3-4 单一纹理的方向插值

方向插值算法的关键在于准确预测出丢失宏块中存在的边缘,传统的方法一般是采用边缘检测算子在丢失宏块的相邻宏块进行边缘检测,根据所提取的边缘信息预测出穿过丢失宏块的边缘。Sobel 算子的抗噪性能较好,模板公式为:

$$s_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad s_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (3-6)$$

对丢失块周围可用块中的每一像素点  $y_{ij}$  求取在水平和垂直方向的梯度分量值:即像素  $y_{ij}$  对应一个边缘矢量,由垂直分量和水平分量组成:

$$G_x(i, j) = \text{vec}(S_x)^T \text{vec}(F(i, j)) \quad (3-7)$$

$$G_y(i, j) = \text{vec}(S_y)^T \text{vec}(F(i, j)) \quad (3-8)$$

即:

$$G_x(i, j) = y_{i+1,j-1} - y_{i-1,j-1} + 2 * y_{i+1,j} - 2 * y_{i-1,j} + y_{i+1,j+1} - y_{i-1,j+1} \quad (3-9)$$

$$G_y(i, j) = y_{i-1,j+1} - y_{i-1,j-1} + 2 * y_{i,j+1} - 2 * y_{i,j-1} + y_{i+1,j+1} - y_{i+1,j-1} \quad (3-10)$$

其中  $F(i, j)$  是以  $y_{ij}$  为中心的可用像素点的集合:

$$F(i, j) = \begin{bmatrix} y_{i-1,j-1} & y_{i-1,j} & y_{i-1,j+1} \\ y_{i,j-1} & y_{i,j} & y_{i,j+1} \\ y_{i+1,j-1} & y_{i+1,j} & y_{i+1,j+1} \end{bmatrix} \quad (3-11)$$

每一点的梯度幅度  $|G(i, j)|$  和方向  $\theta_g(i, j)$  可由下式计算得出:

$$|G(i, j)| = \sqrt{G_x^2(i, j) + G_y^2(i, j)} \quad (3-12)$$

$$\theta_g(i, j) = \arctan\left(\frac{G_y(i, j)}{G_x(i, j)}\right) \quad (3-13)$$

得到了丢失块周围的纹理方向后,把纹理方向划归到预先设置的多个方向中的一个方向上,划归的方法为就近原则,即把纹理方向划归为距离它最近的上面的一个角。对每个

方向设置一个计数器，计数器值最大的方向为最终的掩盖方向，纹理方向角为：

$$\theta(x, y) = \arctan(G_y / G_x) \quad (3-14)$$

最后在该方向上进行线性插值恢复丢失像素。

### 3.3.2 基于多纹理方向的方向插值算法

基于单一纹理方向的方向插值算法是对整个丢失块只提供一个主方向，但一个宏块中可能存在多个边缘的交汇，所以该方法的应用受到限制<sup>[40]</sup>。

基于多纹理方向的插值算法利用了相邻像素之间边缘信息的高度相关性，对于恢复包含丰富边缘且与相邻块的边缘连续的块的恢复效果较好。丢失块的周围有一些未丢失信息的像素，该算法把周围像素划分为 $\{0^\circ, 22.5^\circ, 45^\circ, 67.5^\circ, 90^\circ, 112.5^\circ, 135^\circ, 157.5^\circ\}$  8个方向<sup>[41]</sup>，如图 3-5 所示。首先利用检测器检测周围像素并判断哪些方向的边缘信息丰富，该过程使用 Sobel 算子计算每个方向边缘梯度，若这些梯度大于某个阈值，则判断为边缘信息丰富的方向；然后在这些方向进行空间插值，沿每一个主方向各自得到一个掩盖恢复块，构成和丢失块大小相同的一组块，每个块都描述了各自方向的边缘信息，空间插值采用像素加权的方法，权值为周围像素到丢失像素的距离；最后抽取这些块的边缘细节，将其整合成一个块，这样就恢复了一个包含多个插值边缘的块。

当周围像素暗示丢失的块有多余一个边缘方向，并且周围像素含有高度相关的边缘信息时，该方法的恢复效果较好。

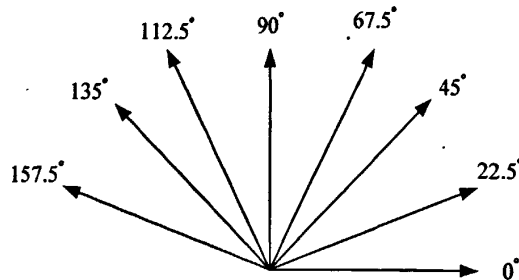


图 3-5 归并的划分方向

### 3.3.3 基于内容划分的方向插值算法

为了使用最合适的误码掩盖方法，需要根据正确接收块的特性来估计每个丢失块的内容。可以将丢失块的内容划分成三类<sup>[42]</sup>：平滑块：块内像素点的灰度值保持不变或近似不变，即没有明显的边缘在块中；边缘块：该块包含两个或更多具有不同灰度级部分的边界，通常由于块的尺寸较小，总体上图像沿边缘方向的变化不大；纹理块：不论灰度级还



是边缘方向都有明显的变化。

利用 Sobel 算子卷积计算出丢失块周围正确接收宏块的梯度幅度和以及方向,将方向也归并到图 3-5 所示的 8 个方向之一。在每一个方向上存在一个计数器  $d_k(k=0,1,\dots,7)$  来存储相应方向的梯度幅度和,每个方向存在边缘的可能性就由计数器的值来决定。

令  $d_{\max}=\max\{d_1,d_2,\dots,d_8\}$ 。对  $d_k$ , 如果满足条件:  $d_k>T*d_{\max}$ , 并且  $d_{\max}\geq T_d$ , 则认为该方向存在强边缘。这里  $T$  和  $T_d$  是两个确定的阈值,由实验数据观察得出,文献中提出的值为  $T=0.55$ ,  $T_d=3000$ 。具有强边缘的方向个数用  $N_d$  记录。根据这些计数器的值,丢失宏块的内容可以按如下条件分成三类:

情况一:  $d_{\max}<T_d$ 。在这种情况下,  $d_{\max}$  小于阈值  $T_d$ , 这说明所有 8 个方向上都没有强边缘出现,即所有像素点的灰度级保持不变或变化很小。因此可以认为丢失块是平滑块,采用简单像素加权插值。

情况二:  $d_{\max}\geq T_d$  且  $N_d\leq T_N$ 。这种情况意味着在丢失块内某些方向上存在着强边缘,但它们的个数不超过给定的阈值  $T_N$ 。所以可以将这个丢失块认为是边缘块,采用基于单一纹理方向的插值。

情况三:  $d_{\max}\geq T_d$  且  $N_d>T_N$ 。这种情况指示着有超过阈值个数的方向存在着强边缘。相对于情况二,这种情况下的丢失块被认为是纹理块,采用基于多纹理方向的插值。

### 3.3.4 基于方向插值的差错掩盖算法分析

每种误码掩盖算法往往只针对某些特定类型的图像序列,在另一些情况下效果可能就不尽如人意,很难有一种通用的方法能对所有类型的图像都获得良好的效果。基于单一纹理方向的插值算法对整个丢失块只提供一个主方向,该方法的应用范围受到限制。基于多纹理方向的插值算法考虑了多边缘情况,得到图像中存在的各个边缘方向,然后恢复各边缘,最后将各边缘融合恢复丢失块。它的问题是易受到噪声的影响,容易造成错误的边缘延伸,很难保证边缘匹配的准确。总体而言,这两种方向插值算法适用于具有明显边缘的图像,而对比较复杂的纹理区域也不能获得良好的恢复效果。而自然图像通常会同时具有几种特性,因此需要综合使用多种的误码掩盖算法,在具有不同图像特性的区域自适应的使用相应的误码掩盖算法,以获得最佳的效果,由此产生了基于内容划分的方向插值算法。

总体上说,方向插值算法虽然能够有效地恢复出图像中的物体边缘,然而它对任何图像区域总是沿特定的方向进行插值,因此容易插值出虚假的边缘,或产生不连续的边缘,造成图像失真。

### 3.4 本章小结

要在带宽有限的通信信道上传输视频，必须对视频进行压缩，而通信信道一般是有差错信道，压缩视频的数据特性加剧了信道误码所产生的视频质量的下降，这就要在视频解码时进行差错掩盖。本章回顾了几种典型的基于统计学的帧内差错掩盖方法：简单像素加权插值、基于块匹配的算法和基于方向插值的算法，分析了这些方法存在的优势和局限性，进而给出这些方法的改进措施。其中重点介绍了块匹配方法中的边缘检测和匹配准则，以及基于方向插值算法中的边缘方向划归。

## 第四章 基于模糊推理的视频帧内差错掩盖

模糊推理是模拟人的日常推理的一种近似推理，是采用模糊逻辑由给定的输入到输出的映射过程，它是由L.A.Zadeh首先提出的。在逻辑推理中，命题一般称为判断。所谓推理就是从一个或几个已知的判断前提出发推出另一个新判断结论的思维形式。当推理所用判断具有模糊性时，也就是在推理规则前提中含有模糊命题时，就称之为模糊推理。模糊推理所得的结论为一个新的模糊判断（模糊命题）。例如，如水温偏高，则加一些冷水；现水温稍高，所以加少许冷水。这种推理是基于模糊逻辑的推理，此时事实与前提的条件不完全等同，所以，模糊推理本质上是一种近似推理。

图像信息不同于语音信息或文本信息，它具有一定的模糊性，尤其是用于观赏的图像更是如此。因此将模糊推理的方法应用到图像处理中去，比较符合人眼的视觉特性。文献[43]提出了一种方法：采用基于内插的HCIE（Hierarchical Compass Interpolation/Extrapolation）方法来恢复丢失块的低频系数，再用模糊逻辑推理的方法来恢复高频系数，并用滑动窗口的滤波方法来综合调整，取得了比较好的效果。但是这种方法包含三个处理步骤，这三个处理方法的作用实质上是有所重叠的，而且实现起来其计算复杂度比较高。本章提出了一种新的较为简便的基于模糊推理的误码恢复FRBEC (Fuzzy Reasoning Based Error Concealment)方法来对采用块编码的图像进行误码恢复。

### 4.1 基于模糊推理的图像处理系统的组成

模糊推理系统一般可以由隶属度函数库、行为规则集、算法库、模糊推理和反模糊化处理五个模块组成，如图 4-1 所示。

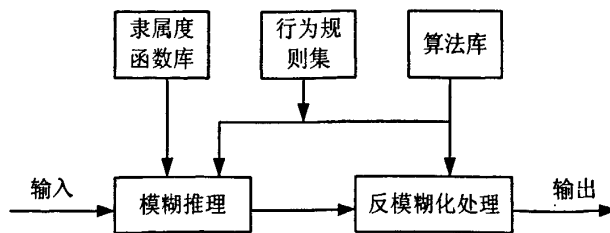


图 4-1 模糊推理系统的一般模型

在实际控制系统中，系统输入常常是有实际物理量构成的输入矢量；行为规则集是由多条“*If-Then*”形式的语句构成，代表推理规则；模糊子集隶属度函数库是对行为规则中的模糊概念的数学理解和定义，用于对输入物理量的模糊化；算法库则包含模糊推理过程中用到的模糊算子以及将模糊输出反模糊化为确定数值的算法；系统的输出常常对应于

可以改变系统状态的物理量。

一个实际的模糊推理系统的输入、输出变量可以有多个；规则的多少、规则的重叠程度、隶属度函数的形状等都是可以灵活设置的。这虽然增加了控制上的灵活性，但是也使得系统的建立和调整不容易把握，需要经过反复的修改和调整才能得到比较满意的结果。

在图像处理方面，由于图像中的局部状况的描述往往是模糊的，很难设定在何种条件下采用何种滤波器，因此一个滤波系统应该具有根据模糊和不确定信息进行推理的能力，将模糊逻辑引入图像处理不失为一种合理的方法。现有的模糊图像处理算法大都集中在模糊自适应滤波器的研究。这类方法常常是通过研究中心像素所处邻域的某些性质来确定滤波器的参数（如窗口尺度），从而达到自适应滤波的效果。然而作为一种单一的算子，要保证其在自适应参数的可调范围内都可以达到很好的处理效果是很难的，因此改进算法将 If-Then 规则引入图像处理，根据推理的规则来决定采用何种滤波器，而不是直接输出某一固定算子的参数值，这样就尽量使得图像处理算子被应用在效果较好的范围内，从而使整体的处理效果较高。这种方法的流程图如图 4-2 所示：

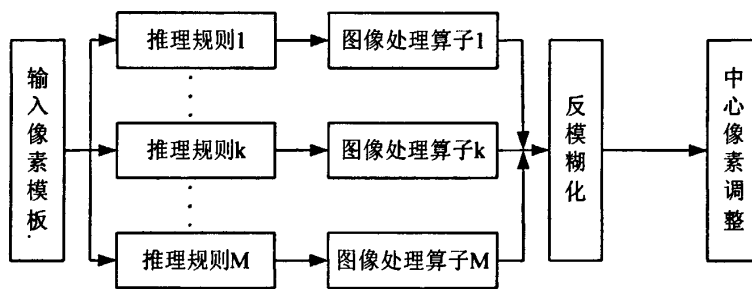


图 4-2 基于模糊推理的图像处理系统

这种模糊图像处理系统的一个特点是通过模糊逻辑的引入可以充分利用人类视觉的先验知识，从而可以使算法更接近于人脑对图像信息的处理，也就可以使得增强的效果得到较好的主观评价。

这种方法的另一个特点是具有很强的灵活性，主要表现在以下几个方面：

- 1) 对不同图像处理任务可以设计不同的模糊处理系统，如：模糊边缘检测系统、模糊平滑系统等；
- 2) 对某个给定的图像处理任务，输入变量的维数容易扩充，这种扩充对应于推理规则中多个条件的组合，最终反映为模糊集合的交并补运算；
- 3) 对某个给定的图像处理任务，每个推理规则的结论对应的图像处理算子的选择很灵活，可以对不同的图像采用不同的算子；
- 4) 隶属度函数的形状和反模糊化方法都可以有多种选择，可以根据需要进行修改。

这种方法也有一定的局限性：

- 1) 模糊系统对于不同的图像处理任务采用不同的行为规则库，所以适应性小，需要为不同的需要设计不同的规则；
- 2) 同一个模糊系统的规则需要一个调整的过程。模糊规则引入了人类的先验知识，而这种先验知识带有明显的经验性，规则的确定需要反复的修改和调试才能得到满意的结果。

## 4.2 基于模糊推理的差错掩盖

### 4.2.1 基于模糊推理的差错掩盖流程

如图 4-3 所示，基于模糊推理的差错掩盖算法分为三个步骤：首先，采用像素匹配的方法对丢失块周围一圈的像素进行匹配。然后，采用模糊推理的方法对匹配的结果进行模糊分类。最后，如果恢复准确度高，则接受结果；如果恢复准确度低，则进行像素逐点修正处理以改善掩盖效果。

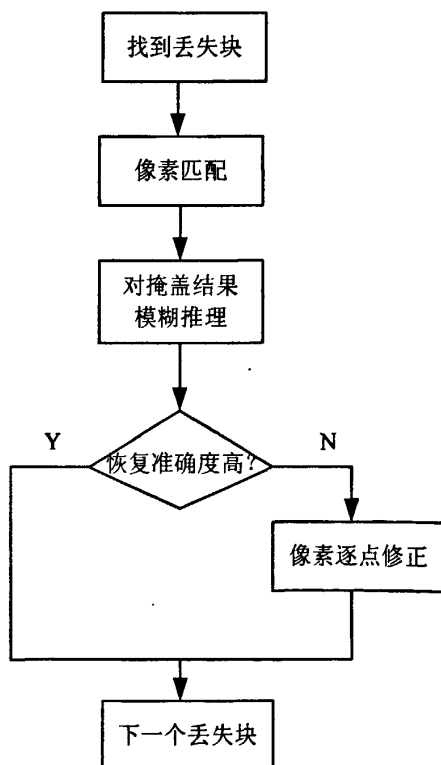


图 4-3 基于模糊推理的差错掩盖流程

#### (1) 像素匹配

考虑到人眼的视觉特性，当某帧图像中出现丢失块时，如果只给出丢失块周围 8 个块

的信息（小区域），那么，人脑在判断丢失块的信息时的准确度就比较低，但是，如果给出丢失块周围较大区域里的图像信息，那么人脑就能够充分地利用这个区域里的图像冗余信息来对图像信息进行恢复，其准确度也就越高。这就是说，在自然图像中往往存在着一定的远程相关性，在不少情况下，丢失块和图像中的一些远程图像块具有相似性。如果利用这些相似的块来恢复丢失块的话就可以获得较好的效果。在这里我们利用远程相关性来寻找图像中的相似块，最简单的像素匹配就是采用第三章介绍的基本块匹配方法。

图 3-3 显示了像素匹配的过程，以丢失块为中心确定一个  $40 \times 40$  的操作窗  $W(p,q)$ ， $p,q=1,2,\dots,40$ ，在这个范围内进行像素匹配。将丢失块及其周围一圈的像素构成一个  $10 \times 10$  的本地窗  $y(i,j)$ ， $i,j=1,2,\dots,10$ ，用这个本地窗去和操作窗口中不包含丢失块的任意一个  $10 \times 10$  的匹配窗  $y_0(i,j)$ ， $i,j=1,2,\dots,10$ ，进行匹配，计算它们周围一圈像素均方误差，并将其定义为匹配差异度。匹配差异度最小的窗就称为最佳匹配窗，最佳匹配窗中除外面一圈像素外中间的  $8 \times 8$  块就称为最佳匹配块  $x(m,n)$ ， $m,n=1,2,\dots,8$ 。

按照公式 (3-2) 计算周围一圈像素方差 MSE 并定义匹配差异度 MDD (Matching Difference Degree) 为：

$$MDD = \min\{MSE\} \quad (4-1)$$

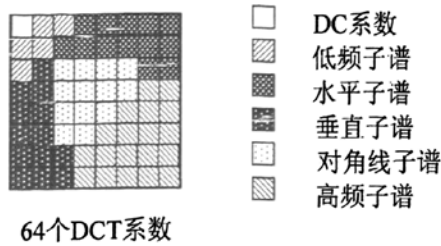
如果匹配差异度较小，则丢失块恢复的准确度就较高；如果匹配差异度较大，则丢失块恢复的准确度就较低；但是对于中间大部分的匹配度差异度，丢失块恢复的准确度很难确定，因为它还和块的纹理平滑度 TSD (Texture Smoothness Degree) 有一定的关系。因此，就需要对匹配的结果进行合理的分类，这通过模糊推理来完成。

## (2) 模糊逻辑推理

影响恢复准确度的因素有匹配差异度 MDD 和纹理的平滑度 TSD，但是这两个量在判断它们大小的时候具有一定的模糊性，比如纹理平滑度，纹理平滑大小之间并没有很明显的数字分界，更多的只是程度上的不同，比如纹理较平滑或较不平滑，即它具有一定的模糊性。因此，我们采用模糊推理的方法来对匹配的结果进行模糊分类。

这里模糊推理用到的特征量有两个，即匹配差异度 MDD 和最佳匹配块的纹理平滑度 TSD。推理的结果为恢复的准确度 RED (Recover Exactness Degree)。

对于一个  $8 \times 8$  的块，其 DCT 变换系数可以分成直流系数 (dc) 和交流系数 (ac) 两部分，其中 ac 系数又可以分成 5 个子谱，即低频子谱、水平纹理子谱、垂直纹理子谱、对角线纹理子谱和高频纹理子谱。如图 4-4 所示。

图 4-4  $8 \times 8$  块的 DCT 系数子谱分布

我们用归一化的低频能量  $E_1$  来定义该块的纹理平滑度, 即

$$TSD = E_1 = \frac{13}{5} \sum_{i \in \Psi_1} C_i^2 / \sum_{j \in \Omega} C_j^2 \quad (4-2)$$

其中  $C_i$ 、 $C_j$  是 ac 系数,  $\{\Psi_1\}$  为子块的低频子谱 ac 系数的集合。  $\Omega$  表示除了低频子谱 ac 系数外的所有 ac 系数的集合。也就是说一个块的纹理复杂度是由该块的低频子谱 ac 系数与其他纹理子谱 ac 系数的比值决定。

当 TSD 较大时表示是一个浅纹理块, 当 TSD 较小时表示是一个复杂纹理块。接下来对恢复准确度 RED 进行模糊推理。

推理规则为:

- ① 如果匹配差异度小, 则恢复准确度高;
- ② 如果匹配差异度中, 且纹理平滑度大, 则恢复准确度高;
- ③ 如果匹配差异度中, 且纹理平滑度很小, 则恢复准确度高;
- ④ 如果匹配差异度大, 则恢复准确度低;

图 4-5 是恢复准确度 RED 的模糊推理过程的示意图。

图中推理结论的反模糊化采用重心法, 也就是将 4 条规则得到的推理结果进行合成, 然后将合成后的图像采用重心法计算出重心点, 并将这一点对应的 RED 值作为输出的 RED 值。

### (3) 像素逐点修正

对于恢复准确度低 (如  $RED < 0.3$ ) 的匹配块, 恢复的图像和周围的像素的差异较大, 人眼看上去就能看出很明显的方块差异, 尤其是对于一些纹理平滑度较大的块, 这种方块差异更是明显。因此, 需要对这些块进行边界像素的平滑处理, 由于这种处理只能修正块的低频信息, 所以这种方法对于具有较复杂纹理的块并不太适合。通过像素修正后, 恢复的丢失块和周围像素之间的平滑性变好了, 减少了部分的方块差异。

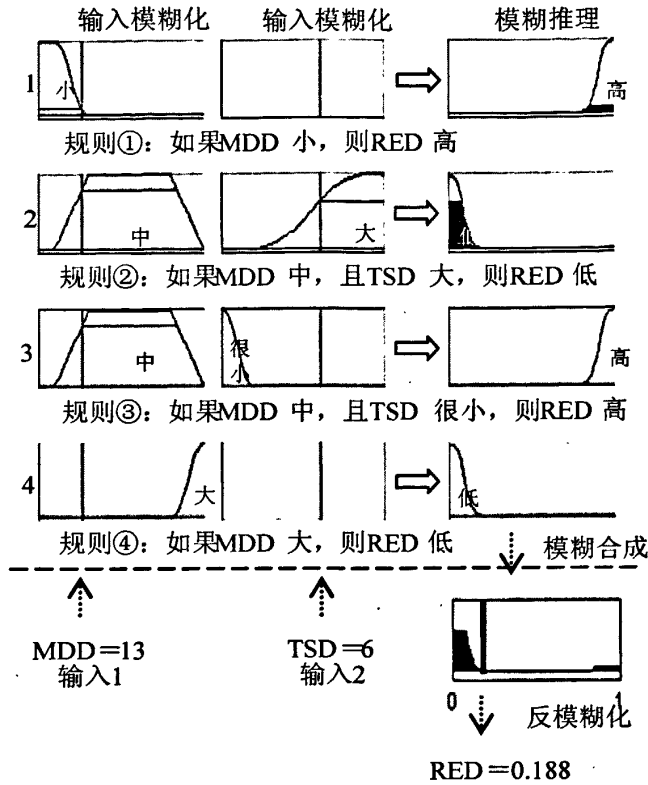


图 4-5 推理过程示例: 当 MDD=13, TSD=6, 推理得到 RED=0.188, 即认为恢复准确度低

文献[43]提出了一种像素逐点修正法, 将最佳块  $x(m, n)$ ,  $m = 1, \dots, 8; n = 1, \dots, 8$  分为 4 个子块, 分别从四个角对最佳块进行系数修正, 如图 4-6。图中的箭头方向表示计算次序。

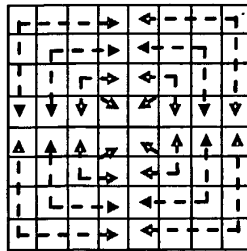


图 4-6 像素计算次序

本文在这种像素逐点修正算法上进行了改进, 即在进行像素修正的时候, 定义了一个修正参数  $C$ , 对逐点像素修正进行改进。采用修正系数  $C$  主要是为了保留最佳块中的一些匹配较好的像素, 使其不参与平滑过程 (尤其是当这些像素是复杂的纹理时, 还可以保留下部分复杂纹理)。

以左上角为例, 对于第一行和第一列的像素, 对应的  $C$  按式 (4-3) 计算, 它是最佳匹配块  $x(m, n)$  的边界像素和本地窗  $y(i, j)$  的周围一圈对应像素的差值 (如图 3-3 所示)。



$$\begin{cases} C = x(1,n) - y_0(1,n+1), n=1,2,3,4 & (\text{行}) \\ C = x(m,1) - y_0(m+1,1), m=2,3,4 & (\text{列}) \end{cases} \quad (4-3)$$

对于其他行和列的像素， $C$ 按式(4-4)计算：

$$\begin{cases} C = x(m,n) - y(m-1,n), m \leq n & (\text{行}) \\ C = x(m,n) - y(m,n-1), m > n & (\text{列}) \end{cases} \quad (4-4)$$

其中， $m, n=2,3,4$ 。

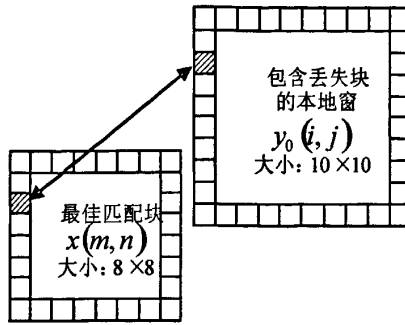


图 4-7 最佳匹配块(8×8)的边界像素与匹配窗(10×10)相应边界像素的差值  $C$

修正前，先计算像素对应的  $C$  值，并进行如下的判断：

- ① 当某像素对应的  $C > T$ ，对该像素进行修正；
- ② 当某像素对应的  $C \leq T$ ，该像素不修正，且当该像素位于行（列）次序时，对于此行（列）后面的其他像素也不进行修正。

其中， $T$  为由像素差值导致的主观感觉灰度发生跳变的可见度阈值，可由实验事先测得。对于某个灰度  $P_i$ ，改变这个灰度值，直到观察者主观感觉到灰度发生了跳变，这时对应的像素灰度差值就是在这个灰度下的可见度阈值  $T_i$ 。

其他的子块按类似的方法进行修正。通过修正，恢复块和周围像素之间的平滑性变好了。同时，也保留了部分复杂的纹理。

#### 4.2.2 模糊推理规则及隶属度函数的选取

第二章已经指出，隶属程度的思想是模糊集理论的基本思想，隶属度是主观臆造的还是客观存在是首要解决的问题，因为应用模糊集理论方法的关键在于建立符合实际的隶属度函数。我国学者汪培庄教授提出的随机集落影理论对于相当一部分模糊集理论的隶属度函数的客观实在性给出了满意的解释，基于这一理论的模糊统计方法是确定一类模糊集的隶属度的有效方法<sup>[44]</sup>。

隶属度函数的确定方法有模糊统计方法、指派方法、借用已有的“客观尺度”和二元

对比排序法。模糊统计方法是对抽样样本进行模糊统计实验，统计某个元素在各个变动的模糊子集中出现的次数作为其隶属度频率，该隶属度频率的稳定值即为该元素的隶属度。这种方法与概率统计的区别是：概率统计是“变动的点（元素）”是否落在“不动的圈（集合或子集）”内，而模糊统计是“变动的圈”是否盖住“不动的点”。指派隶属度函数的方法普遍被认为是一种主观的方法，它可以把人们的实践经验考虑进去。如模糊集定义在实数域  $R$  上，模糊集的隶属函数便称为模糊分布。模糊分布又分为偏小型模糊分布、偏大型模糊分布和中间型模糊分布。所谓指派方法，就是根据问题的性质套用现成的某些形式的模糊分布，然后根据测量数据确定分布中所含的参数，这种方法给出的隶属度函数只是近似的，需要在实践中修改。借用已有的“客观尺度”的方法常用于经济管理和社会科学中，直接借用已有的尺度（经济指标）作为模糊集的隶属度。而二元对比排序法是指对于某些模糊集合，很难直接给出隶属度，但通过两两比较，容易确定两个元素相应隶属度的大小，先排序，再用数学方法加工得到隶属度函数，实际上，这是隶属度函数（模糊分布）的一种离散表示方法。

在本文中，隶属度函数的确定结合了模糊统计和指派隶属度函数的方法，即通过实验观察的方法确定。在大量的实验数据上，利用了人眼的视觉特性进行观察识别，然后根据常规隶属度函数的分布，为其指定合理的隶属度函数，从而确定隶属度函数。例如纹理平滑度的隶属度函数得到确定，具体的做法是，首先计算出某图像块的纹理平滑度 TSD，然后对这个图像块进行主观观察，如果认为该块的纹理平滑，则认为该块 TSD 的隶属度大；如果该块的纹理很复杂，则认为该块的 TSD 隶属度很小。以此类推，通过大量的实验数据，就可以大致确定 TSD 的隶属度函数的形状，然后再组合常用的隶属度函数的分布，最终得到 TSD 的隶属度函数。

本文中的模糊推理采用了 Z 型和 S 型的隶属度函数，其定义分别如 (4-5) 和 (4-6) 所示：

$$y = zmf(x, [a, b]) = \begin{cases} 1, & x \leq a \\ 1 - 2\left(\frac{x-a}{b-a}\right)^2, & a \leq x \leq \frac{a+b}{2} \\ 2\left(\frac{b-x}{b-a}\right), & \frac{a+b}{2} \leq x \leq b \\ 0, & x \geq b \end{cases} \quad (4-5)$$

$$y = sigmf(x, [c, d]) = \frac{1}{1 + e^{-c(x-d)}} \quad (4-6)$$

其基本形状如图 4-8 所示，其中  $a$ ,  $b$ ,  $c$ ,  $d$  为待确定的参数。S 型的隶属度函数适合于表示“特别大”和“特别小”的模糊语言，Z 型的隶属度函数则适合于渐变的模糊

语言。即 S 型隶属度函数的输入量为高分辨率模糊变量，微小的差异会引起输出的强烈变化，控制灵敏度高；而 Z 型隶属度函数的输入量为低分辨率的模糊变量，控制特性就比较平缓，灵敏度低，系统比较稳定，这就意味着在输入量偏差较大的情况采用低分辨率的模糊变量，而在偏差很小或接近于 0 时采用高分辨率模糊变量。人眼对图像的感知具有一定的模糊性，尽管从理论上说数字图像中物体对象的每一个点的亮度（包括颜色）的值是确定的，但是，由于摄像设备的差异、成像环境的不同以及人眼的视觉敏锐度所限，必定会产生一些误差。但是人眼不是对这些误差是无极限的容忍的，因此在一定范围之外，人眼对这些误差变得极为敏感。由 Z 型和 S 型函数的性质可以看出，这两种函数的组合适合于人眼的特性，因此我们将这两种函数组合成输入、输出的隶属度函数。组合的方法是使用模糊逻辑算子“或”，最后的隶属度函数在低差异范围内渐变，而在超出该范围则呈指数级变化。参数值由实验确定，最后我们确定输入输出的隶属度函数组合分别为：

$$MDD: \begin{matrix} zmf[5,15] \\ sigmf[10,5] \end{matrix}; \quad TSD: \begin{matrix} zmf[0.01,1] \\ sigmf[11,9] \end{matrix}; \quad RED: \begin{matrix} zmf[0,1] \\ sigmf[0.5,0.5] \end{matrix} \quad (4-7)$$

最终得到的隶属度函数如图 4-9 所示。

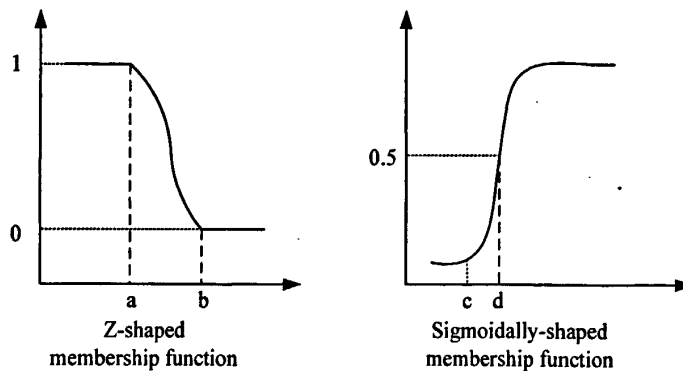


图 4-8 Z 型和 S 型隶属度函数

隶属度函数确定后，还需要选取推理规则。模糊推理系统的另一个关键性问题是生成合适的控制规则，模糊推理系统是一种基于规则的系统（rule-based system），其知识库中主要存储的是规则，这类系统接近人的判断，易接受；各种规则和运算结果一一对应，直观简单，推理速度快。推理规则的产生方法有询问调查法、操作数据法、过程模糊模型法和自组织法。询问调查法是通过询问调查总结熟练操作者的经验知识，并描述成用语言构成的控制规则。操作数据法主要是收集、整理、总结操作者所使用的数据记录，对这些数据进行相关分析，决定准备输入的变量，然后将熟练操作者的操作结果与模糊推理的输

出进行比较,调整模糊规则中的输入变量和隶属度函数,使得上述结果之差的平方和最小,从而获得一个比较合理的模糊推理规则。过程模糊模型法是一种通过生成过程的模糊模型来进行模糊控制的方法。自组织法的特点是能对控制规则进行自动学习,它通过取样评价控制性能,根据其结果修改控制规则或生成新的规则。过程模糊模型法和自组织法比较复杂,为了算法简便,我们采用了前两种方法制定推理规则。均方误差 MDD 可以反映图像客观质量的差别,而纹理平滑度 TSD 可以反映人眼对图像细节感觉的差异即主观质量的差别,因此使用这两个指标衡量恢复的准确度。4.2.1 节中的四条推理规则是根据实验的统计结果和人的主观经验总结得到的,首先判定均方误差起到主导作用,若均方误差小,则恢复的准确度大;若均方误差大,则恢复的准确度低。在均方误差中等的情况下,再根据纹理的平滑程度判断恢复的准确度,根据人眼的主观感觉,若纹理平滑度大,即图像的高频细节分量较少,则人眼容易判断出方块效应,从而判断出恢复的准确度低;否则说明图像存在许多高频细节分量,人眼易于接受掩盖效果,则恢复的准确度高。对于恢复准确度低的情况,需要修正和调整掩盖结果,本文采用的恢复准确度的阈值为 0.3。

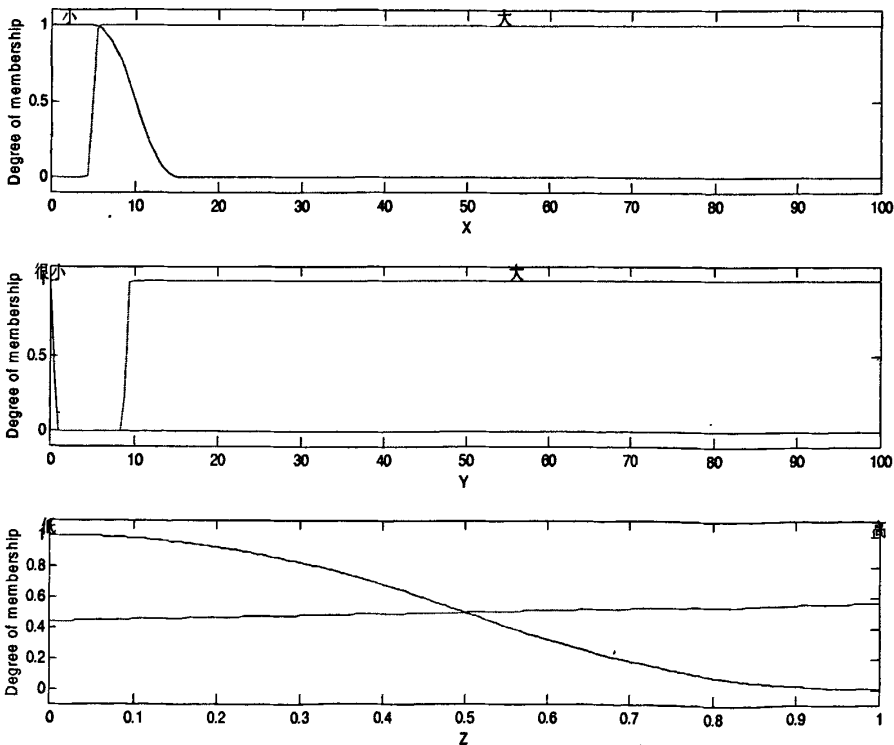
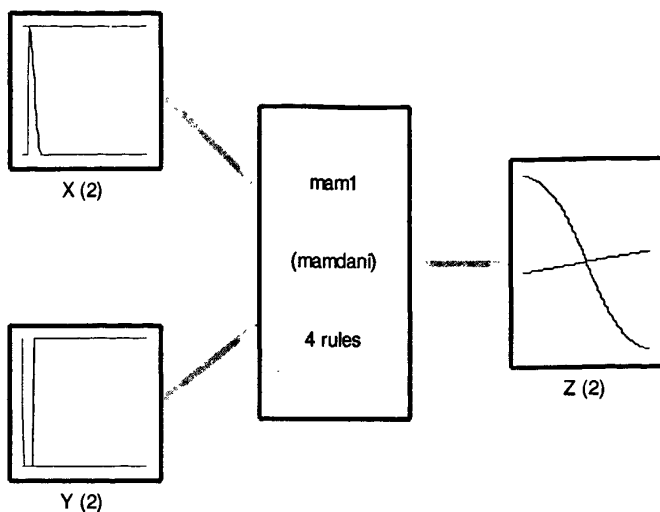


图 4-9 隶属度函数的确定

确定了隶属度函数和推理规则后,就确定了整个推理系统,最终得到的基于模糊推理的差错掩盖系统如图 4-10 所示。



System mam1: 2 inputs, 1 outputs, 4 rules

图 4-10 用于差错掩盖的模糊推理系统

### 4.3 算法的分析

本章中提出的基于模糊推理的差错掩盖方法 FRBEC 首先用基本的块匹配算法掩盖图像，然后将模糊逻辑引入对初步差错掩盖结果的评价上，最后根据评价结果决定是否对掩盖结果进行修正。因为模糊推理采用的评价指标是结合客观质量评价的均方误差 MDD 和主观质量评价的纹理平滑度 TSD，所以对于具有复杂纹理和精细的细节的图像，FRBEC 方法不仅在主观上可以获得较好的恢复效果，而且在 PSNR 也可以有较大的提高。FRBEC 方法对于不同的方块尺寸都适用，只是匹配窗和操作窗要作相应的调整。块尺寸越小，恢复的效果也越好，但相应的计算量也增加，为了折中计算量和恢复的准确度，我们采用  $8 \times 8$  的块尺寸。

对于所有的视频图像，都可以用均方误差和纹理平滑度来进行推理，即所有的视频序列可以采用 4.2.1 节中所提出的同一套推理规则，只需要调整相应的参数，而不需要针对每个视频序列设计不同的推理规则，适应范围广泛，避免了 4.1 节中的模糊系统对于不同的图像处理任务采用不同的行为规则库的局限性。

但是对于同一个模糊推理系统，仍需要一个规则参数调整的过程。推理规则引入了人类的先验知识和主观感觉，而这种先验知识和主观感觉因人而异，带有明显的经验性，参数的确定需要反复的修改和调试，这是本算法的一个局限性。

## 4.4 本章小结

由于图像信息存在模糊性，因此可以将模糊逻辑引入差错掩盖。本章采用一种新的基于模糊推理（FRBEC）方法来对误码图像进行恢复，它在像素匹配的基础上采用模糊推理的方法对匹配结果进行模糊分类，并对恢复准确度低的匹配块进行修正处理。

本章首先介绍了模糊推理系统及其在图像处理中的应用，并分析其优点和存在的局限性；然后将模糊推理引入差错掩盖，根据匹配差异度和纹理平滑度推理恢复的准确度，以确定是否需要初步的差错掩盖结果进行修正，其中详细介绍了隶属度函数的选取和推理规则的制定；最后总结该算法针对所有图像序列能够避免重复制定推理规则、适应范围广泛的优点和存在参数调整的问题。

## 第五章 基于模糊聚类的视频帧内差错掩盖算法

基于方向插值的差错掩盖算法对视频帧内比较平滑的区域或者纹理细节较少的区域有比较好的效果，但是对具有很多细节和纹理的区域则会造成图像模糊或者错误的边缘。而在基于块匹配的方法中，首先寻找与差错块相似的相邻块，然后用拷贝块的方法进行掩盖，它对运动比较平滑的块的掩盖有较好的效果，但这种方法在纹理区域有明显的块效应，同时在寻找相似块的过程中，它采用的单一相似特征量为最小绝对差（MAD）或统计量（均值和方差），这种采用单一特征量的硬判决方法找到的块可能不是最佳相似块，使得掩盖后的差错块外边界不能和周围像素平滑连接，从而使掩盖后的区域出现边缘截断现象和明显的块效应。本章提出了一种基于模糊聚类的多特征量软判决方法来较好地解决这个问题，使得差错掩盖效果有明显改进。

由本论文第 2.3 节知道，基于模糊聚类的分析方法有四种，其中基于等价关系的模糊聚类算法比较简单，且能有效地综合对象的多个特征量进行分类，所以，在本章中把它应用到相似块的多特征聚类中去，对差错块周围的块进行分类，选择和差错块同类的块进行后续的掩盖工作。

### 5.1 算法的流程

当检测到一个丢失块后，先找到该丢失块的周围相邻的 8 个同样大小的块（这几个块并不一定全部存在），对这些块提取特征矢量和聚类分析。因为基于简单像素加权的内插方法在比较平滑的区域有较好的掩盖效果，并且计算简单，所以首先判断该包含丢失块和其相邻块的区域是否为平滑区域，若为平滑区域，则直接采用简单像素加权的内插的方法恢复丢失块以简化计算量；否则进行下一步的细分。先进行垂直方向的细分，对垂直方向细分的块同样进行特征量提取和模糊聚类分析，判断这些块是否相似，若相似，则根据相似块的分布进行多方向插值以恢复丢失的块；否则进行水平方向的细分，在水平方向上进行特征量提取和分析。若水平方向也没有相似块，则说明该丢失块与周围的相邻块存在很小的相关性，只能采用基于平滑滤波器的插值方法即简单像素加权内插的方法恢复丢失块。

概括来说，本章提出的基于模糊聚类的差错掩盖算法分 3 步掩盖：首先提取待分类块的特征矢量；其次，按照选定的相似性的度量对待分类块进行模糊聚类分析，找到与丢失块相似的块；最后根据相似块的位置分布进行多方向线性内插恢复丢失块。该算法的流程如图 5-1 所示：

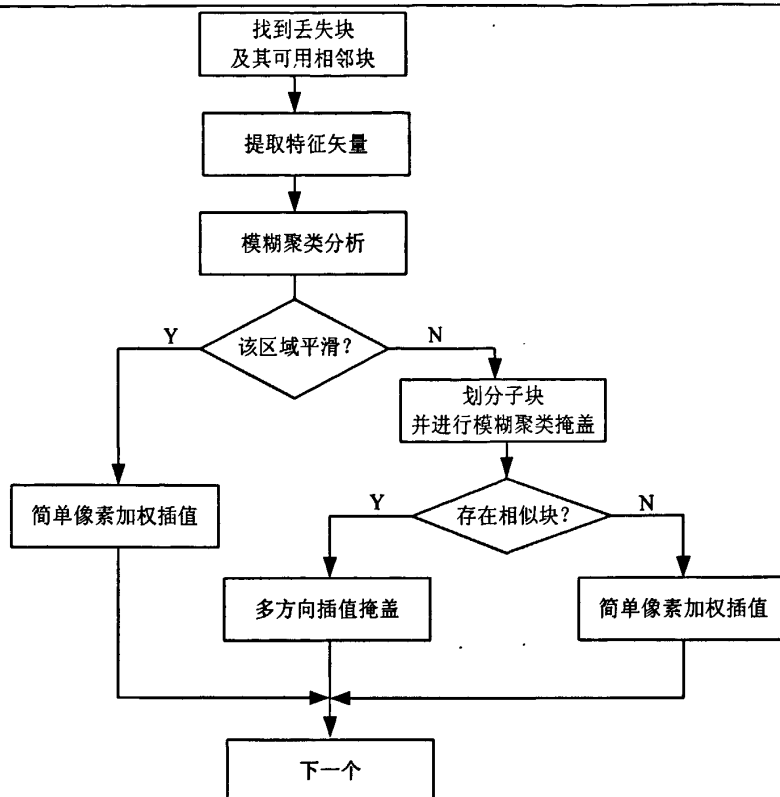


图 5-1 基于模糊聚类的帧内差错掩盖算法流程

## 5.2 算法的具体实现

### 5.2.1 聚类特征分量的提取

丢失块的掩盖必须满足三个条件：光滑性、边缘连续和一致性约束。由这三个约束条件，参考文献[45]从像素值的角度出发度量块的平滑程度，选取了匹配块的亮度、边缘梯度的均值和方差组成六维的特征矢量  $C = [m_l, \sigma_l, m_{\alpha}, \sigma_{\alpha}, m_{\beta}, \sigma_{\beta}]$ 。其中  $m_l$  和  $\sigma_l$  是匹配块的亮度均值和方差， $m_{\alpha}$ 、 $m_{\beta}$  和  $\sigma_{\alpha}$ 、 $\sigma_{\beta}$  是使用 Sobel 算子检测的  $x$  和  $y$  方向的边缘梯度的均值和方差<sup>[46]</sup>。像素的均值  $m_l$  反映了一个块的平均亮度，方差  $\sigma_l$  反映了一个块的对比度，边缘梯度的均值和方差反映了边缘的强度和方向及边缘强度的分布。这种方法为每个块引入多个匹配特征分量，能够一次寻找到多个相似块，减少了多次匹配的计算量，并提高了匹配的准确度。但是这种方法单纯从像素值出发提取特征分量，提取的特征矢量具有片面性，不能充分描述块的纹理复杂度；同时，该方法为每个特征分量分配了相同的权值，没有体现出重要的聚类分量。



为了充分度量块的纹理复杂度, 本文结合像素值和 DCT 系数, 在像素值特征分量的基础上, 从能量和块的预测模式的角度出发再增加 2 个特征分量  $e$  和  $p$ , 为每个块分配一个包含 8 个分量的特征矢量:  $C=[m_l, \sigma_l, m_{ex}, \sigma_{ex}, m_{ey}, \sigma_{ey}, e, p]$ 。其中,  $e$  为归一化化的低频能量  $E_l$ , 其定义和计算方法在本文的 4.2.1 节中详细介绍。如果  $e$  大于 1, 则说明该块低频能量大于高频能量, 是一个低纹理块; 否则说明该块含有许多高频分量, 细节较多, 是一个复杂的纹理块。

$p$  为帧内预测模式, 隐含着块的平滑程度的信息。帧内预测技术作为提高 I 帧压缩性能的一项核心技术, 在 ITU-T 的 H.263<sup>+</sup> 建议中首次被引入<sup>[47]</sup>。其原理是利用图像内部子块与子块之间的冗余性提高帧内编码效率, 即使用当前子块上方和左方的相邻块进行预测, 得到当前块的预测值, 编码时只对像素值和预测值的残差进行编码, 由于残差在量值上通常较小, 因而熵编码后占用的码字也较少, 能有效地提高帧内编码的效率。帧内预测技术分为空域帧内预测和频域帧内预测, H.264 采用的空域帧内预测技术, 采用像素值进行预测; MPEG-4 和 H.263 采用频域帧内预测技术, 采用 DCT 量化系数进行预测。引进帧内预测技术主要是为了提高压缩效率, 但同时也从另一方面反映了编码块的纹理复杂程度。如在 H.263 中高级帧内预测模式存在三个选项: DC 预测, 垂直 DC 和 AC 系数预测, 水平 DC 和 AC 系数预测。若仅用 DC 系数预测, 隐含着该块为平滑块, 高频细节少; 若用 DC 和 AC 系数联合预测, 则说明该块存在复杂的纹理边缘。同样, 在 H.264 中也存在帧内预测模式, 它的帧内预测模式分为  $16 \times 16$  和  $4 \times 4$  共 13 种模式进行, 若预测模式  $16 \times 16$ , 则说明该块为平滑块, 因此可以将该信息作为块的纹理复杂度的一个度量。

$$p = \begin{cases} 1 & \text{(若块为平滑块)} \\ 0 & \text{(若块为纹理块)} \end{cases} \quad (5-1)$$

这样为每个块分配一个包含 8 个分量的特征矢量:  $C=[m_l, \sigma_l, m_{ex}, \sigma_{ex}, m_{ey}, \sigma_{ey}, e, p]$ , 在下一节中, 本文对每个分量归一化后赋以权值  $w_i (i=1, 2, \dots, 8)$ , 其中,  $e$  和  $p$  比其他 6 个分量能更为直接的反应块的纹理复杂度, 将为其分配较大的权值。

## 5.2.2 相似块的模糊聚类

相似块的模糊聚类使用了基于等价关系的聚类方法, 为了建立分类对象集合  $X$  上的模糊等价关系  $\tilde{R}_c$ , 通常需要首先计算各个分类对象之间的相似性统计量, 建立分类对象集

合  $X$  上的模糊相似关系  $\tilde{R}_s = [r_{ij}]_{n \times n}$ ,  $0 \leq r_{ij} \leq 1, i, j = 0, 1, \dots, n-1$ ,  $r_{ij}$  表示分类对象  $x_i$  与  $x_j$  相似的程度。常用的计算各个分类对象的相似性统计量的方法有贴近度法、夹角余弦法、数量积法、相关系数法、指数法等。

假设  $n$  个待分类的块组成一个矢量  $X = \{x_1, x_2, \dots, x_n\}$ , 其中  $x_k (k=1, 2, \dots, n)$  表示待分类的块, 每个待分类的块具有 8 个特征分量  $[m_l, \sigma_l, m_{ex}, \sigma_{ex}, m_{ey}, \sigma_{ey}, e, p]$ , 分别表示为  $[x_{i1}, x_{i2}, x_{i3}, x_{i4}, x_{i5}, x_{i6}, x_{i7}, x_{i8}]$ 。分类采用基于等价类的模糊聚类方法, 将  $X$  分为两两互不相交且覆盖整个集合  $X$  的  $c$  个子类, 划分依据等价关系  $R$  (即待分类子块间的等价关系), 这个过程分 3 个步骤进行:

### 步骤 1: 归一化和权值分配

1) 归一化: 在实际应用中, 由于所获得的分类对象数据不一定是  $[0, 1]$  区间中的数, 因此首先把各个原始数据按照公式 (5-2) 标准化, 得到  $x_i$  块所对应的标准化特征矢量:

$$[x'_{i1}, x'_{i2}, \dots, x'_{i8}], i=1, 2, \dots, n$$

$$x'_{ik} = \frac{x_{ik} - \bar{x}_k}{S_k}, i=1, 2, \dots, n; k=1, 2, \dots, 8$$

$$\text{其中, } \bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_{ik} \quad (5-2)$$

$$S_k = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_{ik} - \bar{x}_k)^2}$$

2) 权值分配: 对每个特征分量分配一个权值  $w_i (i=1, 2, \dots, 8)$  来表示其对块平滑程度的贡献。本文采用固定的权值分配, 其中  $e$  和  $p$  比其他 6 个分量能更为直接的反映块的纹理复杂度, 分配较大的权值:

$$w_i = \begin{cases} 1/12, i=1, \dots, 6 \\ 1/4, i=7, 8 \end{cases}, \sum_{i=1}^8 w_i = 1 \quad (5-3)$$

从而得到加权的特征矢量  $C = [w_1 x'_{i1}, w_2 x'_{i2}, w_3 x'_{i3}, w_4 x'_{i4}, w_5 x'_{i5}, w_6 x'_{i6}, w_7 x'_{i7}, w_8 x'_{i8}]$ 。

### 步骤 2: 相似性度量

得到每个块的特征矢量后, 求出待分类块之间的相似程度。记  $r_{ij} = R(x_i, x_j)$  为待分类块  $x_i$  和  $x_j$  之间的相似程度 ( $0 \leq r_{ij} \leq 1 (i, j=1, 2, \dots, n)$ ), 从而确定了论域  $X$  上的一个相似关系, 其对应的矩阵为  $R$ 。

分类效果的好坏关键在于合适地选择相似性的统计指标, 本文选取贴进度法作为相似程度的统计指标以比较掩盖的效果。设特征分量  $x_{ik}, x_{jk} \in [0, 1] (k=1, 2, \dots, m)$ , 则  $x_i, x_j$  可看作模糊向量:  $x_i = (x_{i1}, x_{i2}, \dots, x_{im}), x_j = (x_{j1}, x_{j2}, \dots, x_{jm})$ , 分别选择  $x_i, x_j$  之间相似程度  $r_{ij}$  为:

$$r_{ij} = \begin{cases} 1, & i = j \\ \bigvee_{k=1}^m (x_{ik} \wedge x_{jk}) \wedge \left[ 1 - \bigwedge_{k=1}^m (x_{ik} \vee x_{jk}) \right], & i \neq j \end{cases} \quad (5-4)$$

符号  $\wedge, \vee$  分别是取最小值和取最大值。其中,  $\bigvee_{k=1}^m (x_{ik} \wedge x_{jk})$  表示样本  $x_i, x_j$  的内积,  $\bigwedge_{k=1}^m (x_{ik} \vee x_{jk})$  表示样本  $x_i, x_j$  的外积。内积越大, 这两个特征矢量越贴近; 外积越小, 这两个特征矢量也越贴近。但单独用内积或外积都不能刻画贴进度, 所以要用二者相结合的贴进度来刻画这两个特征矢量的相似程度。若  $r_{ij} = 0$ , 表示块  $x_i$  与块  $x_j$  截然不同, 毫无相似之处; 若  $r_{ij} = 1$ , 表示块  $x_i$  与块  $x_j$  完全相似(或等同)。当  $i = j$  时, 则是  $x_i$  自己与自己相似的程度, 恒取 1, 从而确定了论域  $X$  上的一个相似关系, 其对应的矩阵为:

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \cdots & r_{nn} \end{bmatrix} \quad (5-5)$$

其中  $r_{ij} = R(x_i, x_j)$  为被分类块之间的相似程度的统计量。

### 步骤 3: 聚类

经过上述步骤得到的模糊关系矩阵  $R$  是满足自反性和对称性的, 但不满足传递性, 故  $R$  只是相似矩阵, 还要把它改造成等价矩阵。构造等价矩阵的方法是自乘, 通过自乘依次求出  $R^8, R^{16}, \dots$ , 直至出现如下关系:  $R^{2^t} = R^{2^t} \bullet R^{2^t}$ , (其中的相乘为模糊乘积), 则得到等价矩阵  $R^{2^t}$ 。当确定了模糊等价矩阵后, 对于任意的  $\lambda \in [0, 1]$ , 所得的  $\lambda$  截矩阵  $R_\lambda$  刻画了普通集上的一个等价关系, 从而确定了  $\lambda$  水平的分类。

### 5.2.3 差错块的插值恢复

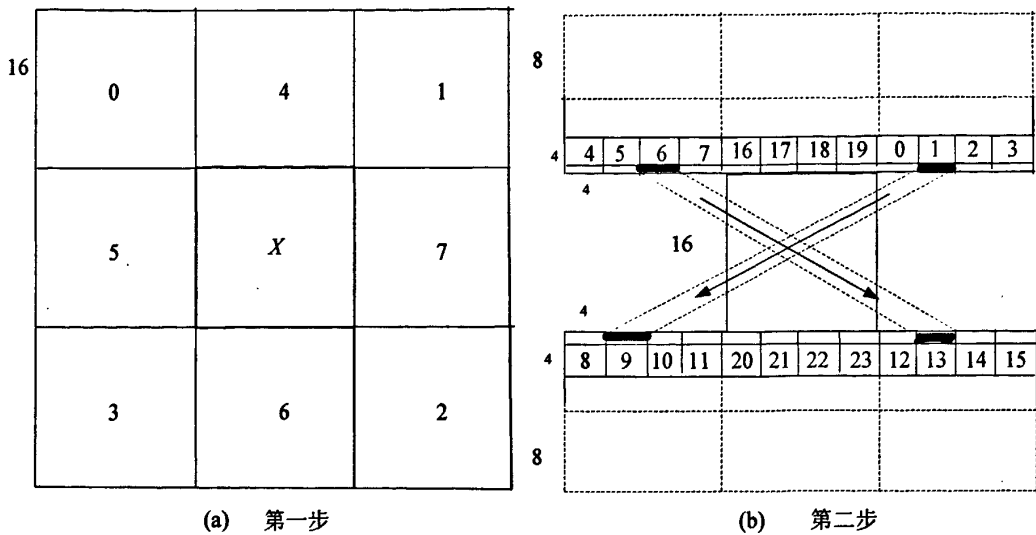


图 5-2 基于模糊聚类的多方向插值

具体的帧内差错掩盖过程可分为 4 步进行：

1) 首先对图 5-2(a)中 (0, 1, 2, 3, 4, 5, 6, 7) 8 个 16×16 块进行聚类分析。每个块的初始特征矢量为  $[m_l, \sigma_l, m_{ex}, \sigma_{ex}, m_{ey}, \sigma_{ey}, e, p]$ ，其中的边缘梯度使用 Sobel 算子进行计算，继而得到梯度的均值和方差  $m_{ex}, \sigma_{ex}, m_{ey}, \sigma_{ey}$ ；然后进行归一化和权值分配，得到最终的聚类特征矢量  $C = [w_1x'_{i1}, w_2x'_{i2}, w_3x'_{i3}, w_4x'_{i4}, w_5x'_{i5}, w_6x'_{i7}, w_7x'_{i7}, w_8x'_{i8}]$ 。根据公式 (5-4) 的贴近度相似性度量，计算每两个块的相似程度，得到相似关系矩阵  $R$ 。设定  $\lambda$ ，当矩阵中的元素  $r_{ij} \geq \lambda$  时，则表示块  $x_i$  和块  $x_j$  相似。鉴于不是所以相邻块可得，设定阈值  $RATIO$ ，并设相似的相邻块数目为  $n_1$  和所有可得的块数目为  $n_2$ ，当  $n_1/n_2 \geq RATIO$  时，表明该区域为平滑区域，使用像素值的加权平均进行恢复，否则进入第二步。

2) 第二步如图 5-2(b)所示，对垂直方向进行划分，对图中的 (0, 1, 2, 3, 4, 5, ..., 23) 24 个 4×4 的块进行与第一步类似的模糊分类。分析的结果存在以下两种情形：( I ) 若  $n_1/n_2 < RATIO$  时，说明底部的 4×4 块与上部的 4×4 块没有相似的块，则判断为丢失块的内部有水平方向的边缘细节，则进行水平方向的划分及聚类分析；( II ) 当  $n_1/n_2 \geq RATIO$  时，并且相似块分别分布在底部和顶部，则进行如图 2(b)所示的像素点多方向线性内插插值。插值公式为

$$f_x^i = f_m \cdot \frac{d2}{d1+d2} + f_n \cdot \frac{d1}{d1+d2} \tag{5-6}$$

其中： $m \in \{0, 1, 2, 3, 4, 5, 6, 7, 16, 17, 18, 19\}$ ;

$n \in \{8, 9, 10, 11, 12, 13, 14, 15, 20, 21, 22, 23\}$ ;

$d_1, d_2$  分别表示待插值的像素点到顶部和底部的距离。

3) 在第二步的线性内插时, 存在重叠的区域, 因此要把重叠区域的内插出来的像素值进行求均值合并。设置一个二维数组来存放掩盖标志, 数组的下标对应像素点的坐标, 元素值标志该像素点有无被掩盖过, 0 表示没有被掩盖, 1 表示已经被掩盖。扫描存放标志的二维数组, 若有元素值为 0, 则表示该像素没有被掩盖到, 采用像素值的加权平均恢复。

4) 在 2) 的 (I) 中, 若水平方向分类的结果也没有相似块, 则仍然采用像素值的加权平均的方法进行插值。

### 5.3 算法的分析

本章提出的基于模糊聚类的差错掩盖算法改进了相似块的匹配过程, 采用基于多特征量的软判决方法简化寻找相似块的过程, 最后根据相似块的分布采用逐像素插值而不是块拷贝的方法恢复丢失块, 提高恢复的准确度。该算法的关键在于阈值的设置和计算复杂度。

由图 5-1 掩盖算法流程可知, 该算法需要两次判断, 一是需要设定阈值 *RATIO* 判断待掩盖区域是否为平滑域; 另外一个选定阈值  $\lambda$ , 判断待匹配的两个块是否为相似块。*RATIO* 由实验统计观察得出, 一般相似块在一半以上, 该区域就被认为是平滑区域, 因此设定 *RATIO* 为 0.5。而在模糊聚类分析中, 对于各个不同的  $\lambda \in [0, 1]$ , 可以得到不同的分类, 从而形成一种动态聚类图, 这对全面了解样本的分类情况是比较直观和形象的。但在许多实际情况中需要选定某个最佳阈值  $\lambda$ , 确定样本的一个具体分类, 本章算法就属于该类。 $\lambda$  的确定有两种方法<sup>[44]</sup>: 一种方法是按照经验, 根据实际需要在动态聚类图中, 调整  $\lambda$  的值以得到适当的分类, 而不需要事先准确地估计样本应分成几类; 另一种方法是用 *F* 统计量的方法确定  $\lambda$  最佳值。*F* 统计量的方法首先计算各个样本的特征分量的均值分别作为各个样本的中心向量; 然后用这些中心向量计算各分类中心的距离的统计量 *F*, *F* 越大, 类与类之间的差异越大, 分类越好; 确定一个合适的 *F*, 即可确定分类数, 从而确定  $\lambda$  值。基于 *F* 统计量的方法需要迭代计算, 复杂度较高, 因此我们采用第一种方法确定  $\lambda$  值。

最后, 我们从计算复杂度上分析该算法。该算法用到了求梯度的均值、方差和聚类运算, 假设图像块的大小为  $m \times m$ , 参与聚类的块的个数为  $n$ 。求均值和方差都是逐像素获

得, 计算复杂度的量级为  $O(m^2)$ ; 聚类的过程中求相似性的计算只涉及简单的比较大小和求和运算, 计算量可以忽略, 但是从相似性矩阵转化为等价矩阵涉及矩阵的合成运算, 其计算复杂度的量级为  $O(n^2)$ ; 获得相似块后, 逐像素插值掩盖, 复杂度与块的大小有关, 复杂度的量级为  $O(m^2)$ ; 因此总体计算复杂度为  $O(2m^2 + n^2)$ 。简单像素加权插值算法是逐像素恢复丢失块, 其复杂度为  $O(m^2)$ ; 基于纹理方向的插值算法需要计算纹理方向, 也需要计算各像素的梯度, 计算梯度所需要的时间是  $m^2$ , 得到纹理方向后也是基于逐像素插值恢复丢失块, 总体的复杂度为  $O(2m^2)$ ; 基于块匹配的算法需要首先寻找匹配块, 设操作窗的大小为  $M \times M$ , 则寻找匹配窗在最坏的情况下需要的时间是  $M^2(m+1)^2$ , 总体的计算复杂度为  $O(m^2 + M^2(m+1)^2)$ 。从以上分析可以看出, 基于模糊聚类的差错掩盖算法的计算复杂度介于像素插值算法和块匹配算法之间, 参与聚类的块尺寸越大, 块数越少, 计算量越小; 在大尺寸的块中找不到相似块的情况下, 需要细分子块, 但并不是意味着块的尺寸可以无限制划分。在块的尺寸过小的情况下, 参与计算的像素值越少, 块之间的差异越小, 容易把不相似块误判为相似块, 因此我们划分子块尺寸的下界是  $4 \times 4$ 。

## 5.4 本章小结

模糊聚类利用每个样本属于各个类别的不确定性程度, 建立起样本属于各个类别的不确定性的程度的描述, 能够较客观地反应现实世界。而根据人的视觉特性, 图像信息存在一定的模糊性, 因此可以将模糊聚类的理论用于图像的后处理中, 如差错视频的误码掩盖。本章提出了一种基于模糊聚类方法的帧内差错掩盖新方法, 提取更多的匹配特征分量来表现块的纹理复杂度, 并根据各个分量对纹理复杂度的贡献赋以不同的权值, 选择合理的度量寻找相似块。

本章首先介绍了基于模糊聚类的差错掩盖算法的基本思想; 然后介绍算法的具体实现, 包括聚类特征分量的提取、聚类分析和差错块的插值恢复过程; 最后重点讨论该算法的关键问题, 包括最佳阈值的选取和计算复杂度分析, 指出该算法的最佳阈值是根据实验确定的经验值、算法的计算量介于像素插值算法和块匹配算法之间。

## 第六章 实验结果与分析

### 6.1 实验平台介绍

本文采用 H.264 标准草案的测试参考模型 JM8.6 作为测试平台, 并使用 PC 机实现上述各章节所述的基于经典数学和基于模糊数学的帧内误码掩盖算法, 测试条件为: CPU, Pentium 1.6GHz; RAM, 512M; 操作系统, Windows XP。虽然各类算法基于 JM8.6 实现, 但是并不局限于 H.264 标准, 这些算法都可以适应于现存的各种视频编解码标准。

测试序列为 YUV4:2:0 格式的 CIF(352×288)图像, 编码档次为 baseline。在编码端需设置: OutFileMode=1, 即使用 RTP 打包; 并且设置 IntraPeriod=1, 即全部使用帧内编码方式, 图像序列中只存在 I 帧。同时还可以通过修改 FramesToBeEncoded 的值改变待编码的帧数等。在解码端需设置: NAL mode=1, 采用 RTP 解包; H.264 编码比特流和输出文件是 YUV 格式。

产生丢包误码环境的方法采用 ITU-T 的 VCEG-N80 信道仿真软件 MobileIP, 它具有 18681.3, 18681.4, wcdma\_64kb\_3kph\_5e-04.bin, wcdma\_64kb\_50kph\_2e-04.bin, cdma\_128kb\_3kph\_5e-04.bin 和 cdma\_128kb\_50kph\_2e-04.bin 等 5 种误码模式文件, 通过模拟无线移动信道的特性产生丢包误码, 但这种方法难以控制误码率和错误产生的位置。因此误码的检测和定位采用 JM8.6 中的方法。在 H.264 标准草案测试模型 JM8.6 中, 所有出错或不完整的 slice 在解码前被丢弃, 不被解码, 因此不执行完整性检测和比特错误检测。JM8.6 首先对所有正确接收的 slice 进行解码, 然后应用误码掩盖技术对丢失片进行掩盖。在解码端首先设置一个宏块状态图, 初始状态设置为错误, 利用宏块结构中的 ei\_flag 标记和 ercStartSegment 和 ercStopSegment 函数来标识宏块正确接收与否, 若解码端收到一个 slice, 则将其中接收到的宏块标记为正确, 没有接收到的宏块仍然标记为错误状态。在掩盖过程中, 宏块有四种状态: ERC\_BLOCK\_OK, 表示正确接收的宏块; ERC\_BLOCK\_CONCEALED 表示错误接收但已被掩盖过的宏块; ERC\_BLOCK\_CORRUPTED 和 ERC\_BLOCK\_EMPTY 表示错误接收或已经丢失的宏块。在帧内差错掩盖函数 ercConcealIntraFrame 中, 查询 ercVariables\_t->concealment, 该变量指示该帧图像是否有丢失的块, 即该帧的状态图上是否有标记为“丢失的”宏块。ercVariables\_t->nOfCorruptedSegments 指示该帧中待掩盖的宏块数目。若有要掩盖的宏块, 则从该帧的第一个像素开始逐行扫描, 每个像素点的状态存储在一个一维指针数组 condition 中, 扫描的过程即为检查像素点的状态的过程, 若该状态值低于阈值, 则说明

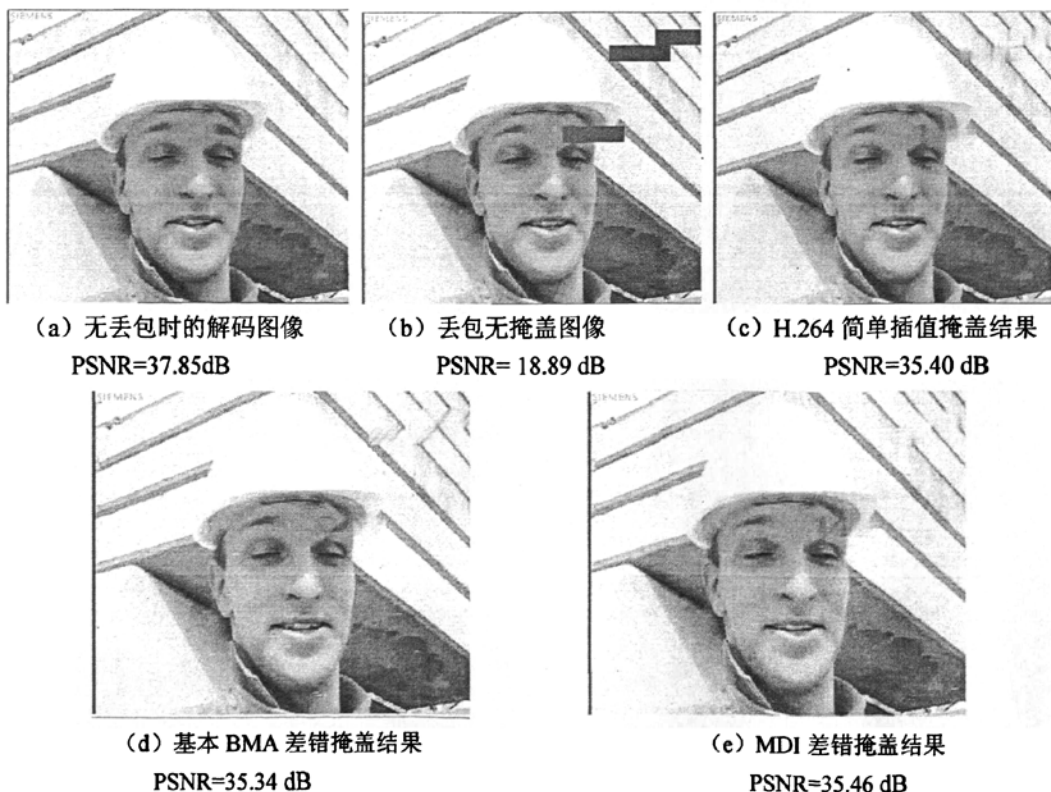
找到了一个丢失的宏块的第一个像素点。然后找到这个丢失宏块的相邻块，进行相应的差错掩盖过程。

## 6.2 实验结果分析

帧内差错掩盖的效果与图像本身的特性，如纹理区域的大小、边缘的强烈等有关，这不同于帧间方法，帧间方法通常与运动的剧烈程度等有关。因此本文选择具有不同特性的序列进行实验。对边缘较明显，纹理区域不多的情况，选取 foreman 和 highway 这两个序列进行测试，而对纹理较多的区域，则选取 mobile 和 flower 两个序列进行测试。每种序列分别测试了不同丢包率的情况。本章采用了各类基于经典数学的算法作为本文算法的比较对象。

### 6.2.1 不同算法下平滑序列的测试结果

图 6-1 是视频序列“foreman”不同掩盖算法结果的比较。(a) 是无丢包时的解码图像，(b) 是丢包无掩盖的图像，(c) 是 H.264 参考算法即基于简单像素加权内插的掩盖结果，(d) 是基于基本块匹配 BMA 算法的掩盖结果，(e) 是基于多纹理方向插值 MDI 算法的掩盖结果，图 (f) 和 (g) 分别是基于模糊推理和模糊聚类的差错掩盖结果图像。







(f) 基于模糊推理的差错掩盖结果  
PSNR=35.44 dB



(g) 基于模糊聚类的差错掩盖结果  
PSNR=35.93 dB

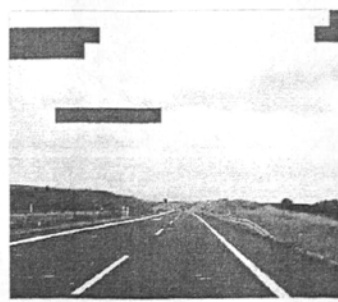
图 6-1 foreman 在丢包率为 3%时的解码图像

从图中可以看出，基于简单像素加权插值的方法在平滑的墙壁区域有较好的掩盖效果，但是对于墙壁纹理和帽子等边缘区域造成模糊现象，而且像素插值不准确，造成颜色交错混杂，可以从帽子和脸部边缘看出来；基于基本块匹配的方法的掩盖图像中出现了明显的虚假边缘，这是因为基于单一的匹配特征量（最小方差 MSE）寻找到的相似块并不是最佳的，因此拷贝过来的像素值虽然使得 MSE 最小，但是产生了错误的纹理方向；基于多纹理方向插值的算法则在墙壁区域产生了良好的掩盖效果，因为墙壁的纹理方向比较单一，能够根据梯度方向计算出丢失块的纹理方向，但是在帽子边缘和人脸细节处的掩盖效果不好；基于模糊推理的方法优于 264 和块匹配的方法，但是信噪比低于基于多纹理方向插值的方法，这是因为模糊推理在墙壁边缘区域产生了错误的结论；基于模糊聚类的掩盖方法的主观质量和客观质量最好，人脸和帽子边缘的细节也恢复出来。

图 6-2 是视频序列“highway”在丢包率为 10%时的掩盖图像。(a) 是无丢包时的解码原图像，(b) 是丢包无掩盖的图像，(c) 是 H.264 参考算法即基于简单像素加权内插的掩盖结果，(d) 是基于基本块匹配 BMA 算法的掩盖结果，(e) 是基于多纹理方向插值 MDI 算法的掩盖结果，图 (f) 和 (g) 分别是基于模糊推理和模糊聚类的差错掩盖结果图像。



(a) 无丢包时的解码图像  
PSNR=39.51 dB



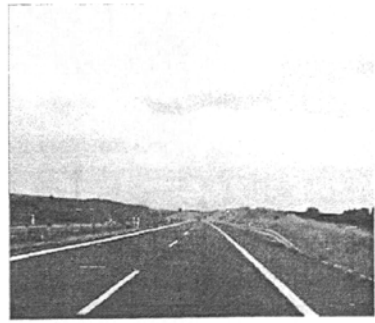
(b) 丢包无掩盖图像  
PSNR=13.53 dB



(c) H.264 简单插值掩盖结果  
PSNR= 37.55 dB



(d) 基本 BMA 差错掩盖结果  
PSNR=36.08 dB



(e) MDI 差错掩盖结果  
PSNR=36.93 dB



(f) 基于模糊推理的差错掩盖结果  
PSNR=37.64 dB



(g) 基于模糊聚类的差错掩盖结果  
PSNR=37.75 dB

图 6-2 highway 在丢包率为 10%时的解码图像

从图中可以看出，丢包的位置处于比较平滑的位置，各种算法的主观掩盖质量相当。在左上角的云朵被完全丢掉以后，各种算法都无法恢复。但是基于模糊数学的掩盖方法在下方的云朵的纹理细节处获得了更好的掩盖效果，灰度插值更为准确，因此客观质量更高。

### 6.2.2 不同算法下复杂序列的测试结果

图 6-3 是视频序列“mobile”在丢包率为 9.5%时的掩盖图像。(a) 是无丢包时的解码图像，(b) 是丢包无掩盖的图像，(c) 是 H.264 参考算法即基于简单像素加权内插的掩盖结果，(d) 是基于基本块匹配 BMA 算法的掩盖结果，(e) 是基于多纹理方向插值 MDI 算法的掩盖结果，图 (f) 和 (g) 分别是基于模糊推理和模糊聚类的差错掩盖结果图像。



(a) 无丢包时的解码图像



(b) 丢包无掩盖图像



(c) H.264 简单插值掩盖结果

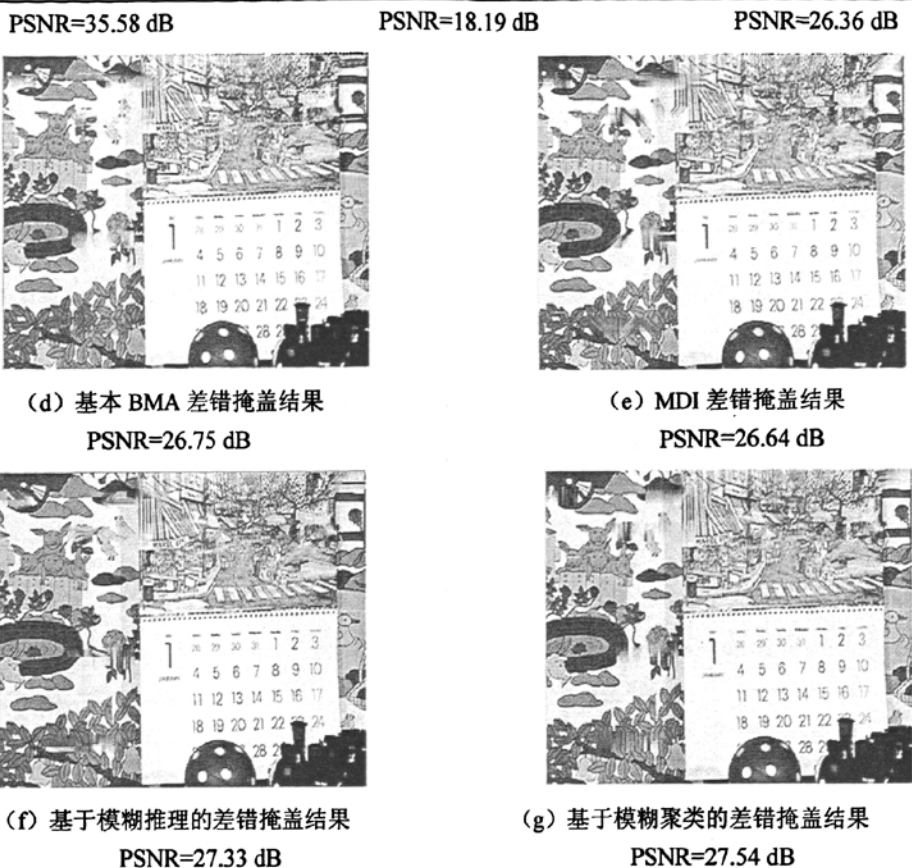


图 6-3 mobile 在丢包率为 9.5% 时的解码图像

从图中可以看出，基于简单像素加权插值的方法在细节和纹理区域造成模糊现象，而且像素插值不准确，造成颜色交错混杂；基于基本块匹配的方法对于细节丰富的区域用块拷贝的方法取得了较好的效果，如叶子区域，但由于寻找的相似块未必最准确，出现了明显的虚假边缘，如车轮和羊头附近，拷贝过来的像素值使得部分细节消失或产生错误的细节；基于多纹理方向插值的算法在纹理走向不多和边缘明显的区域能够产生较好的效果，如羊头附近的圆圈，但是在细节丰富的区域则产生很多错误的纹理方向分割，影响了主观掩盖质量；基于模糊推理的方法和基于模糊聚类的掩盖方法在平滑区域和细节丰富的区域都能够产生良好的掩盖效果，信噪比也高于基于经典数学的掩盖方法。

图 6-4 是视频序列“flower”在丢包率为 15% 时的掩盖图像。(a) 是无丢包时的解码图像，(b) 是丢包无掩盖的图像，(c) 是 H.264 参考算法即基于简单像素加权内插的掩盖结果，(d) 是基于基本块匹配 BMA 算法的掩盖结果，(e) 是基于多纹理方向插值 MDI 算法的掩盖结果，图 (f) 和 (g) 分别是基于模糊推理和模糊聚类的差错掩盖结果图像。

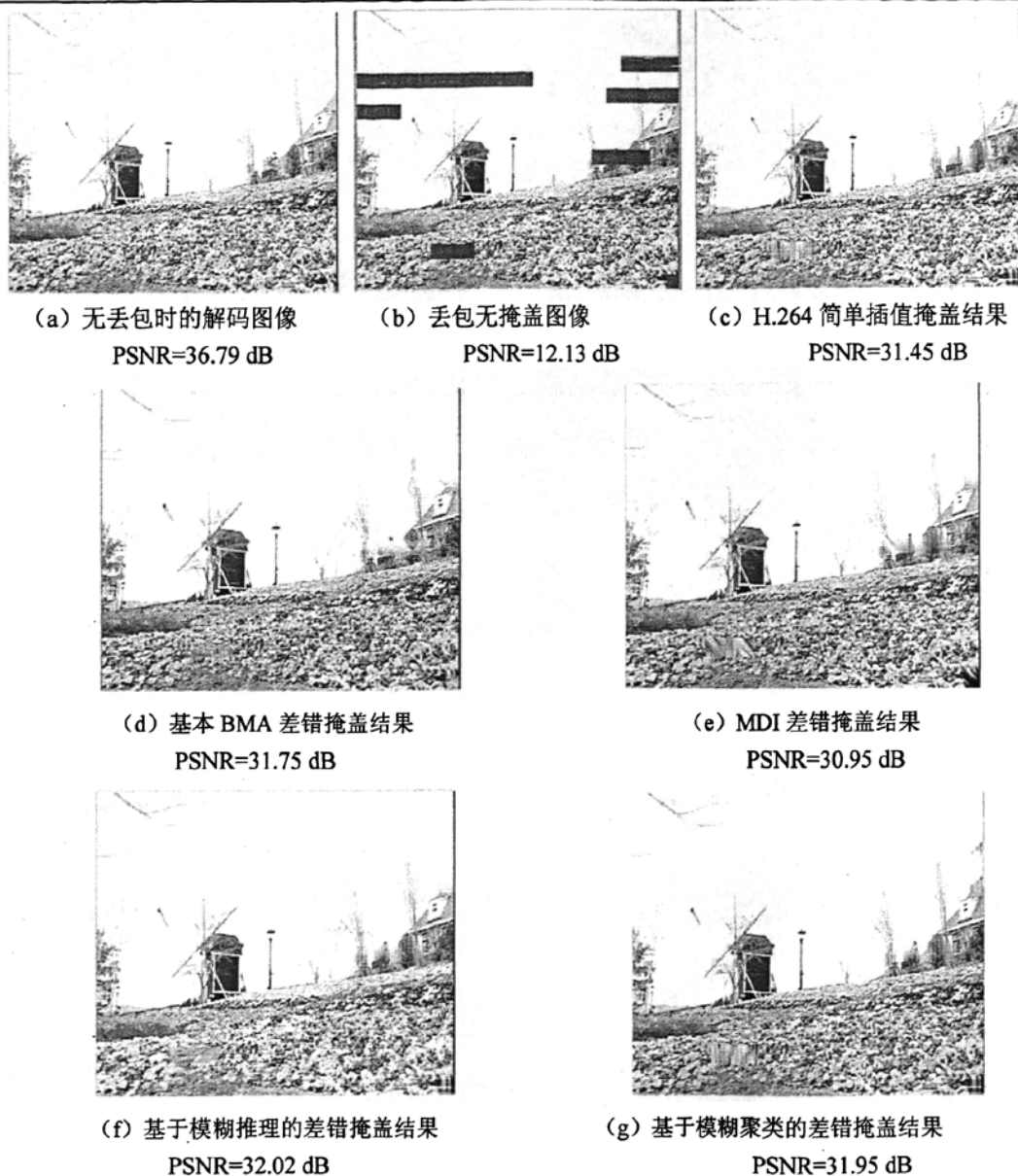


图 6-4 flower 在丢包率为 15% 时的解码图像

从图中可以看出,上述各种算法在平滑区域的掩盖效果一致,但在细节和纹理丰富的区域产生差别;基于简单像素加权插值的方法在细节和纹理区域造成模糊现象,而且像素插值不准确,造成颜色交错混杂;基于基本块匹配的方法对于细节丰富的区域用块拷贝的方法取得了较好的效果,如花朵区域,但在树干和房屋区域的相似块不准确,掩盖质量下降;基于多纹理方向插值的算法在树干区域产生了错误的纹理方向,在花朵区域的纹理方向过少,导致掩盖质量下降;基于模糊推理的方法和基于模糊聚类的掩盖方法在平滑区域和细节丰富的区域都能够产生良好的掩盖效果,信噪比也高于基于经典数学的掩盖方法,其中基于模糊推理的方法在房屋上空的掩盖质量优于基于模糊聚类的方法,这是由于基于

模糊推理的方法存在像素修正的过程，对掩盖明显出错的地方能够加以修正。

### 6.2.3 不同丢包率下的统计结果分析及结论

为了充分观察各个方法的掩盖效果，我们选取 6 个不同平滑程度的序列进行实验，对各个序列编解码 30 帧，统计平均丢包率和各个方法掩盖后的平均信噪比，结果如表 6-1 所示。

表 6-1 不同丢包率下的统计平均信噪比

| 序列名称<br>(cif 格式) | 平均丢包率 | 平均信噪比(dB) |        |        |       |       |
|------------------|-------|-----------|--------|--------|-------|-------|
|                  |       | H.264 掩盖  | BMA 掩盖 | MDI 掩盖 | FR 掩盖 | FC 掩盖 |
| foreman          | 5.8%  | 29.49     | 29.79  | 30.06  | 29.96 | 30.42 |
| highway          | 3.1%  | 37.43     | 34.43  | 37.49  | 37.94 | 38.97 |
| hall             | 4.7%  | 30.04     | 29.77  | 31.52  | 31.76 | 31.68 |
| mobile           | 4.0%  | 30.36     | 30.85  | 29.50  | 31.46 | 31.01 |
| flower           | 4.5%  | 31.59     | 31.73  | 31.21  | 33.06 | 32.24 |
| waterfall        | 3.8%  | 33.50     | 33.73  | 33.03  | 34.21 | 34.13 |

由以上的图形显示和表格数据分析可知：H.264 所采用的基于简单像素加权内插的方法即平滑滤波的方式在图像平滑的区域效果还是比较好的，但在有丰富细节或边缘的地方就会使图像模糊；基于块匹配的方法由于采用了基于统计量 MSE 最小的准则去寻找匹配块，对于纹理细节丰富的区域有较好的恢复效果，但是对于单一纹理方向的边缘区域会产生虚假边缘或边缘不连续的情况；基于多纹理方向插值的算法对于纹理方向较少而且明显的边缘区域能够准确的判断出纹理方向并插值恢复，但是对于细节比较丰富的区域恢复效果不佳。文中所提出的基于模糊数学的方法不论对于平滑区域还是纹理细节比较丰富的区域，总体说来效果较好，恢复后的图像的信噪比值也有上升。即使信噪比略有降低，但是由于客观度量并不完全代替人的视觉系统，从图形上看，主观质量仍可以得到改善。其中，由 5.3 节的分析可知，基于模糊聚类的方法的计算量介于基于像素插值恢复的方法和基于块匹配恢复的方法之间，但基于模糊推理的方法采用了块匹配的过程寻找相似块并在恢复准确度低的情况下，需要逐像素修正，因此计算量略高于基于模糊聚类的方法。

### 6.3 本章小结

本章首先介绍了实验环境与平台,然后分别测试了具有不同特性的视频序列的 I 帧在不同宏块丢失率的情况下,使用本文的基于模糊数学的差错掩盖方法和各种基于经典数学的掩盖算法恢复视频序列。与基于经典统计学的各类方法相比,本文算法在主观和客观质量上均有一定程度的提高和改善。本文算法能够恢复复杂的边缘和纹理,平均信噪比比 H.264 标准推荐的算法高出 1 个 dB 左右。

## 第七章 总结与展望

视频压缩编码是视频传输与通信等应用的必要条件,为此,人们已提出了多种视频压缩编码标准,由于去除了视频在时间和空间上的相关性,压缩码流对传输错误更加的敏感。在众多的抗误码技术中,基于视频解码端的差错掩盖是经常使用的有效改善误码图像质量的方法之一。目前的帧内差错掩盖技术往往是利用图像的空间上的相关性来进行的,它对解码后的视频进行统计分析、计算和处理,从而得到反映视频特征的特征量作为约束变量,在使它满足某个约束条件的情况下,估计差错块的像素值。国内外相关文献表明,常用的差错掩盖方法中所用到的约束条件主要有五个方面:一致性、光滑性、统计相关性、边缘连续性、运动矢量的连续性。这种建立在信号的统计分析、计算机数值计算和问题求解基础上的技术体系,对确定性、完全性的信息对象已形成了一系列传统的信息处理技术。但从其处理方法上看,它们大部分都是利用满足某单一的约束条件来得到差错块的掩盖块,这种方法能得到某一约束条件的满足,但不可能同时满足其它的约束条件。模糊数学在控制论和人工智能方面取得了巨大成功,近年来也逐渐应用于图像分割、边缘检测等图像处理系统中,这给了我们启示:图像信息中存在模糊性,因此可以将模糊集理论引入差错掩盖。

模糊推理和模糊聚类分析是模糊集理论的两大重要分支:模糊推理根据推理规则和输入的模糊变量,得到输出模糊变量,它是一种近似推理,推理规则的制定可以符合人的视觉系统的特征,对差错掩盖结果进行模糊判决以修正差错掩盖结果。而模糊聚类分析能综合视频多个相似性度量的特征量,对视频的某些区域进行划分,然后用同类的区域进行视频的差错掩盖,就能提高掩盖后视频的质量,这是因为在掩盖的过程中同时满足了视频差错掩盖的多个约束条件。

本论文的主要工作包括:

(1) 研究和分析各类基于经典统计学的帧内差错掩盖算法,包括:简单像素加权插值算法,重点介绍了 H.264 的帧内差错掩盖流程;基本的块匹配算法和改进的块匹配算法;基本的基于纹理方向的插值算法和改进基于内容划分的插值算法;并分别分析这些算法的优点和主要缺陷。

(2) 研究基于模糊推理的手段处理视频差错掩盖问题,提出了基于模糊推理的视频帧内差错掩盖方法,将模糊推理应用于是否接受初步差错掩盖结果的判决上,合理修正掩盖结果,使得边缘连续。经过实验论证,这种方法能够改善虚假边缘和边缘不连续的情况,提高了图像的掩盖质量。其中 4.2.2 节重点讨论了推理规则的制定和隶属度函数的选取,

推理规则主要是根据人的主观视觉选择,隶属度函数也选择了 Z 型隶属度函数和 S 型隶属度函数的组合,进行合理地模拟人类视觉系统。

(3) 提出基于模糊聚类的帧内差错掩盖方法,它首先利用帧内差错块周围相邻块的平滑程度和纹理方向等约束条件,对差错块周围的相邻块进行模糊分类,然后用相似块的像素插值得到差错块的像素值。经过实验测试,不论是在细节丰富和运动复杂的场合还是在细节较少和运动不复杂的场合,该算法掩盖后的视频质量有明显地提高。第五章重点讨论了最佳阈值的设定和算法的计算复杂度。最佳阈值的准确设定可以根据统计量的方法,但为了折中计算量和准确度,我们选用了实验观察确定的方法。

(4) 在 H.264 参考软件 JM8.6 平台上实现上述各类算法,并使用具有不同特性和不同误码率的序列对这些算法进行测试,比较这些算法的优缺点。

因为传输误码的不可避免,如何有效充分利用空间和时间冗余恢复错误区域是一个复杂而长期的研究内容。由于图像的复杂性和人眼视觉的特殊性,人们虽然已经提出了各种帧内方法进行差错掩盖,但仍然有较大的改进空间。实验表明本文提出的帧内差错掩盖算法的主客观效果有一定的提高,但算法也有一定的局限性,因此还需要进行优化和改进。

下一步的工作包括:

(1) 在基于模糊推理的方法中,对于所有的图像序列采用了同一套推理规则,从而简化了推理规则的制定;但针对某个特定的序列需要调整参数,参数的调整与实验者的经验有关,下一步可以设定自适应的参数调整系统。

(2) 对算法进行优化,进一步降低算法的复杂性,使本文算法能达到实时解码系统的要求。

(3) 本文目前主要研究对象是帧内差错掩盖技术,下一步工作应该结合较好的帧间差错掩盖方法以获得更佳的恢复质量。



## 参考文献

- [1] T. Wiegand, G. J. Sullivan, G. Bjontegaard, A. Luthra. Overview of the H.264/AVC video coding standard [J]. IEEE Transactions on Circuits System for Video Technology, 2003, 13: 560-576.
- [2] Joint Video Team (JVT) of ISO/IEC MPEG&ITU-T VCEG. Draft ITU-T recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264-ISO/IEC 14496/10 AVC). May, 2003.
- [3] Ye-Kui Wang, Hannuksela, M.M., Varsa, V., Hourunranta, A., Gabbouj, M.. The error concealment feature in the H.26L test model [C]. 2002 International Conference on Image Processing, Sept. 2002, 2: II-729 - II-732.
- [4] 龚时镠, 朱秀昌. 视频通信中的误码掩盖技术初探[J]. 电视技术, 2000, 6:71-73.
- [5] 梁凡. 差错信道下视频传输关键技术的研究: [博士学位论文]. 2001, 广州: 中山大学.
- [6] 江舟. 视频通信容错技术的研究与实现: [硕士学位论文]. 2002, 武汉: 武汉大学.
- [7] 廖彬, 冯穗力, 叶梧. 基于 H.264 的差错掩盖技术的研究[J]. 电视技术, 2005, 1:26-28.
- [8] Li-Wei Kang, Jin-Jang Leou. A hybrid error concealment scheme for MPEG-2 video transmission based on best neighborhood matching algorithm [C]. 2004 IEEE International Conference on Multimedia and Expo., June 2004, 2: 1355-1358.
- [9] Nemethova, O., Al-Moghrabi, A., Rupp, M.. Flexible error concealment for H.264 based on directional interpolation [C]. 2005 International Conference on Wireless Networks, Communications and Mobile Computing, June 2005, 2: 1255-1260.
- [10] H. Sun, W. Kwok. Concealment of damaged block transform coded images using projection onto convex sets [J]. IEEE Transactions on Image Processing, 1995, 14(4): 470-477.
- [11] Shahram Shirani, Faouzi Kossentini, Rabab Ward. An adaptive Markov random field based error concealment method for video communication in an error prone environment [C]. IEEE International Conference on Acoustics, Speech and Signal Processing, Mar 1999, 6: 3117-3120.
- [12] 王保平. 基于模糊技术的图像处理方法研究: [博士学位论文]. 2004, 西安: 西安电子科技大学.
- [13] L.A. Zadeh. The concept of a linguistic variable and its application to approximate reasoning [J]. Information Sciences, 1975, Part 1: 8:199-249; Part 2: 8: 301-357; Part 3: 9:43-80.
- [14] L. A. Zadeh. Fuzzy Sets [J]. Information and Control, 1965, 8: 338-353.
- [15] Kaufmann, A., Gupta, M. M.. Introduction to fuzzy arithmetic: theory and applications [M]. 1985, New York: Van Nostrand Reinhold Company Inc.

- [16] 王国俊. 关于模糊推理的若干问题[J]. 工程数学学报, 2000, 17(5): 9-13.
- [17] 王国俊. 模糊推理与模糊逻辑[J]. 系统工程学报, 1998, 13(2):1-16.
- [18] 王国俊. 模糊推理的一个新方法[J]. 模糊系统与数学, 1999, 13(3):1-10.
- [19] 林国辉, 门爱东, 全子一. 模糊算子在图像后处理中的应用[J]. 电视技术, 2001, 1:24-25.
- [20] 何清. 模糊聚类分析理论与应用研究进展[J]. 模糊系统与数学, 1998, 12(2):89-94.
- [21] 高新波, 谢维信. 模糊聚类理论发展与应用的研究进展[J]. 科学通报, 1999, 44(21):2241-2251.
- [22] Bezdek J C, Hathaway R J. Recent convergence results for the fuzzy c-means clustering algorithm [J]. Journal Classification, 1988, 5(2): 237-247.
- [23] 鲍正益. 模糊聚类算法及其有效性研究: [硕士学位论文]. 2006, 厦门: 厦门大学.
- [24] 高新波. 模糊聚类分析及其应用[M]. 2004, 西安: 西安电子科技大学出版社.
- [25] Bezdek J C, Hathaway R J, et al. Convergence and theory for fuzzy c-means clustering: counter examples and repairs [J]. IEEE Transactions on PAMI, 1987, 17(5): 873-877.
- [26] 王晟. 模糊聚类算法的研究与实现: [硕士学位论文]. 2006, 南京: 南京理工大学.
- [27] Batuwita, K.B.M.R., Bandara, G.E.M.D.C.. Fuzzy recognition of offline handwritten numeric characters [C]. 2006 IEEE Conference on Cybernetics and Intelligent Systems, 2006, 1-5.
- [28] Schulte, S., De Witte, V., Kerre, E. E.. A fuzzy noise reduction method for color images [J]. IEEE Transactions on Image Processing, 2007, 15(6): 1425-1436.
- [29] Hanmandlu, M., Jha, D.. An optimal fuzzy system for color image enhancement [J]. IEEE Transactions on Image Processing, 2006, 10(15): 2956-2966.
- [30] Xinjian Chen, Jie Tian, Xin Yang. A new algorithm for distorted fingerprints matching based on normalized fuzzy similarity measure [J]. IEEE Transactions on Image Processing, 2006, 15(3): 767-776.
- [31] Yao Nie, Barner K.E. The fuzzy transformation and its applications in image processing [J]. IEEE Transactions on Image Processing, 2006, 15(4): 910-927.
- [32] Avesta N.. Wavelet-Fuzzy-Stochastic Kalman filtering for image compression [C]. 2006 IEEE International Conference on Multimedia and Expo, July 2006, 717-720.
- [33] Jinghong Zheng, Lap-Pui Chau. An efficient spatial domain error concealment method for H.264 video [C]. Proceedings of the 2003 Joint Conference of the Fourth International Conference on Information Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia, Dec 2003, 1:26-30.
- [34] 徐锦亮. 基于 H.264 的视频差错掩盖技术研究: [硕士学位论文]. 2007, 四川: 西南交通大学.
- [35] Zhou Wang, Yinglin Yu, Zhang, D.. Best neighborhood matching: an information loss restoration

- technique for block-based image coding [J]. IEEE Transactions on Image Processing Systems, 1998, 7(7): 1056-1061.
- [36] Mei-Juan Chen, Che-Shing Chen, Ming-Chieh Chi. Recursive block-matching principle for error concealment algorithm [C]. ISCAS'03, Proceedings of the 2003 International Symposium on Circuits and Systems, 2003, 2: 528-531.
- [37] S. Tsekeridou, L. Pias. MPEG-2 error concealment based on block matching principles [J]. IEEE Transactions on Circuits Systems for Video Technology, 2000, 10(4): 646-658.
- [38] Yen-Chi Lee, Yucel Altunbasak, Russell M, Mersereau. Coordinated application of multiple description scalar quantization and error concealment for error-resilient MPEG video streaming [J]. IEEE Transaction on Circuits and Systems and Video Technology, 2005, 5(4): 457-468.
- [39] J.W.Suh, Y.S.Ho. Error concealment based on directional interpolation [J]. IEEE Transactions on Consumer Electronics, 1997, 43(3): 295-302.
- [40] Shih-Chang Hsia. An edge-oriented spatial interpolation for consecutive block error concealment [J]. IEEE Signal Processing Letters 2004, 11(6): 577-580.
- [41] Jianle Chen, Jilin Liu, Xingguo Wang, Guobin Chen. Modified edge-oriented spatial interpolation for consecutive blocks error concealment [C]. Proceedings of IEEE International Conference on Image Processing, Sep 2005, 3: III - 904-7.
- [42] Rong-fu Zhang, Yuan-hua Zhou, Xiao-dong Huang. Content-adaptive spatial error concealment for video communication [J]. IEEE Transactions on Consumer Electronics, 2004, 50(1): 335-341.
- [43] Xeroxing Lee, Ya-Qin Zhang, Leon-Garcia A.. Information loss recovery for block-based image coding techniques-a fuzzy logic approach [J]. IEEE Transactions on Image Processing, 1995, 4(3): 259-273.
- [44] 谢季坚, 刘承平. 模糊数学方法及其应用[M]. 2006, 武汉: 华中科技大学出版社.
- [45] 郭秀江, 朱秀昌. 一种基于模糊聚类的视频帧内差错掩盖[J]. 南京邮电大学学报, 2006, 26(4):41-43.
- [46] 郭秀江. 基于模糊数学的视频差错掩盖技术研究: [博士学位论文]. 2006, 南京: 南京邮电大学.
- [47] 倪伟, 郭宝龙, 王勇. H.26L中新的帧内预测技术研究[J]. 计算机工程与应用, 2004, 8:42-45.

## 致 谢

在论文搁笔之时，我要感谢所有关心和帮助我的老师、同学和家人。

首先要向我的导师朱秀昌教授表示深深的感谢，朱老师在学业上给予我悉心指导和大力帮助，生活上给予关心和照顾，使我得以顺利完成论文工作。在这期间，导师朱秀昌教授严谨的治学态度，灵活的学术方法，活跃的思维方式，求实的学术风格，扎实的科研精神，为我们营造了一种良好的学术氛围。我置身其间，耳濡目染，学到了新的思维方法，掌握了常用的研究方法，……。这些宝贵的财富将使我终生受益。导师严以律己、宽以待人、朴实无华、平易近人的人格魅力，无微不至、感人至深的人文关怀，乐观向上、积极主动的生活态度，令我如沐春风，明白了许多待人接物与为人处世的道理。总之，回想和导师在一起的时光，使我倍感温暖，对导师道一句真诚的祝福：愿导师合家欢乐，一生平安！

与此同时，我还要深深地感谢干宗良博士、郭秀江博士，师兄吴明星、詹学峰，图像实验室的宋建新教授、胡栋教授、刘峰教授以及所有其他老师对本人在学术、科研上的支持和指导帮助，在我学术上遇到困难的时候，他们为我指出了问题解决的可能方向；其次，感谢同师门的同学在学习和研究过程中与我所作的有益探讨，感谢室友冯荃、高小云及其他同学和朋友对我的关心和帮助，使我度过了难忘和快乐的两年半的时光。

最后，特别感谢我的父母，他们崇尚知识的理念和热爱生活的态度对我影响至深，为我营造了良好地学习生活条件，给予我无条件的支持和信任，正是他们无私的付出和帮助，使我得以完成学业，他们无私的爱和“春蚕到死丝方尽、蜡烛成灰泪始干”的奉献精神使我深受感动！

沈寒

二零零八年三月

## 攻读硕士论文期间参加的科研项目和发表的论文

### 发表论文

(1) Han Shen, Xiu-chang Zhu, A novel spatial information recovery algorithm based on fuzzy clustering in H.264. 2007 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS 2007).

### 参与的科研项目

(1) 参与项目“基于 Fuzzy Sets 的视频差错掩盖技术研究”的研究（国自基金 60672134）。