

ICS 01.140.40  
CCS A 19



# 中华人民共和国新闻出版行业标准

CY/T 270—2023

## 静态图像识别与检索技术规则

Technical rules for still image recognition and retrieval

行业标准信息服务平台

2023 - 06 - 16 发布

2023 - 08 - 01 实施

国家新闻出版署 发布

## 目 次

前言	III
1 范围	1
2 规范性引用文件	1
3 术语和定义	1
4 图像类型与质量指标	1
4.1 图像类型	1
4.2 图像质量指标	2
5 图像识别与指标	2
5.1 对象识别	2
5.2 识别过程	2
5.3 识别指标	3
6 图像检索与指标	4
6.1 检索过程	4
6.2 检索指标	5
附录 A (资料性) 图像识别与检索应用场景	6
A.1 图像识别应用场景	6
A.2 图像检索应用场景	7
参考文献	8

行业标准信息平台

## 前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由全国新闻出版标准化技术委员会（SAC/TC 527）提出并归口。

本文件起草单位：北京大学、中国建筑出版传媒有限公司、北京拓标卓越信息技术研究院、商务印书馆有限公司、人民卫生出版社有限公司、中国新闻出版研究院、北大方正信息产业集团有限公司。

本文件主要起草人：崔晓瑜、汪智、安秀敏、尹翔宇、陈永刚、杨继全、张倩影、周长岭、黄肖俊、吕肖庆、王勇涛、裴飞。

行业标准信息服务平台

# 静态图像识别与检索技术规则

## 1 范围

本文件规定了静态图像的类型与指标、图像识别与指标、图像检索与指标等。  
本文件适用于新闻出版及相关领域的图像处理与应用。

## 2 规范性引用文件

本文件没有规范性引用文件。

## 3 术语和定义

下列术语和定义适用于本文件。

### 3.1

**图像检索 image retrieval**

基于输入图像特征信息提取满足相似度阈值图像的技术。

### 3.2

**图像识别 image recognition**

基于图像处理、分析和理解，区分、定位、提取不同模式目标和对象并转换成所需结果的技术。

## 4 图像类型与质量指标

### 4.1 图像类型

#### 4.1.1 图像内容类型

按图像内容类型划分：

——字符图像；

——公式图像；

——表格图像；

——图案图像，主要包括：

- 实景图像，含人物、动植物和自然场景等；
- 工程效果图，含建筑效果图和产品效果图等；
- 绘画和设计图像，含美术创作、建筑设计图和电路设计图等；
- 医学影像（静态）；

——其他。

#### 4.1.2 图像色彩类型

按图像色彩类型划分：

——二值图像；

——灰度图像；

——彩色图像；

——其他。

## 4.2 图像质量指标

### 4.2.1 清晰度

图像清晰度用以描述图像上各细部影纹及其边界清晰的程度，可采用 DPI/PPI 表示（“DPI”是“每英寸点数 (Dots Per Inch)”，“PPI”是“每英寸像素数 (Pixels Per Inch)”， $1\text{in}=25.4\text{mm}$ ）。清晰图像的分辨率应不低于 300DPI。图像清晰度宜采用如下算法：

- a) 基于频率域特征的评价函数，如频谱函数；
- b) 基于统计特征的评价函数，如熵函数；
- b) 基于空间域特征的评价函数，如 TenenGrad 函数（特南梯度函数）、Brenner 函数（布伦纳梯度函数）、方差函数、平方梯度函数、Vollath 函数（沃拉特函数）、加窗梯度函数等。

### 4.2.2 干扰度

图像干扰度用以描述图像数据中不必要的或多余的噪声信息的程度。干扰度越大，图像质量越差，导致图像识别与检索结果指标越低。

### 4.2.3 规整度

图像规整度用以描述图像裁切、倾斜度和色差等规整的程度。规整度越高，越利于图像的识别与检索。

## 5 图像识别与指标

### 5.1 对象识别

#### 5.1.1 字符

字符图像识别用于自动定位及提取图像中的字符，通过图像识别技术进行字符识别，识别结果包括文字、数字、其他符号等。

#### 5.1.2 公式

公式图像识别用于自动定位及提取公式的符号、版式结构和语义信息，将识别结果转换为 LaTeX、MathML 等描述的结构化公式。

#### 5.1.3 表格

表格图像识别用于自动定位及提取表格的单元格结构和单元格内容信息，将识别结果还原成数字类型的表格，或结构化描述语言（如 html 定义表格结构的标签）描述的结构化表格。

#### 5.1.4 图案

图案图像识别用于从图像中识别出目标对象，比如人脸识别、商品识别、医疗影像的病灶识别等。

### 5.2 识别过程

#### 5.2.1 识别过程概述

图像识别由模型构建和识别执行过程两个环节组成，如图 1 所示。图像识别应用场景见 A.1。

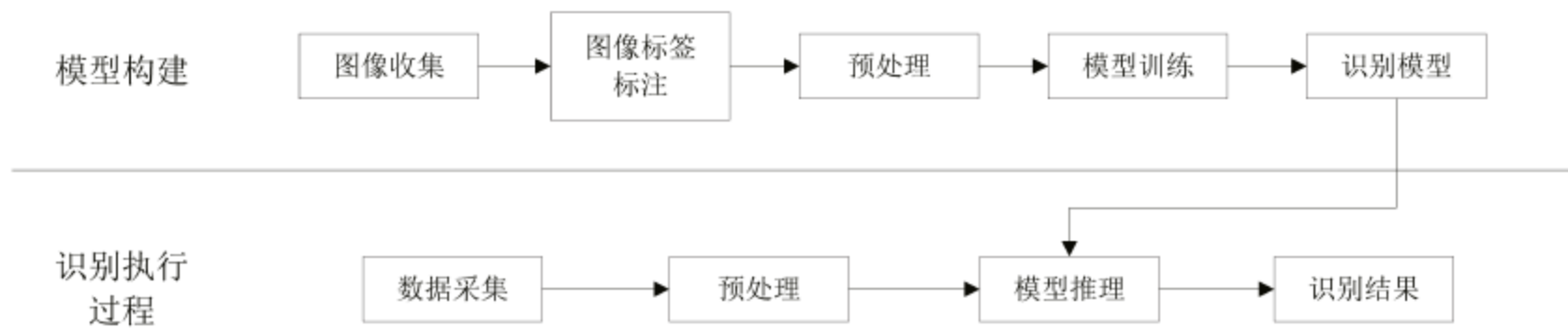


图1 图像识别过程

### 5.2.2 模型构建

模型构建过程宜包括：

- a) 图像收集：收集图像数据，用于图像识别的模型训练；
- b) 图像标签标注：在图像数据上添加一个或多个属性标签，用于后面的图像模型训练；
- c) 预处理：对图像进行灰度化、图像增强、滤波、归一化等处理以克服图像干扰因素；
- d) 模型训练：构建图像识别模型，训练模型参数，生成最优的图像识别模型；
- e) 识别模型：根据学习结果，输出最终生成的图像识别模型。

### 5.2.3 识别执行过程

识别执行过程宜包括：

- a) 数据采集：获取图像数据，通常借助于数字摄像机、扫描仪、数码相机等设备经过采样数字化得到的图像；
- b) 预处理：对图像进行灰度化、图像增强、滤波、归一化等处理以克服图像干扰因素；
- c) 模型推理：使用图像识别模型对图像数据进行推理分析并得到识别结果；
- d) 识别结果：输出图像识别结果。

## 5.3 识别指标

### 5.3.1 正确率

正确率是衡量识别系统质量优劣的重要指标之一，定义如下：

$$\text{Accuracy} = \frac{TP+TN}{P+N} \times 100\%$$

式中：

- Accuracy —— 正确率；  
 TP —— 正确识别的正样本数；  
 TN —— 正确识别的负样本数；  
 P —— 正样本数；  
 N —— 负样本数。

常用的是 Top- $k$  正确率，即对于图像类别预测标签的概率向量中前  $k$  个中包含物体正确的分类标签的正确率，如 Top-1 正确率和 Top-5 正确率。

### 5.3.2 误识率

误识率是图像被错误判别为某类图像的概率，定义如下：

$$\text{FAR} = \frac{N_{\text{FA}}}{N_{\text{IRA}}} \times 100\%$$

式中：

FAR——误识率；

$N_{FA}$ ——错误判别样本的次数；

$N_{IRA}$ ——类间样本判别总次数。

### 5.3.3 拒识率

拒识率是图像被错误拒绝识别为某类图像的概率，定义如下：

$$FRR = \frac{N_{FR}}{N_{GRA}} \times 100\%$$

式中：

FRR——拒识率；

$N_{FR}$ ——错误拒绝的次数；

$N_{GRA}$ ——类内样本判别总次数。

## 6 图像检索与指标

### 6.1 检索过程

#### 6.1.1 检索过程概述

图像检索过程由索引建立和检索执行两个环节组成，如图 2 所示。图像检索应用场景见 A. 2。

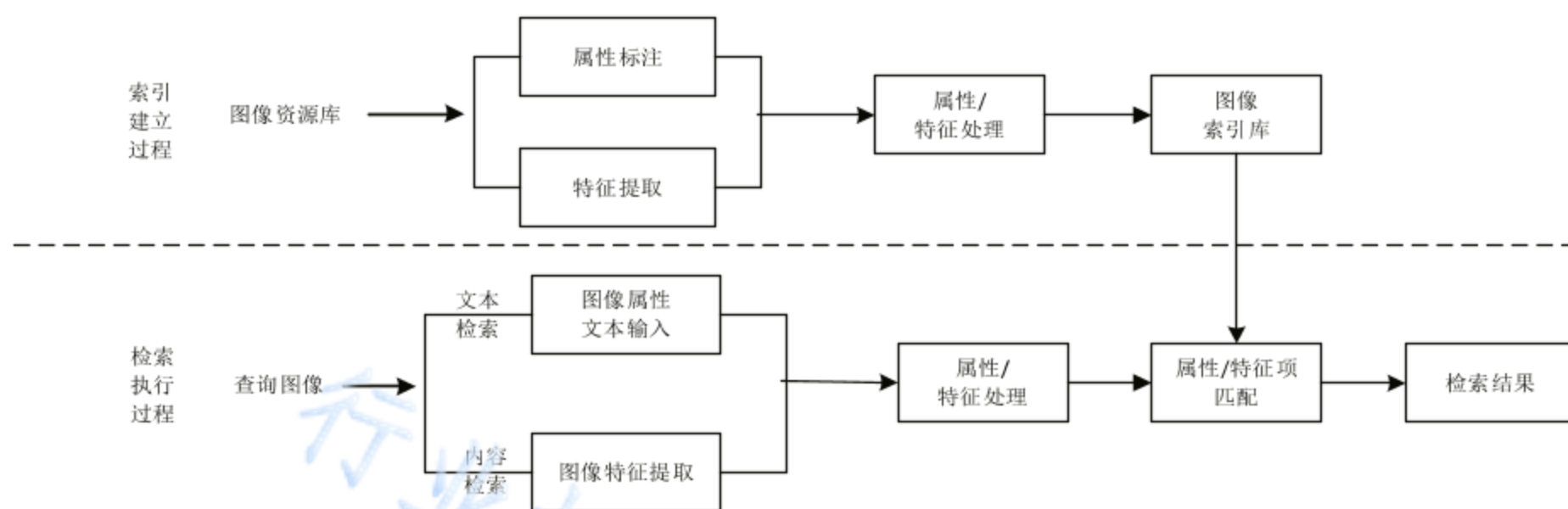


图 2 图像检索过程

#### 6.1.2 索引建立过程

索引建立过程宜包括：

- 属性标注：对图像资源库中的图像进行属性标注；
- 特征提取：对图像资源库中的图像提取图像特征数据；
- 属性 / 特征处理：抽取最能反映图像本质的属性或特征，进行编码与聚合等操作，生成待索引指标；
- 图像索引库：将待索引指标入库，建立图像索引库。

#### 6.1.3 检索执行过程

##### 6.1.3.1 检索模式

图像检索模式包括：

- 文本检索：通过文本的形式输入检索需求，获取图像检索结果；

- b) 内容检索：通过图像内容语义的形式输入检索需求，获取图像检索结果；
- c) 混合检索：通过文本和图像内容语义相结合的形式输入检索需求，获取图像检索结果。

注：采用内容检索与混合检索时对输入的图像内容语义，后续需进行图像特征提取操作。

### 6.1.3.2 执行过程

图像检索执行过程宜包括：

- a) 检索信息输入：按照上述图像检索模式，输入相应的检索信息；
- b) 属性 / 特征处理：使用索引库建立时的属性或特征处理方法对数据进行分析并得到其特征值；
- c) 属性 / 特征值匹配：将得到的特征值与图像索引库进行匹配操作；
- d) 检索结果：根据匹配规则，输出图像检索结果。

## 6.2 检索指标

### 6.2.1 查准率

查准率用以衡量图像检索结果的准确性，定义如下：

$$\text{Precision} = \frac{TP}{TP+FP} \times 100\%$$

式中：

Precision —— 查准率；

TP —— 检索到的相似图像个数；

FP —— 检索到的非相似图像个数。

### 6.2.2 查全率

查全率用以衡量图像检索结果的完整性，定义如下：

$$\text{Recall} = \frac{TP}{TP+FN} \times 100\%$$

式中：

Recall —— 查全率；

TP —— 检索到的相似图像个数；

FN —— 没有检索到的相似图像个数。

行业标准信息服务平台



附录 A  
(资料性)  
图像识别与检索应用场景

## A.1 图像识别应用场景

### A.1.1 数字资源加工

数字资源作为文献信息的一种表现形式是将计算机技术、通信技术及多媒体技术相互融合而形成的以数字形式发布、存取、利用的信息资源综合。图像数字资源加工是运用图像扫描与处理、图像识别以及数字化信息的各种再加工技术，将大量已存在的，以不同形式和载体存储的信息资源转化为能够再利用的数字化信息。在图像数字资源加工过程中，图像文字识别技术是主要使用的核心技术之一。

图像数字资源加工主要处理步骤如图 A.1 所示。

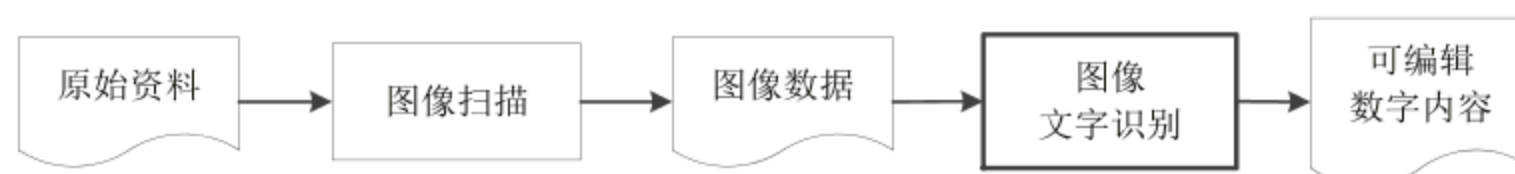


图 A.1 数字资源加工主要处理步骤

图像数字资源加工主要处理步骤描述：

- a) 通过扫描设备对待加工资源进行扫描处理，为下一步图像识别做准备；
- b) 对图像进行文字识别、校对等操作，获得准确的文字编码数据以及字体、字号、标题级别等版式信息；
- c) 经过质量检验的结果数据能够为后续应用提供可编辑数字内容资源。

图像文字识别技术融合高速扫描和全文检索技术，能将数以万计的纸质文件、手稿、档案、资料和文书等快速、自动地转换成数字化文本和影像，并实现海量信息高速检索和查询。数字资源加工涉及领域包括图书馆、出版社、杂志社、报社、档案馆、银行、保险、财会及大型企事业单位等。

### A.1.2 实体对象识别

实体对象识别是对三维世界的客体及环境的感知和认识，运用图像处理与识别技术，结合人工智能、系统学等学科的研究成果，广泛应用在各种行业应用中，例如人脸识别、病灶识别、建材清点、场景检测、植物识别等。实体对象识别的评价指标常用的有鲁棒性、正确性、效率和范围。

实体对象识别主要处理步骤如图 A.2 所示。

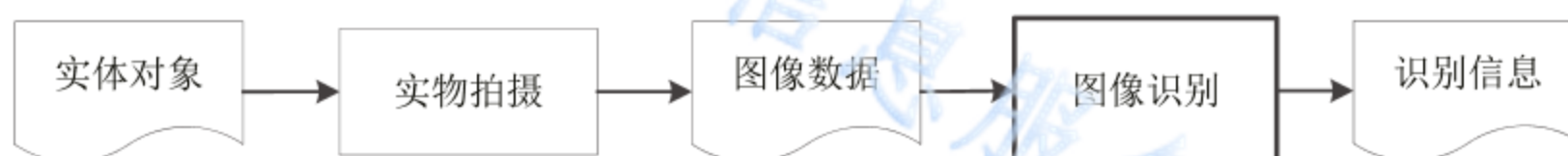


图 A.2 实体对象识别主要处理步骤

实体对象识别主要处理步骤描述：

- a) 拍摄待识别实体对象，将实体对象转换为数字图像信息；
- b) 利用图像识别技术识别出图像中有什么物体，并报告出这个物体在图像表示的场景中的位置和方向等相关信息；
- c) 输出识别结果。

## A.2 图像检索应用场景

### A.2.1 图像资源甄别

图像资源甄别是从海量图像资源中甄别出所需信息。在图像资源甄别应用中，基于内容的图像检索技术是运用的关键技术之一。

图像资源甄别主要处理步骤如图 A.3 所示。

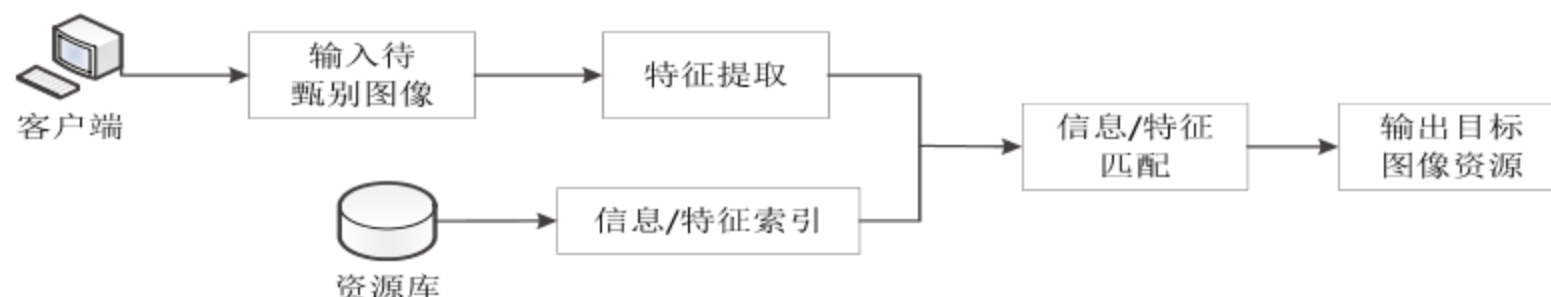


图 A.3 图像资源甄别主要处理步骤

图像资源甄别主要处理步骤描述：

- 用户输入图像形式的待甄别信息；
- 对待甄别信息进行预处理和特征提取等操作；
- 将处理后的待甄别信息与目标资源库的特征信息索引库进行匹配计算和比较；
- 根据匹配结果输出目标图像资源。

### A.2.2 论文查重

论文查重应用于对学位论文、学术论文等进行一致性检测，防止学术不端行为。早期论文查重只针对文本内容的比对，随着图像处理技术的发展，公式、分子式、图片等基于图像的比对检测在论文查重系统中也开始广泛应用。

论文查重主要处理步骤如图 A.4 所示。

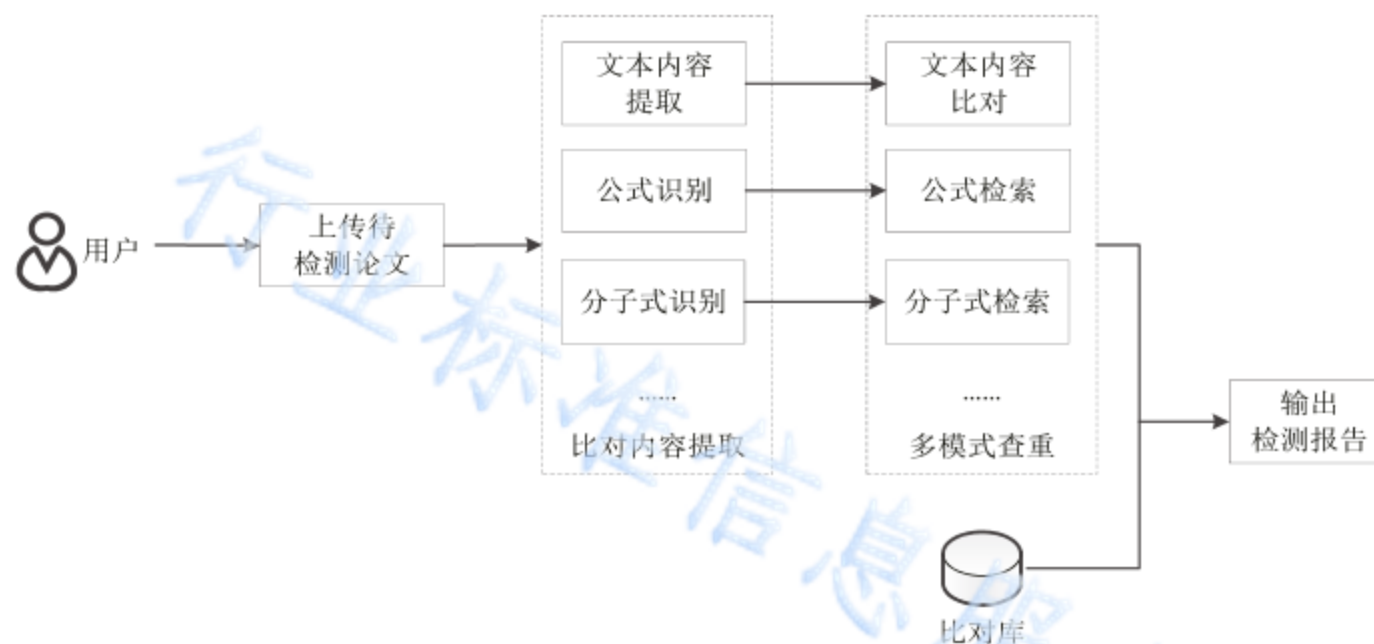


图 A.4 论文查重主要处理步骤

论文查重主要处理步骤描述：

- 用户上传待检测论文；
- 论文查重系统从待检测论文中提取出文本、公式、分子式等比对内容；
- 比对内容按照不同模式与比对库进行相应的比对与检索操作；
- 将不同模式下的比对结果进行汇总与整理，输出论文检测报告。

### 参考文献

- [1] GB/T 31219.3—2014 图书馆馆藏资源数字化加工规范 第3部分：图像资源
- [2] GB/T 35678—2017 公共安全 人脸识别应用 图像技术要求
- [3] ISO/IEC 15938-1:2002 Information technology—Multimedia content description interface—Part 1: Systems
- [4] ISO/IEC 15938-2:2002 Information technology—Multimedia content description interface—Part 2: Description definition language
- [5] ISO/IEC 15938-3:2002 Information technology—Multimedia content description interface—Part 3: Visual
- [6] 张春红, 唐勇, 邵珂. 图像资源数字加工标准及其应用 [J]. 现代图书情报技术 (京), 2010, 26 (12): 9-14.
- 

行业标准信息服务平台