

中国光学文献数据库系统

李士范 赵桂云 于晓光*

摘要: 中国光学文献数据库 (COA) 是一个实用的计算机情报检索系统, 包括1985年以来中国光学文献2.3万条。利用MICRO-CDS/ISIS软件实现建库。用户可通过分类、标题、作者、单位、文献源、年限、主题词等12种途径实现检索, 还可以实现ISO-2709标准格式数据转换。系统全部采用菜单操作。介绍了COA的主要内容和功能, 描述了它的数据结构及其实现方法。最后介绍了它的检索方法。

一、系统概述

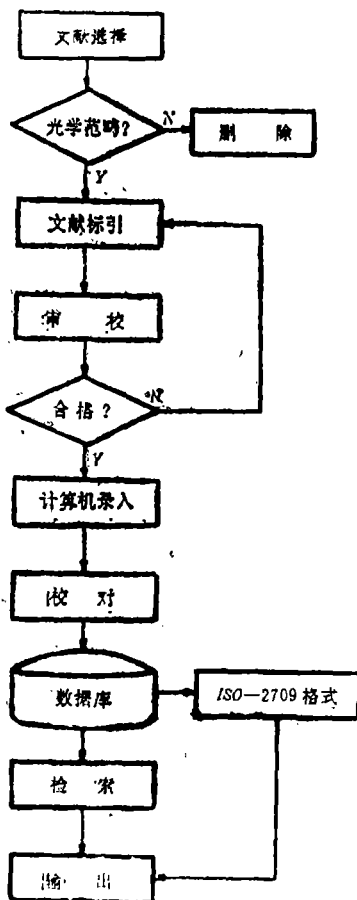


图1 COA建库流程图

中国光学文献数据库 (Chinese Optical Abstracts Database) COA是一个实用的计算机存贮与检索系统。现累积记录 2.3万条 (1985年以来), 文献增长率 4000条/年, 全部含有文摘, 题录部分为中英文对照。初步实现了对国产光学文献的全面采集, 处理, 计算机建库和检索。用户可通过分类、类型、标题、作者、单位、文献源、主题词 (包括题内关键词), 年限等多种途径实现精确词检索、逻辑检索、右截断检索、字段限定检索、相邻检索、ANY检索、全文检索等, 还可在屏幕上显示检索字典, 在词典列表上任意选择检索项和组合检索式。对检索结果, 可进行屏幕浏览和打印输出。

COA收录中国源光学文献有科技期刊, 会议录, 学位论文, 专利, 图书等。囊括了光学, 光电子学及其应用领域。专业范围包括: 光学概论、光度学与色度学、辐射与发光、几何光学、物理光学、光谱学与仪器、激光物理、器件与应用、非线性光学、信息光学、集成光学、纤维光学、薄膜光学, 摄影光学、光电器件与技术、光通信、光学测试与仪器、光学材料与工艺、光学仪器与设备、光学应用等。

系统使用国际上较为先进的通用微机检索软件 MICRO-CDS/ISIS, 结合光学文献数据库的特点开发而成。全部采用菜单操作, 可方便地实现数据录入, 增删、更新、倒排、打印和原始数据的ISO-2709标准格式输出等功能。系统建立于AST/P286微机上。

*注: 参加此项工作的还有冯丽杰、邵亚荣。

原始数据的收集,充分发挥了中科院在我国光学领域的优势,采用分散收集和处理,集中录入、建库与服务的方式。数据处理采用“光学叙词表”的受控标引和使用统一的“光学文献标引手册”。文献前处理进行了严格的质量控制,建立了完善的审校制度和严格的管理网络,从而保证了系统的整体质量。

二、数据库结构与系统建立

COA是在MICRO-CDS/ISIS软件的基础上开发而成的。该软件系统功能包括:

一个总控程序:

IS.BAT

批处理文件

IS.EXE

各功能模块的调用文件

四个用户程序:

ISISRET.EXE

数据库的检索程序

ISISUTL.EXE

数据库的录入程序

ISISPRT.EXE

数据库的输出打印程序

ISISINV.EXE

数据库的索引文件维护程序

四个系统程序:

ISISDEF.EXE

数据库结构的定义与维护程序

ISISUTL.EXE

数据库的实用程序

ISISXCH.EXE

数据库主文件维护程序

ISISPAS.EXE

高级语言接口程序

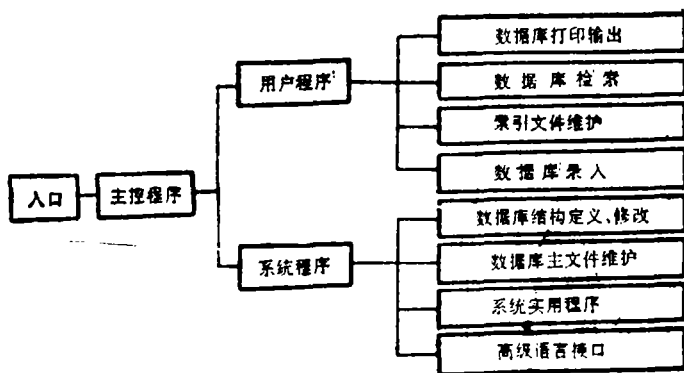


图2 程序结构框图

中国光学文献数据库系统建立

要建立COA检索系统,必须首先在CDS/ISIS中建立四个系统文件。

- * COA.FDT COA字段定义表
- * COA.FST COA字段选择表
- * COA.FMT COA数据录入工作单
- * COA.PFT COA显示格式

1. COA字段定义表 (COA.FDT):

COA.FDT包括COA中的全部数据项,是对所有这些项目的预定义。

表1 COA字段定义表

标识	名称	长度	类型	重复	分界符/模式
1	顺序号	6	N		
89	分类范畴	10	X	R	
83	中图法分类	24	X	R	
41	语种	1	A		
42	类型	1	A		
43	密级	1	N		
245	中文标题	200	X		
248	英文标题	200	X		
100	著者与单位	220	X	R	abc
101	著者与单位	300	X	R	abc
216	刊名与副刊名	60	X		
243	英译名	100	X		
247	A年B卷C期	30	X		abc
244	页码	40	X		
37	馆藏单位	40	X		
520	文摘	800	X		
650	主题词	220	X	R	
652	自由词	140	X	R	
653	主题词(英文)	250	X	R	
654	自由词(英文)	150	X	R	
111	会议名称	50	X		
112	英译名	100	X		
113	地点	20	X		
114	届次	10	X		
115	主办单位	50	X		
116	日期	20	X		
117	页次	20	X		
502	学位论文机构名称	50	X		
503	学位	1	X		
504	提出日期	20	X		
505	论文页码	20	X		
506	论文号码	20	X		
251	出版者	30	X		
252	出版地	10	X		
20	ISBN	25	X		
253	出版日期	15	X		
254	版次	2	X		
255	开本	4	X		
256	页码	10	X		
350	单价	10	X		
450	专利号	20	X		
453	国际分类号	40	X		
456	发行国	10	X		
452	申请日	20	X		
455	授权日	20	X		
457	代理人	10	X		
458	专利权人(申请人)	30	X		
451	设计人	20	X		
22	ISSN	10	X		
2	记录生成日期	10	X		9999-99-99
655	文摘员	10	X		
521	英文文摘	800	X		
522	作者单位	40	X		

2. COA数据库录入工作单 (COA.FMT)

COA根据不同类型的文献, 设置了五种不同类型的工作单, 以实现COA全部数据的录入。它们是:

COA.FMT	期刊录入工作单 (缺省)
HYL.FMT	会议录录入工作单
XWKW.FMT	学位论文录入工作单
ZL.FMT	专利录入工作单
TS.FMT	图书录入工作单

工作单格式:

中国光学文献数据库工作单 (COA)

顺序号:	分类范畴:	中图法分类:
语种:	类型:	密级:
中文标题:		
英文标题:		
著者与单位:		
著者与单位 (英):		
刊名与副刊名:		
英译名:		
A年B卷C期:		
页码:		
馆藏单位:		
文摘:		
主题词:		
自由词:		
主题词 (英):		
自由词 (英):		
记录生成日期:		文摘员:

3. COA数据显示格式 (COA.PFT)

COA根据CDS/ISIS所提供的格式语言编写了COA数据显示格式, 此格式可以随机修改。

显示格式如下:

mfn' _____
 _____', /"顺序号:"v001, /"分类范畴:"d089, (|/| + v089), x5, "专利号:"v450, x20, "馆藏单位:"v37, mhl, /v245, " /"d100, (|, | + v100 ^ a, |(|v100 ^ b|)|), d100 ^ b, d100 ^ c, v100 ^ c, " /"v111, ". -"v113, " /"v116, " -"v117, " /"v505, " /"v246, " -"v247 ^ a, " , "v247 ^ b, " ("v247 ^ c") "247 ^ c" . -"v244, " /"v451, " . -"v452, " . -"v455, " . -"v458, " /"v25 , " . -"v252, " . -"v253, d" . -"v256, /v520(4, 0)%#, /

在此种显示格式下的记录输出:

000662

顺序号: 870662

分类范畴: 13D

馆藏单位: SH07

光学薄膜的宽带膜厚监控装置 / 杨本祺, 许绍基, 金林法 (中科院上海光机所) // 光学学报, - 1987, 7 (7). - 613~617

装置是由配有微机的快速扫描分光光度计组成, 在镀膜时不断测量膜层的光谱透射率曲线以监控膜厚并作了监控三层膜的试验。装置的性能: 测量波长宽度333nm; 波长点数: 50点; 透射率精度: ±1%; 光谱透射率扫描速度: 12c/s。图5参5

4. COA字段选择表 (COA.FST)

COA根据用户需求, 对需要的检索字段, 列入字段选择表, 此表可进行随时修改。

字段选择表实例:

标识	抽词技术	数据抽取格式
100	1	mpu, (/v100 ^ a)
100	1	mpu, (/v100 ^ b)
100	1	mpu, (/v100 ^ c)
246	0	mpu, (/v246)
247	1	mpu, (/v247 ^ a)
652	0	mpu, (/v652)
650	0	mpu, (/v650)
89	0	mpu, (/v089)
1	0	mpu, (/v001)

上述四个表建立, 系统予以确认, COA即可实现建库。

三、检 索

COA可以检索1985年以来的全部中国源光学文献, 包括科技期刊、图书、专利、会议录、学位论文等。用户可以用中英两种文字, 用多种检索途径, 诸如: 分类、文献类型、标题、作者、单位、文献源、年限、主题词等, 检索到自己所需要的文献。

1. 精确词检索

该检索手段是指使用数据库中定义的任何一个检索元素, 它们可能是: 主题词、关键词、标题、作者等。例如: 染料激光器、韩昌元、12A (分类号)、光学机械(刊名)等均为精确检索词。精确检索词要求使用的检索词和计算机内的词形完全一致。

2. 右截断检索

该检索手段允许用户使用检索词的左半部分实现检索, 在所出现词头的后面加上字符

“\$”。该检索手段将自动地在带有该词头的检索词之间进行逻辑或 (OR) 运算。例如:

检索表达式:“气体\$”,可检索出含有: 气体发光、气体放电、气体激光器、气体分析等内容的全部文献。

3. ANY检索

ANY词是一个集合项,代表一组预先定义的检索词,当使用ANY检索时,系统将自动地把所有和这个ANY词有关的词进行逻辑或 (OR) 运算。例如:

检索表达式:“ANY长春光机所”,则凡是含有与之同义的词如:中科院长春光机所、中科院长春光学精密机械研究所、中国科学院长春光机所等检索词的文献均命中。

4. 逻辑检索

逻辑检索包括逻辑与 (AND)、逻辑或 (OR)、逻辑非 (NOT),或者三种逻辑算符的任意组配均可检索,逻辑检索可提高文献查准率和查全率。例如:

检索表达式:

(光学设计 + 镜头数据库) * 翁志成

则可检索到由翁志成所著有关“光学设计”和“镜头数据库”的全部文献。

5. 全文检索

全文检索包括数据库中的任意字段,只要该文献的著录项中含有需要检索的检索词(该词可在任何字段中),则此篇文献均可检出。

四、结 语

COA经过几年的努力,现已初具规模,今后需要稳步发展,长期坚持下去。不断改善性能和增加数据量。提高数据库质量,搞好检索服务,为我国科学技术的发展做出贡献。COA系统的建立是中国科学院光学六所一厂共同努力和中科院数据库工程办公室的直接领导和支持的结果,并得到中科院文献中心的大力支持,在此,一并致谢。

Chinese Optical Abstracts Database Systems

Li Shifan Zhao Guiyun Yu Xiaoguang

Abstract

The Chinese Optical Abstracts (COA) is a practical computer information retrieval systems. It contained 23,000 records which included title, author source, classification, subject words etc. It has been established by MICROCDS/ISIS software on the AST/P286 computer. Users may find every document by 12 entrances, also may realize input and output through ISO-2709 standard data format. Systems are operated by menu. COA's contents, functions, data structure, and programs structure are described and discussed followed by retrieval methods.