

一种预测起始搜索点的运动图象 分块运动匹配新算法

许洁斌 韦岗

布礼文

(华南理工大学电子与通信系, 广州 510641) (香港城市大学电子工程系)

摘要 在以往的运动图象分块运动匹配算法中,分块运动向量的起始搜索点是固定的。本文根据各分块运动向量的分布统计特性和相邻分块运动向量的相关特性,提出了一种基于预测起始搜索点的运动图象分块运动匹配新算法。通过用相邻分块运动向量的相互关系预测当前分块的起始搜索点,利用基于中心搜索模式的分块运动匹配算法寻找匹配的运动向量。实验表明,该算法预测步骤简单,令搜索更快接近全局极值,与其它快速分块运动匹配算法相比有效地降低了帧间运动补偿误差和所需搜索运算量,效果较好。

关键词 运动图象编码,分块运动匹配,四步搜索法

1 引言

近年来,随着计算机和网络技术的发展,运动图象和语音也被用于数字通信中。为适应高清晰度电视、电视会议和可视电话的需要,国际上已提出多个运动图象压缩编码标准,如 ITU H. 261^[1], MPEG-1^[2]和 MPEG-2^[3]等。这些标准的一个共同特点是它们都采用 DCT 变换去除空间冗余,用分块运动匹配算法(BMA)去除时间冗余。

在 BMA 算法中,每帧图象被分为众多 $N \times N$ 大小的方块($N=8$ 或 16),然后对参考图象进行搜索,在给定匹配标准下找出补偿误差最小的方块,从而得到该对应方块的运动向量,这个过程就是分块运动匹配。以这个运动向量描述整个方块的图象内容来达到信息压缩,这个过程就是分块运动补偿。匹配标准有均方误差(MSE)、平均绝对误差(MAD)等,其中平均绝对误差不需作乘法运算且能得到与均方误差相近的效果,所以被广泛使用。全搜索法(FS)是最直接使用的 BMA 算法,它在搜索范围内检查所有候选点,因此能得到全局最优点。但其运算量庞大,不适于实时应用。为此,一系列次优快速

BMA 算法如三步搜索法(3SS)^[4]、正交搜索法(OSA)^[5]、二维对数搜索法(2D-LOGS)^[6]、预测法^[7,8]等得到了发展。最近,根据分块运动向量基于中心分布(center-biased)的特点,又提出新三步法(N3SS)^[9]、四步搜索法(4SS)^[10]和梯度下降搜索法(BBGDS)^[11],有效降低了运动补偿误差,减少了平均搜索点。但这些算法都是以分块所在位置为中心的一个窗口(大小为 ± 7 或 ± 15)内进行搜索,且都从该中心点开始搜索。所以如果物体运动较快,全局最优点离中心点较远时,这些算法的平均搜索步数会较大,而且更有机会落入局部极值陷阱。国际电信联盟(ITU)有关运动图象编码最新测试标准,即 Video Coding Experts Group 的 Video Codec Test Model 测试模型 TMN8 在运动估值方面作了几方面建议:(1)在基本模式下采用了半像素搜索和帧内/帧间模式选择,(2)在增强预测模式下还增加了 8×8 子分块的运动估值,(3)非限制运动向量模式下,先在 16×16 分块作全搜索得到一个预测因子,然后以这个预测因子为中心在一定的搜索范围作 16×16 和 8×8 的全搜索。但是在产生预测因子时并没有利用相邻分块运动向量的相关特性,计算量仍然很大。我们通过分析相邻分块运动向量的相关

特性,重新选择起始搜索点,使搜索一开始就接近全局极值,据此提出了基于预测起始搜索点的运动图象分块运动匹配新算法。

2 分块运动向量的运动关联性

在BMA算法中,分块运动向量的估计可被假设为一个二维空间的寻优问题。在搜索过程中,假设匹配误差随着搜索点向全局最优点的逼近而单调下降,在3SS中还使用统一分布的搜索点模式。但现实世界中,由于存在局部极值点,而且分布规律的复杂,使得上述假设并不正确。实验表明,真实世界里图象的分块运动是平滑的,而且变化较慢。这使得全局最优解呈现中心分布的规律^[9~11]。所以采用侧重中心点的不均匀搜索模式,如N3SS、4SS和BBGDS等算法就取得较好效果。但当物体运动较快,全局最优解不在中心点附近时,如何选择适当的起始搜索点,使全局最优解落入搜索范围并尽量接近起始点的问题就显得十分重要。

文献[8]的研究表明由于在现实图象序列中,一个运动物体会覆盖多个分块,所以空间相邻分块的运动可能较相似。又由于物体运动在时间上的连续性,前后帧图象的时间相邻分块的运动也具有相似性。文献[8]就是利用分块运动的关联性进行运动向量搜索的。但是当物体运动较快时,其运动的方向和速度并不与图象序列的帧采样率保持稳定,所以在时间轴上利用前后帧分块运动的关联性进行运动预测效果并不好,而且要保存前帧的分块运动向量数据会占用解码器较多的存储空间,不利于在硬件上实现。因此本文将集中研究分块运动的空间关联性。

图1给出了全搜索法在±7搜索范围得到的图象序列“Tennis”在第50帧的分块运动向量分布图,当时图象情景包括了物体快速运动和摄像镜头推拉

等效果。从图可看出当分块处于同一个物体或背景,它们与空间相邻分块的运动向量的大小和方向是非常相似的。如果空间相邻的分块运动向量已知,那么当前分块运动向量的近似大小和方向就很容易被预测了。在本文,我们选择的4个空间邻块如图2所示。 B_0 是当前块, B_1, B_2, B_3, B_4 分别是水平和垂直方向的空间邻块。

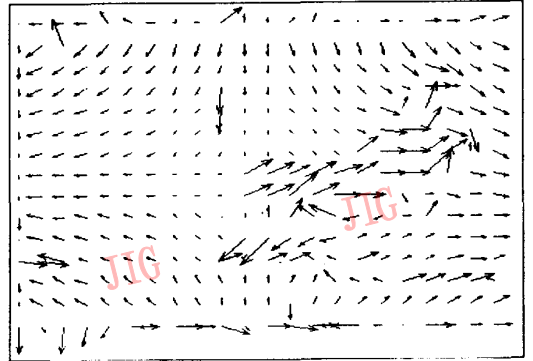


图1 “Tennis”第50帧运动向量分布图

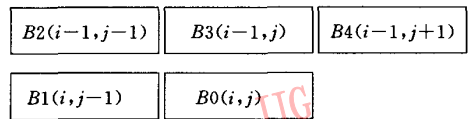


图2 当前分块与其4个相邻分块

我们定义当前块与4个邻块的运动关联性为

$$\vec{D}_{mv} = \vec{D}_{B_0} - \frac{1}{4} \sum_{i=1}^4 \vec{V}_{B_i} \quad (1)$$

其中 \vec{D}_{mv} 是一个偏移向量, $\vec{V}_{B_i}, i=0, 1, 2, 3, 4$,是上述分块对应的运动向量。

式(1)描述了4个空间邻块的运动中心与当前分块的真实运动的偏差情况, \vec{D}_{mv} 越小表明当前分块与4个邻块的运动越有关联。我们对2组标准图象序列“Football”和“Tennis”作了的分布特性统计(见表1、图3)。其中式(1)的 $\vec{V}_{B_i}, i=0, 1, 2, 3, 4$,是在±7搜索范围对16×16分块使用全搜索法所得,因此,

表1 \vec{D}_{mv} 分布统计

测试图象序列		Football		Tennis	
		所有情形	运动情形	所有情形	运动情形
分块在所有区域(±14)的 \vec{D}_{mv} 统计	分块数(百分比)	22 400(100%)	7 341(100%)	22 400(100%)	16 224(100%)
	距真实向量的平均偏移	$ x =0.420, y =0.393$	$ x =0.973, y =0.913$	$ x =0.904, y =0.732$	$ x =1.195, y =0.930$
分块在5×5区域的 \vec{D}_{mv} 统计	分块数(百分比)	22 086(98.60%)	7 063(96.21%)	21 732(97.02%)	15 579(96.02%)
	距真实向量的平均偏移	$ x =0.365, y =0.338$	$ x =0.836, y =0.775$	$ x =0.789, y =0.624$	$ x =1.048, y =0.792$
$\max_{i=1 \rightarrow 4} \ \vec{V}_{B_i} - \vec{V}_m\ < T$ 满足时的 \vec{D}_{mv} 统计	分块数(百分比)	21 849(97.74%)	6 959(94.80%)	20 423(91.17%)	14 464(89.15%)
	距真实向量的平均偏移	$ x =0.370, y =0.345$	$ x =0.873, y =0.184$	$ x =0.731, y =0.579$	$ x =0.990, y =0.754$

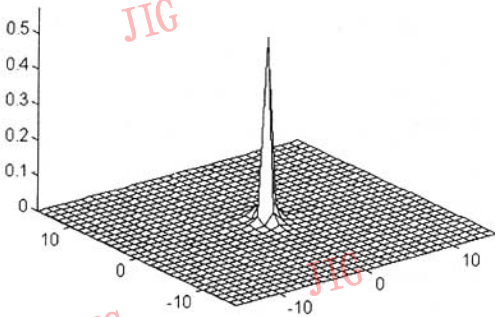


图 3 “Tennis”相邻分块的运动关联统计分布图

\vec{D}_m 的最大偏移是 ± 14 。该分布特性是从每个序列取 80 帧,每帧取 280 个分块,所以共得到 22 400 个数据。由于图象序列运动向量的中心分布特性,许多分块的运动向量为零。所以表 1 分所有测试分块和运动分块 2 种情况考虑,以分别考察所有情形和只有运动时的分布。从表 1 可看出在这 2 种情形,4 个邻块的运动中心都非常接近当前块的真实运动向量,特别是在中间 5×5 区域,2 种情形集中了近 97% 和 96% 的分布。这表明 4 个邻块的运动提供了一种很好的预测当前分块运动的信息。然而当 4 个邻块不在同一运动物体时,其运动中心不但不能给出有效的预测,反而会引起误差扩散。因此,我们定义判断式 $\max_{i=1 \rightarrow 4} \|\vec{v}_{Bi} - \vec{v}_m\| < T$ 来考察 4 个邻块运动的一致性。从表 1 第 3 行的统计数据看出,满足上述判断式的分块,其运动预测很接近真实运动,其平均偏移比第 1 行用 4 个邻块运动中心作预测的要小,而非常接近第 2 行在 5×5 区域的平均运动偏移量。通过这些数据我们认为这种判断方法在未知真实运动向量的情况下,已能简单有效地预测出 5×5 中心区域内绝大部分的真实运动向量。

3 预测起始搜索点的分块运动匹配算法

算法分为预测起始搜索点和基于中心分布的搜索两大步。

3.1 预测起始搜索点

在这一步,我们首先判断当前分块与其周围 4 个邻块的运动关系。令 $\vec{v}_{Bi}, i=0,1,2,3,4$, 为如图 1 所示的 5 个分块的运动向量, \vec{v}_{init} 为当前分块起始搜

索点距离分块原点的位移量。我们用来描述 4 个相邻分块的平均运动。当这 4 个邻块的运动向量都很接近时,我们就假定这 5 个分块很大机会是在同一个图象物体或都在背景里。此时, \vec{v}_m 能够给出较好的运动预测信息。具体算法见式(3)~(5),当分块的运动存在关联,则用预测起始位移量 \vec{v}_p 作为 \vec{v}_{init} 。否则,用分块原点作起始搜索点,即:

$$\vec{v}_{init} = \begin{cases} \vec{v}_p, & \text{if } \max_{i=1 \rightarrow 4} \|\vec{v}_{Bi} - \vec{v}_m\| < T \\ (0,0), & \text{otherwise} \end{cases} \quad (2)$$

其中 T 是一个预定的阈值,用以描述 4 个邻块运动的相似程度。

基于分块运动向量的中心分布特性^[9,10],我们提出以下 3 种预测 \vec{v}_p 的方案:

(1)最小运动预测:

$$\vec{v}_p = \operatorname{argmin}_{\vec{v}_{Bi}} \|\vec{v}_{Bi}\|, \quad i = 1, 2, 3, 4 \quad (3)$$

(2)平均运动预测:

$$\vec{v}_p = \operatorname{round}(\vec{v}_m) \quad (4)$$

(3)最小平均运动预测:

$$\vec{v}_p = \operatorname{argmin}_{\vec{v}_{Bi}} \|\vec{v}_{Bi} - \vec{v}_m\|, \quad i = 1, 2, 3, 4 \quad (5)$$

其中 $\operatorname{round}(\cdot)$ 是对向量的所有分量的取整运算。

最小运动预测符合现实世界图象的变化通常是平滑和缓慢的特性,但这种方法不能很快地跟踪一些快速运动。而平均运动预测则非常适合运动分块都在同一个物体的假设成立的情况,它能给出准确的运动匹配。然而,从模拟实验我们发现有时候相邻的 4 个邻块所覆盖的范围太大,不能分辨一些细微的运动。因此当预测有偏差时,这种方法较易引起进一步的误差扩散。最小平均运动预测方法选择离邻块平均运动中心点最近的一个运动向量作为当前分块的初始搜索起点。当上述 4 个邻块同在一个运动物体的情况出现时,这种预测给出了离真实位置(即平均运动中心)最近的估计。当这些分块属于不同的运动物体而式(2)的判断值又未超出阈值 T 时,这种预测所给出的最小位移估计又能体现运动的平滑缓慢的特性。所以这种方法能在运动相关假设的判断正确与否的两种情况下取得较好的平衡。我们的实验也证实了这点,所以在本文的实验中我们只用最小平均运动预测。

新算法对分块运动关联性的判定和起始点的预测共使用了 18 次加法运算、8 次乘法运算和 7 次比较运算,而 BMA 算法对 16×16 的分块每搜索一点需要 511 次加法运算、256 次绝对值运算和 1 次比

较运算,可见预测只占用极小的运算量。

3.2 基于中心分布的搜索

在得到起始搜索点后,无论分块间运动的相关性如何,上述预测算法都能使搜索窗中心 \vec{V}_{imi} 非常接近真实的运动向量。我们使用基于中心搜索模式的 N3SS 或 4SS 就能最有效检测出最终的运动向量,因为这两种算法的第一步使用侧重中心的搜索模式,大大增加了在中心 5×5 区域得到全局最优点的机会,并缩短了搜索步数。

4 实验结果

我们在计算机上对本文的算法与其它几种具有

代表性的和最近提出的快速 BMA 算法作了实验比较,这些算法包括 FS, 3SS, N3SS, 4SS。我们使用了"Football"、"Tennis"和"Garden"3种最常用的标准图象序列的前 80 帧作测试。这些测试序列包含了物体的快速运动和镜头的变焦和移动。序列每帧大小为 352×240 ,每个象素量化为灰度 8 比特。实验使用平均绝对误差(MAD)作分块匹配标准。本文提出的结合 N3SS 和 4SS 作第 2 步的新算法分别记为 AMTN3SS 和 AMT4SS。预定的阈值 T 是 5,使用最小平均运动预测作为第 1 步的起始点预测。在第 2 步中,实验为每个 16×16 分块使用的搜索范围为 ± 7 个象素点,与 N3SS 和 4SS 算法一样。

表 2 给出了 FS, 3SS, N3SS, 4SS, AMTN3SS 和 AMT4SS 对各图象序列前后 2 帧作运动补偿后的

表 2 测试序列的 MSE 和平均运算复杂度(以搜索每个运动向量所需的平均搜索点数来衡量)

搜索算法	测试序列的前 80 帧						测试序列前 80 帧并隔 1 帧取样					
	Football		Tennis		Garden		Football		Tennis		Garden	
	MSE	复杂度	MSE	复杂度	MSE	复杂度	MSE	复杂度	MSE	复杂度	MSE	复杂度
FS	156.98	225	132.12	225	293.32	225	226.19	225	198.11	225	527.22	225
3SS	176.84	25	186.21	25	342.17	25	297.66	25	254.43	25	630.46	25
N3SS	166.04	19.05	164.14	22.55	303.39	23.45	293.90	20.49	251.24	25.57	586.42	27.45
4SS	170.28	18.01	156.68	20.09	320.54	20.15	304.34	18.69	291.28	21.20	701.37	21.51
AMTN3SS	163.35	18.60	150.79	20.94	299.02	19.51	286.08	19.85	229.16	22.98	534.45	21.67
AMT4SS	167.19	17.72	149.97	18.75	304.67	17.97	294.13	18.38	267.61	20.05	553.63	19.03

均方误差(MSE)的统计数据。从 3 个测试序列的结果都可看出 AMTN3SS 和 AMT4SS 有效地提高了对应的 N3SS 和 4SS 的性能。图 4 给出了" Tennis"前 80 帧的均方误差曲线,可看出在运动快速的时候(30 至 80 帧),本文算法的 MSE 更贴近全搜索的结果,这表明算法能以极小的运算量较好地预测出物体的运动趋势。图 4 表 2 同时给出以寻找一个运动

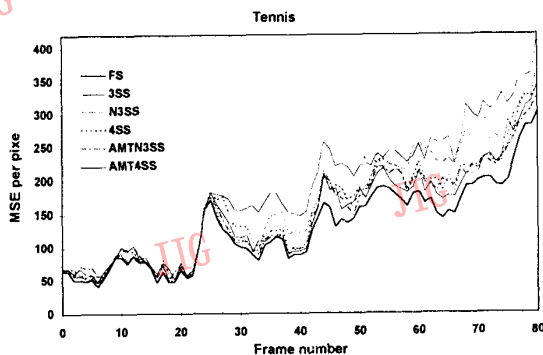


图 4 各种算法对" Tennis"的 MSE 变化曲线

向量所需平均搜索点数作为运算复杂度的实验数据,图 5 是" Tennis"前 80 帧的平均搜索点曲线。

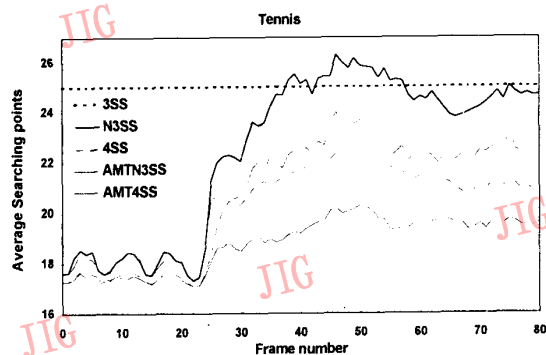


图 5 各种算法对" Tennis"的平均搜索点变化曲线

这些数据显示了本文算法令匹配误差减少的同时也使搜索运算下降,这是由于:(1)新算法能较好地判定分块运动的关联情况并作出适当的运动预测,使起始搜索点总是沿着真实的运动向量方向移动;(2)

起始搜索点的选择范围是 ± 7 ,第2步的搜索范围是 ± 7 ,所以整个算法的搜索范围扩大至 ± 14 ,有更多的候选机会,在运动快速的情况可能比 ± 7 全搜索法有更好的匹配值,而最差情形时的搜索点数仍与第2步 ± 7 的中心搜索法相同(N3SS的33点和4SS的27点)。表2同时给出了帧抽样率下降1/2的实验数据,结果显示在低码率时算法对运动的跟踪预测和起始搜索点的适当选择令匹配效果更好。从图5的曲线还可看出本算法的平均搜索点比较稳定地接近它的最小搜索点(N3SS和4SS都是17点),而其它算法则随图象内容的变化有起伏,这说明本算法的运动预测具有较好的鲁棒性。

我们的实验还表明算法性能对预定阈值 T 的取值不敏感。当选取 T 从 $\sqrt{10}$ 到 $\sqrt{40}$,每隔 $\sqrt{1}$ 取值一次,共作31次不同 T 值的试验,新算法AMT4SS对“Tennis”序列前80帧平均MSE的变化范围在149.97与151.26之间,“Football”序列在168.15与167.04之间。而平均搜索点分别在18.70与18.90之间和17.71到17.73之间。所有的变化都小于1%。实验肯定了新算法对描述分块运动趋势的有效性并不依赖某些预定值,相反在应用上性能保持得很好。

5 结论

本文提出一种基于预测起始搜索点的运动图象分块运动匹配新算法。通过分析图象分块运动向量的分布特性,我们利用空间相邻块的运动关联性预测起始搜索点,结合中心搜索模式的BMA算法能较好地跟踪匹配快速的运动。实验表明这种预测起始搜索点的新算法能有效降低运动补偿误差并同时减少搜索运算量,具有较好的鲁棒性,显著提高所用的BMA算法的性能。可以预见,将这种方法运用到其它BMA算法也将有助它们提高初始搜索的准确性和扩大搜索的范围,而增加的预测运算量极少。这种预测运动的方法能保持物体分块运动固有的连续性,所以它得到的匹配运动向量比全搜索法得到的均方误差意义下的全局最优点有更好的视觉效果。另外对运动趋势的有效描述和在性能上的简单稳定使得算法十分适合应用于实时硬件系统。

参考文献

- 1 CCITT SGXV. Description of referencemode l8(RM8). Document 525, Working Party XV/4, Specialists Group on Coding for Visual Telephony, 1989, 6.
- 2 ISO/IEC11172-2(MPEG-1 Video). Information Technology-Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s; Video, 1993.
- 3 ISO/IEC13818-2|ITU-T H. 262(MPEG-2 Video). Information Technology- Generic Coding of Moving Pictures and Associated Audio Information; Video, 1995.
- 4 Koga T, Iinuma K, Hirano A, et al. Motion compensated interframe coding for video conferencing, Proc. NTC81, 1981(7):c9. 6.1~9.6.5.
- 5 Puri A, Hang H M, Schilling D L. An efficient block matching algorithm for motion compensated coding, Intl. Conf. Acoust., Speech, and Signal Process, 1987, (4):1063~1066.
- 6 Jain J R, Jain A K. Displacement measurement and its application in interframe image coding. IEEE Trans. Commun, 1981(12):1799~1808.
- 7 李华等. 一种用于运动估值的预测搜索算法. 信号处理, 1996, 12(1):46~49.
- 8 Hsieh C. et al. Motion estimation algorithm using interblock correlation. Electronics Letters, 1990, 26(5):276~277.
- 9 Li R, Zeng B, Liu M L. A new three-step search algorithm for block motion estimation. IEEE Trans. on Circuits and Systems for Video Technology, 1994, 4(4):438~442.
- 10 Po L M, Ma W C. A novel four-step search algorithm for fast block motion estimation. IEEE Trans. Circuits Syst. Video Technol, 1996, 6(3):313~317.
- 11 Lurng-Kuo Liu Ephraim Feig. A Block-Based Gradient Descent Search Algorithm for Block Motion Estimation in Video Coding. IEEE Trans. Circuits Syst. Video Technol, 1996, 6(4):419~422.



许洁斌 1992年华南理工大学无线电系获学士学位,1995年华南理工大学无线电与自控所获硕士学位,现正在华南理工大学电子与通信系攻读博士学位。主要研究领域包括信号与图象处理、计算机视觉、图象识别和人工智能、图象通信等。



韦 岗 1984年清华大学无线电系获学士学位,1987年、1990年于华南理工大学无线电与自控所分获硕士、博士学位。现为该校电子与通信系教授、博士生导师。主要研究领域包括信号处理、模式识别、非线性系统理论及应用、神经网络等,已出版学术著作4部,发表论文40多篇。



布礼文 1988年香港城市大学电子工程系获第一荣誉学士学位,1991年香港城市大学电子工程系获该校第一个博士学位。现为该校电子工程系助理教授和城大图象处理实验室主任。在国际核心期刊和会议发表论文数十篇。主要研究领域包括图象处理、矢量量化、运动图象压缩码等。

A Novel Block Motion Estimation Algorithm by Predicted Start Searching Point for Video Coding

Xu Jiebin, Wei Gang

Po L. W.

(South China University of Technology, Guangzhou 510641) (City University of HongKong)

Abstract Most of the fast block matching algorithms (BMAs) use the origin as the initial center of the search. To improve the accuracy of the fast BMAs, a new predicted search algorithm is proposed in this paper. Based on the spatial correlation of motion and the fast center-based BMAs, the four causal neighbour blocks are chosen to predict the start searching point and then the center-based BMA is used to find the final motion vector as the global minimum is more close to the predicted center. Experimental results show that the proposed predicted search algorithm overdrives the performance very well in most fast BMAs such as the new three-step search, the four-step search in terms of motion compensation errors. In addition, the average computational requirement is also reduced.

Keywords Video coding, Block matching algorithm (BMA), Four-Step Search algorithm (4SS)

怎样写参考文献

- 1 在文稿中引用过的参考文献,按在文中出现的先后用阿拉伯数字顺序统一编号。
- 2 参考文献中不应列入未公开发表的文章、著作等,也不应列入内部资料和密级资料。
- 3 在正文中,当需说明引用内容的出处时,应把引用的参考文献编号连同方括号写在所注部分末字的右上角,如:××××^[1]。若所提及的参考文献作为正文中的直接说明语,则其编号连同方括号与正文并排,如:文献[2]所使用的方法是……。
- 4 参考文献的书写格式:

(1) 图书格式为:作者.书名.译者.版本(第1版可略).出版地:出版者,出版年:页数。

例:1 钱伟长.穿甲力学.北京:国防工业出版社,1984.

2 Bao Ting-Yu. The theory of potential equilibrium for interior ballistics and its application. Beijing: National Defence industry Press, 1988.

(2) 期刊书写格式为:作者.文章题名.译者.刊名,年,卷(期):页数。

例:1 朱自强.无激波机翼的设计方法.航空学报,1988,9(11):499~508.

2 Gray R M. Vector quantization. IEEE ASSP Magazine, 1984, 1, 4~29.

3 Bezdek J C, et al. FCM: The fuzzy c-mean clustering algorithm. Comput. Geosciences, 1984, 10(2): 191~203.