

● 陈廷标 夏良正 编著 何振亚 审

邮电高等学校教材

数字图象处理

人民邮电出版社

邮电高等学校教材

数字图象处理

陈廷标 夏良正 编著

何振亚 审

人民邮电出版社

内 容 简 介

本书论述了数字图象处理的基本概念、理论和系统组成。全书共分十一章，主要内容有数字图象处理的基础知识、基本方法和原理，重点讨论数字化、压缩编码、增强、复原、分割、分类和投影重建。最后介绍图象处理系统和设备。各章附有足够的习题以供选用。

本书可作为通信工程、自动控制、信息科学、电子工程、生物医学工程、计算机科学与工程等有关专业的教材，也可供有关专业硕士研究生和科技人员参考。

邮电高等学校教材

数字图象处理

陈廷标 夏良正 编著

何振亚 审

责任编辑 杜士选

*

人民邮电出版社出版

北京东长安街27号

北京交通印刷厂印刷

新华书店北京发行所发行

各地新华书店经售

*

开本：850×1168 1/32 1990年5月 第一版

印张：18 20/32 页数：298 1990年5月北京第1次印刷

字数：493千字 印数：1—2500册

ISBN 7-115-04201-2/G·018

定价：4.40元

前 言

图象技术的突出成绩在古代是印刷技术，到十八世纪是照相以及二十世纪的激光全息技术等。作为电子学方面的应用主要是电视。其它处理技术，由于图象信号信息量大、频带宽等特性，长期以来并未获得多大发展。直到近二十多年，一方面由于数字计算机和大规模集成电路的高速发展，使图象信息的处理、传输和存贮技术有可能得以实际应用。另一方面在现代信息社会中，越来越多的技术课题必须应用图象处理技术。如航天航空技术，通信工程，医疗卫生，工业探伤，气象预报，军事公安以及自动控制，机器人等领域中都有很多的实际应用要求。可以肯定，数字图象处理技术在今后必将获得越来越广泛的应用和更深入地发展。

本教材是为高等院校通信工程、自动控制，计算机科学与工程，电子工程，信息工程等专业学生编写的，亦可供有关专业硕士研究生和工程技术人员参考。书中有*号的章节视读者水平和要求可以增缩。虽然不同的专业或不同的应用要求，可能有不同的侧重，但基本概念和方法原理是共同基础。这就是本书编著的基本出发点。

全书分四部分：

第一部分包括一，二，三章。是准备知识。其中第二章图象信息的基本知识，主要介绍视觉图象信息的产生，传递和加工处理以及视觉模型的研究状况。目的是从仿生学角度出发，给读者深入研究图象技术以某种启发。还介绍了图象质量主客观评价方法的研究状况，最后扼要介绍图象的光学处理方法。第三章是数学预备知识。主要介绍图象的线性运算，图象变换以及二维线性数字滤波器。

第二部分包括四至七章，是数字图象处理技术的基本方法和原理。其中第四章图象数字化，主要介绍取样和量化的基本方法原理。第五章是图象压缩编码。有效地压缩图象信息比特数可以大大降低

对信道容量的要求。对节省存贮容量，缩短处理运算时间也十分必要。为了适应目前广泛应用的二值图象传输和处理，适当介绍了一些实用的二值图象编码方法。第六、七两章都是图象改善技术，前者以人的视觉系统为对象，按照人的主观要求对图象进行某种改善如灰度变换，平滑去噪声，锐化以及伪彩色技术等图象增强方法。后者主要是要将降质图象尽可能地复原到原始图象，往往要寻求降质原因，再将其模型化，然后采用“对症下药”的方法进行改善。即所谓图象复原技术。

第三部分包括八、九、十章，主要介绍图象分割，分类分析，投影重建和模式识别等技术的基本概念。读者可以选择其中与自己相关部分结合文中所提供的资料，有选择地重点学习。

第四部分主要介绍图象系统和设备工作原理以及典型系统的性能。对目前已经出现的平行处理技术等新技术也作了简略介绍。使读者了解一些系统发展的动态。

阅读本书时，我们认为读者已具有积分变换、线性代数、概率论和随机过程、数字信号处理、通信理论、计算机语言等方面的基本知识。本教材以编者对南京几所高校本科生和硕士研究生的讲稿为主，结合外籍专家讲学内容，加上国内外一些单位的科研成果综合而成。根据浙江大学姚庆栋教授、北京邮电学院全子一教授、清华大学王汉生副教授评审意见，邮电高校教材编审委员会评审通过作为通用教材使用。最后请东南大学何振亚教授审定。对此作者深表感谢。

本书第三、七、十、十一章及附录一由夏良正编著，其余各章由陈廷标编著，并统编全书，王岚和杜桂芬参加部分工作。由于作者水平所限，书中难免有错误与不妥之处，恳请读者批评指正。

作者

1989.5

第二章 图象信息基本知识

2-1 图象与视觉

2-1-1 概述

为了有效地设计图象系统，尤其是输出供人观察的照片或屏幕显示这样一些图象系统，必须充分研究人的视觉系统，因为人的视觉系统才是这类图象系统的最后终端，即图象信息的信宿。而且此类系统输出图象最终总是由人的视觉系统来评价。另一方面，从某种意义上讲，人的视觉系统本身就是一个结构复杂、性能优越的图象系统。从仿生学角度出发，视觉原理、视觉特性以及视觉模型

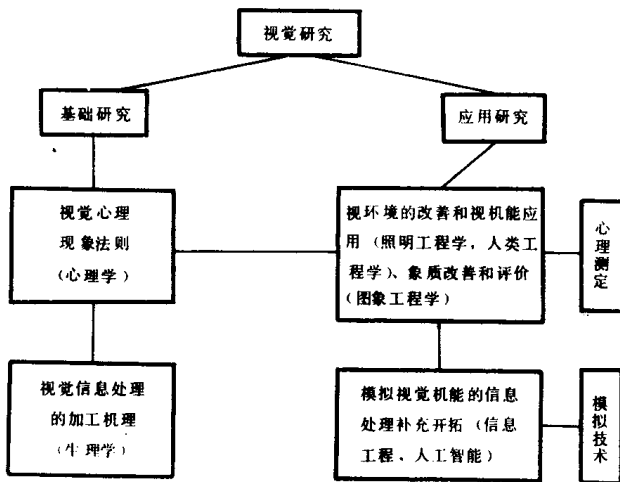


图 2-1-1 视觉研究与图象的关系

的研究，对图象工程技术人员来讲，是非常具有启发性和巨大吸引力的。

视觉研究可分为视觉生理、视觉特性、视觉模型等方面，也可分为视觉基础和视觉应用研究，它们与图象研究的关系见图 2-1-1。

近些年来，人们越来越清楚地认识到视觉系统在图象信息的产生、传输、处理方面具有十分优越的性能，许多方面是目前图象系统还不能具备的。因此，从事图象工程研究的学者不得不花费大量的精力对视觉基础和应用进行深入地研究。尽管时间不长，但也取得了不少有意义的成果〔2·1〕。

(1) 图象质量评价与视觉心理

目前把图象信息看成为二维平面上具有亮暗和色彩变化的若干个象素信息的集合是非常初步的认识。应该说，对图象认识或理解是由感觉和心理状态来决定的，也就是讲，这是与图象内容和观察者的心理因素有关。从图象信息传输角度出发，图象系统评价的真正尺度应该是发信者的意图为受信者所理解的程度，而不是对发信者发出的图象象素信息集合的简单接收。对物体的视觉包括了许多信息来源，这些信息来源超出了当我们注视一个物体时眼睛所接收的信息。它还包括由过去经验所产生的对物体的知识，这些经验不限于视觉，可能还包括其它感觉，如触觉、味觉、温度和痛觉等。由此启发人们去研究建立包括人的因素在内的信息理论。

(2) 画面组成和视觉心理

人的视野相当宽广，左右视角约为 180° ，上下约 60° （现在的电视画面约占 $7^\circ\sim 8^\circ$ ）。但如此宽广的视野中视力好的部位也仅限于 $2^\circ\sim 3^\circ$ 左右。那么，人是如何转动眼球使视线移动，从而适应大的画面和立体景象的呢？为此人们研究了画面取大时的心理效应、眼球运动的特性和任务、中心视和周边视的特性、立体视因素等问题，使得上面的问题已有个概括了解。例如，人眼中心视力分辨率强，可以进行图象细节的认识，但只能认识图象的一小部分；而周边视力分辨率差，但可认识图象的全貌，而且可以将所视目标特征部分

检出，利用检出的目标图象特征去控制眼球运动，必要时可以再用中心视力来进一步认识这一部分图象。对于大画面图象，充分利用周边视产生较强的临场感，而小画面临场感弱，为了产生充分的临场感，画面尺寸一般应有 30° 以上的视野。宽银幕和球幕电影的视觉效果好的原因也就是这个道理。

（3）视觉的时空频率分析

这是视觉研究较有成效的一个方面。许多研究表明，视觉的空间频率特性是影响图象锐度的主要因素。视觉神经具有的 mach 效应和 Roca-Sulzer 微分效应以及中枢神经的 Craik-obrien 积分效应等是视觉信息处理的基础。由此可以将视觉和图象结合起来研究，并应用于图象编码及象质改善技术中。

（4）视觉生理和模型的研究

视觉生理主要是指从视觉信息的产生部分——视细胞（图象信息感受器）和其它神经细胞以及大脑高级中枢的神经系统的信息产生、传输和处理的机理。这就涉及到神经生理学，它对图象工程技术是很有启发性的部分。由于能够插入单个细胞中的微小电极的出现以及数字处理技术的发展，已经能够对部分神经网络进行模型分析。如侧抑制现象，马赫效应等。但涉及到大脑高级神经中枢的“思考过程”，学习、联想，记忆等所谓自己组织化机能的研究，还是刚刚开始，预料今后可能有所突破，这将大大促进图象技术和人工智能的研究。

关于今后研究课题，不少学者认为应从以前的基础研究逐步转到和图象对应的视觉综合化、体系化的研究，可以大致归纳为以下几点：

- （1）搞清决定图象质量的主观因素，作出其总的结构模型。
- （2）找出人脑真正接收信息的容量，大脑有效接收图象的显示方式。
- （3）弄清图象信源和信宿的结构，建立起包括人的因素在内的信息论。

(4) 研究视觉和其它感觉的相乘作用, 即视觉和其它感觉的互相影响。

(5) 开发自己组织作用的综合研究, 进一步建立发展“思考过程”。

(6) 视觉和行为的联系。

由于不少生物、医学书籍对人眼结构、视觉一般特性已有介绍, 因此本章先介绍信息的产生、传递、处理的研究情况, 再对视觉运动特性、时空频率特性扼要地作一些介绍, 最后提出两种视觉模型。

2-1-2 视觉信息的产生、传递和处理

(一) 视觉信息的产生

当外界景象通过眼球的光学系统在视网膜上成象后, 视网膜就产生相应的生理电图象并经视神经传入大脑。在视网膜上如何产生生理电图象的问题, 经过多年的研究还没有肯定的结论, 但有一些倾向性的意见〔2·3〕, 扼要介绍如下: 视网膜可分为若干层, 各层对视觉信息的产生、加工、传递起着不同的作用。视觉信息的产生是在具有杆状细胞和锥状细胞的视细胞层。锥状细胞主要分布在视网膜的中央凹, 它的体积小, 但排列密度很高, 感光灵敏度低, 故形成具有高分辨率的有颜色感觉的明视觉。杆状细胞从中央凹开始向四周慢慢减少, 它的体积大, 感光灵敏度高, 形成具有高灵敏度的无色觉功能的暗视觉。视细胞外层称为光感受器, 其中含有成千上万个的小圆盘, 实验中发现这些小圆盘中含有大量视色素, 不同强度、不同波长的光照使不同数量、不同种类的视色素分解。无光照时在酶的作用下又重新合成。在这种分解和合成的过程中, 不但引起细胞内热能变化, 也引起其电能变化, 这种光—化—电过程就是视觉信息形成的实质。在这反应过程中还产生某种载体, 它将视觉电信息送出视细胞。对于实验中光能转换为生理电能的假说很多, 其中以离子学说目前最为流行, 它认为视色素吸收光子后使视细胞产生被动的阳离子流, 进而兴奋双极细胞, 再传到视神经细胞。

无光照时靠 Na-KATP 酶的作用，恢复视细胞的离子浓度梯度。

关于色觉理论主要有两种，一种是三色学说，认为人眼中有三种视色素，对光的响应峰值分别在红、绿、蓝区，由此综合形成色觉信息。另一种叫四色学说，它认为有三对视色素（白—黑，红—绿，黄—蓝），它们的组合产生色觉信息。

（二）视觉信息传递

视觉信息传递分眼球内和眼球外两部分。

1. 视觉信息在眼球内的传递

视觉信息在视细胞内形成后传给双极细胞，继而再传给视神经细胞，在这两种细胞之间还插入有双向输出、输入的水平细胞和无足细胞，起着双向反馈调节作用。视细胞送出的是与光照有关的连续电位，双极细胞和水平细胞的输入输出也是连续电位，但形式不同。无足细胞输出和输入则为瞬变性的连续电位，而视神经细胞输出却是具有“有”、“无”两种状态的脉冲电位。脉冲的幅度一定，但频率则随所受光刺激的强度变化而变化。视神经细胞以后的视觉信息都变成了电脉冲的形式。在大脑中枢神经中以积分作用再将传来的密度调制脉冲信号变为模拟信号。

2. 视觉信息在眼球外的传递

视觉信息由眼球经视神经送出后，经视神经交叉处分为两路传到两个外侧膝状体，再经视觉放射到达大脑视区（如 17 区）。17 区左右两个半球各自与对侧视网膜呈投射关系。许多实验表明，视网膜上各点和大脑视区各点保持着拓扑对应关系。

3. 视觉信息的加工处理

人和高级动物都有高度发达而完善的检测、识别、记忆视觉信息的能力。这里我们先介绍生物神经系统内普遍存在的侧抑制现象，即刺激某一个神经元使其兴奋，若再刺激该神经元附近的其它神经元时，可以看出后者的兴奋对前者的兴奋有抑制作用。侧抑制的强弱与神经之间距离、兴奋程度等有关。由于视神经的侧抑制作用而产生的视觉马赫现象，有增强图象轮廓提高图象反差作用。Glezer

〔2·1〕认为外侧膝状体中的I细胞对其传递信息的P细胞起反馈抑制作用，而I细胞是受大脑皮层来的离心神经纤维控制，因此人可以主观地控制视觉信息传入大脑视区。在大脑视区中有简单型细胞，作用是检测图象的明暗边界和线条的位置及方向。其中的复杂型细胞对直线和边界的运动方向是敏感的。还有低级超复杂型细胞作用是检测直线和边界的长度、宽度。而高级超复杂型细胞可以检测曲线和边缘曲度等。

Maffei〔2·1〕等从空间频率特性角度研究结果认为，外侧膝状体与神经节细胞都呈低通特性，而视区简单型细胞呈非常窄的带通特性，复杂型细胞则呈带宽稍宽的带通特性。超复杂型细胞对特定空间频率以颞颞（antagonize）形式组成感受野，因此可能对检测纹理边界起作用。另外大脑视区双眼细胞，通过对双眼视差信息提取起着检测距离和深度立体感觉作用。其它感觉信息（如听觉、触觉）和视觉信息相乘作用（互相影响）也发生在大脑视区，关于大脑对颜色信息处理倾向性意见〔2·1〕认为由视4区来承担，与空间信息是分开加工的。对于大脑视区以外部分的视觉信息处理目前还只处于问题的提出阶段。也有一些高级动物试验结果，推测人类也大体相同。如认为上颞沟后面区域中神经元对运动图象检测方向、形状和距离。颞下叶主要与视觉辨认学习有关。这部分如果损坏将不能辨认图形差异，想象不出几何图形，不能辨认容貌。上丘区域根据大脑区信息控制眼球运动等。

关于大脑对图形认识过程，目前在视觉心理上了解很少，但已提出了不少设想，并做了模拟试验〔2·1〕，这里介绍一下大脑全息记忆型设想。Lashley曾切除已学会走迷宫的鼠的大脑皮层，若只切除少部分，并不影响记忆。因此有人推测大脑记忆不同于计算机中的磁芯存储器。就是说一种现象不能用一种神经元状态来记忆，而是分散在许多神经元的不同状态中，和光学全息颇相似，可以从部分全息记忆中恢复整个图象。这些均有待于实验的证实和深入的理论探讨。总之，倾向性意见认为：当我们观察物体时，眼睛将外

界物象编码成神经活动信息送入大脑，这些神经活动借助于神经密码和大脑活动模式，代表着外界条件，而绝没有在脑中形成什么内部图象。

2-1-3 视觉特性

(一) 视觉运动特性

1. 眼球运动及控制

由中央凹中视锥细胞构成的分辨率高的集中视力一般只有 $2^{\circ}\sim 3^{\circ}$ ，而视网膜周围由视杆细胞构成的周边视力分辨率低，不能看清图象细节，但周边视力对图象中运动变化部分很灵敏，有特征抽取作用从而控制眼肌转动视轴，使集中视力对准这些部位，以便看清其细节。但是光靠这种转动是无法形成整个画面的印象，必须依靠周边视力。例如一个很大的物体，放在离眼睛很近的地方，尽管我们转动，仍不能看清物体全貌〔2·6〕。

根据实验〔2·7〕表明，眼球运动有两种方式。一种是按正弦形式进行的平稳追踪运动 (smooth pursuit movement)。另一种是以跳跃形式进行的单牵 (flick) 运动。平稳追踪运动是使视轴运动与物体运动一致，速度约为 (25~30) 度/秒。而单牵运动则起到使中央凹上的目标维持在原来的位置上，好象自动跟踪系统取样值控制似的。跟踪运动和单牵运动是左右眼视轴同方向移动。若两眼观察由近而远或由远而近的物体运动时，则两眼视轴向相反方向运动，产生改变两眼视轴夹角的所谓收敛运动 (Convergence movement)。

另外眼球还有一种无意识的象噪声似的微小运动，称为非随意运动〔2·8〕，包括颤动 (tremor)、单牵运动、漂移运动 (drift)，如果没有这些微小运动，人眼将无法看见静止图象。有人曾做过使非随意运动停止的试验，结果图象在刚刚开始时可以看见，但很快就模糊不清，这是因为视觉系统对不随时间改变的刺激很快适应所致。但平时观察静止图象时，由于眼球不断运动，在视网膜上的图

象转变成随时间变化的活动图象,也就不会由于完全适应而看不见。

2. 注视点的分布

人在观察景象时注视点喜欢集中在什么地方呢?有人〔2·9〕曾用电视眼球标记摄像机组 (television eye-marker camera) 进行试验 (见图 2-1-2)。

当光照到眼球的角膜,用凸透镜收集其反射光,则能把光源实象聚集成小光点,当注视点移动时,由眼球反射出来的小光点随之运动。用一个摄像机将此光点拍出并和从被试者正在看的画面所获得的视频信号在混频器中混合,这样从混频器输出端所接的监视器上就可以既看到画面图象,也看到被试者的注视点,即眼球标记光点移动情况。实验得出:

(1) 注视点主要集中在图象黑白交界的部分,尤其集中在拐角处。

(2) 用闭合的图形进行试验,注视点容易往图形内侧移动。

(3) 注视点容易集中在时隐时现、运动变化的部分。

(4) 图象中若存在一些特别的不规则处,则也是注视点容易集中的地方。

3. 眼球运动与外界不动性

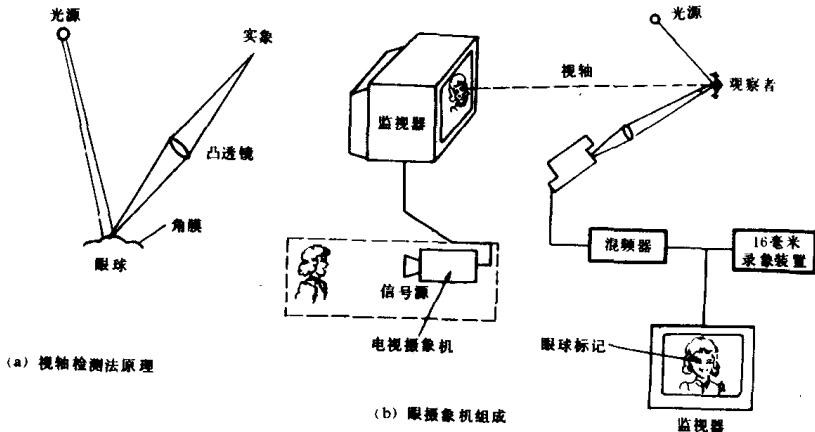


图 2-1-2 电视眼球标记摄像机

当我们在看一幅静止图象时，根据前面介绍，眼球始终在不停地运动。也就是讲视网膜和图象之间一直存在着相对运动，但是我们还是会感觉到图象是不动的，这是为什么呢？显然用我们通常的概念认为视觉系统相当于一个摄像系统是解释不了的。因为如果摄像象机相对于景物不断移动，输出视频信号一定不断改变，得不到一幅静止图象。这一机理虽未完全弄清楚，但一般认为〔2·10〕：在视觉中枢，由眼球运动控制中枢离心性送来与眼球有关的信息，它对从视网膜送到视觉中枢向心性的视觉信号伴随眼球运动而产生的变化，起预测、修正和抵消的作用。起这一作用的离心性信息叫做推理放电（Corollary discharge）。

如果眼球运动不是由脑指挥的，而是被迫的，那么就会感觉图象也是动的。假如我们坐在火车里，向窗外发呆看景物，即大脑不对眼球送出指令性离心信息，那么就会感到外面景物是动的。一旦认真看外面景物时，就会感到外面景物是不动的。显然这是与心理学有很大关系的。

4. 运动视觉的其它特性

在视野范围内物体运动，就会有速度感觉，但相同运动速度的物体，给人的速度感觉大小与物体的大小形状有关。如大的运动物体看上去慢，小的运动物体看上去快，与运动方向水平的棒状物体看上去快，而与之垂直的看上去慢。

还有一种假运动现象，即当实际静止的几个物体时隐时现时，看上去好象物体在运动，如图 2-1-3 所示，有 A、B 两个光源相互距离为 d ，最初 A 消失，经过 t 时后 B 消失，当时间 t 较大时，看上去 A 先消失，而 B 后消失；若 t 很小，看上去 A、B 同时消失，都没有运动感觉。但当 t 和 d 合适时，

看上去就会感到 A 向 B 运动，这种运动就是假运动也称 β 运动。 β 运动与 d 、 t 和光源的光强 I 有关，设 ϕ 为 β 运动感觉，则

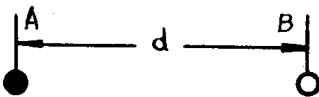


图 2-1-3 假运动现象

$$\varphi = f(d/It)$$

可见 d 、 I 和 t 有互相补偿关系。称此为 korte 法则。

(二) 视觉空间频率特性

我们平时都会感到：当快速运动物体从眼前通过时，是很难看清其细节而只能有个粗略轮廓。只有当物体细节大小、明暗对比度以及在眼中呈现时间长短都比较合适时，才能对物体细节有个清楚的认识。这样一种视觉特性可以用视觉空间频率特性或视觉时间频率特性来定量描述。视觉系统的空间和时间频率特性是互相依赖的。对亮度在空间作正弦变化的正弦光栅，求出人眼对各种节距正弦显示的图案的认识程度，就可以得到视觉空间频率特性。

亮暗在空间上作正弦变化条纹图案的物理对比度和感觉对比度之比随空间频率变化而变化的曲线，称为调制传递函数 MTF (modulation transfer function)，显然 MTF 与测试条件等许多因素有关。对这个问题，多年来许多学者进行了大量实验研究 (2·11、2·12、2·13)，一般情况下典型 MTF 曲线具有带通滤波特性，其峰值在 3 周/度 (即每度视角有三个周期的正弦条纹) 附近。见图 2-1-4。

F. W. Campbell (2·14) 在 1974 年发表“空间信息通过视觉系统的传输”一文。他认为视觉系统是呈多频道性，每频道对应不同的空间频率调制，且各频道带宽不超过一个倍频程。形成视觉系统的这种带通特性，并不是眼球的屈光系统，而是视网膜。其中视网膜中 Y 型神经节细胞对空间频率呈低通特性；而 X 型神经节细胞也呈低通性，但低频部分下降很快，高频部分延伸，截止频率提高。大脑视区中简单细胞是呈窄带通，复杂细胞呈宽带通特性，综合成为图 2-1-4 曲线。

视觉系统的空间频率特性 MTF 受时间因素的影响有：

闪烁：画面的空间平均亮度保持不变，周期性地改变线栅的亮暗就会产生闪烁。实验表明，随闪烁的频率增加 MTF 由带通变为低通，见图 2-1-5。

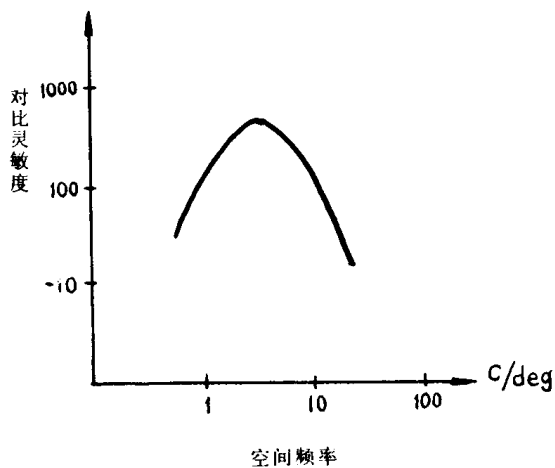


图 2-1-4 典型 MTF 曲线

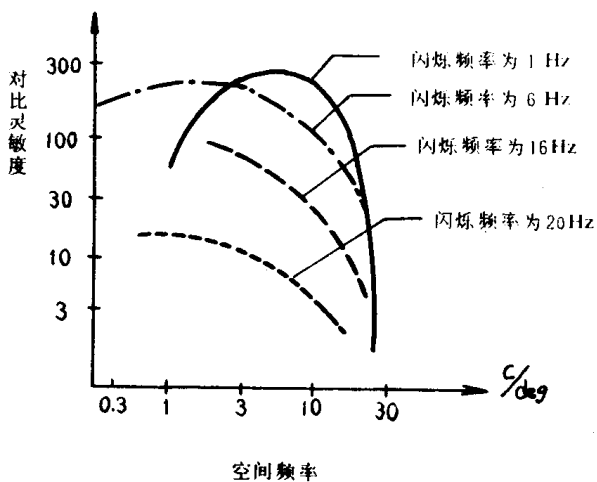


图 2-1-5 闪烁对 MTF 影响

漂移：漂移对 MTF 的影响，定性地讲和闪烁相似。

呈现时间：呈现时间变短，MTF 由带通变为低通。

运动的影响：线栅和视网膜相对运动，随速度增加（约大于 5 周/度）时，MTF 由带通变为低通。

其它还有图象的平均亮度、光栅的颜色等都对 MTF 有一定影响。

（三）视觉时间频率特性

若在视野内将亮度作正弦变化，可以得出与空间频率特性相似的时间频率特性，也是带通特性，见图 2-1-6。

2-1-4 视觉模型

将视觉系统的功能抽象化为简单的模型。并以此模型为基础来对视觉系统的功能进行研究，或工程上加以模仿是很吸引人的。建立模型首先要提出假说，根据假说建立模型，再根据试验修改模型，反复进行使之逐渐完善合理。建立模型并不一定要把视觉系统全部性质机能都模仿出来，针对模型的应用目的只对感兴趣的某种或某

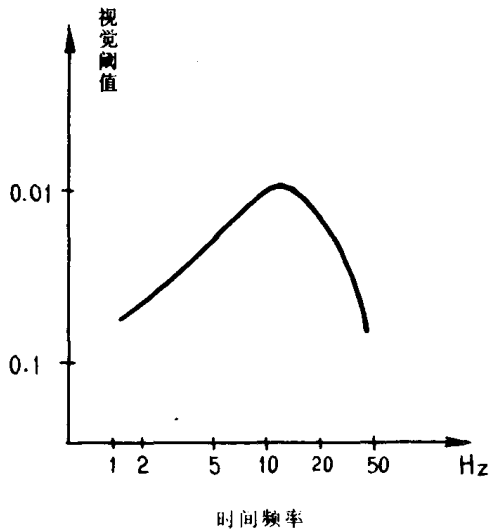


图 2-1-6 视觉时间频率特性