

ICS 01.140
A14

DB32

江苏省地方标准

DB 32/T 1894—2011

档案数字化转换操作规程

Operation Procedures for digitization of records and archives

2011-09-01 发布

2011-11-01 实施

江苏省质量技术监督局

发布

目 次

前言	V
1 范围	1
2 规范性引用文件	1
3 术语与定义	1
4 总则	5
4.1 目标	5
4.2 原则	5
4.2.1 基本原则	5
4.2.2 数字化对象确定原则	5
4.2.3 守法原则	5
4.2.4 保密安全原则	5
4.2.5 价值性原则	5
4.2.6 质量优先原则	5
4.3 档案数字化转换对象的范围	5
4.3.1 基本范围	5
4.3.2 文书（管理）档案	6
4.3.3 基建科技档案	6
4.3.4 照片、录音录像、实物档案	6
4.4 基本环节	6
5 操作规程	6
5.1 纸质档案数字化转换	6
5.1.1 档案准备	6
5.1.2 目录数据库建设	7
5.1.3 全文数据采集	7
5.2 录音录像档案数字化转换	8
5.2.1 档案数字化前处理	8
5.2.2 数字化转换	9
5.2.3 音频视频数字文件存储	9
5.2.4 音视频数字文件处理	10
5.2.5 数据管理	10
5.3 照片底片档案数字化转换	11
5.3.1 档案准备	11
5.3.2 档案还原	11
5.3.3 目录建库	11
5.3.4 数据采集	11
5.3.5 图像修复	11

5.3.6	图像拼接	12
5.3.7	图像处理	12
5.3.8	图像存储	13
5.3.9	版权保护	13
5.4	缩微胶片档案数字化	13
5.4.1	数字化范围	13
5.4.2	检查整理	13
5.4.3	内容检查	13
5.4.4	胶片扫描	13
5.4.5	数据校验	15
5.5	实物档案数字化转换	15
5.5.1	档案整理	15
5.5.2	数字化准备	15
5.5.3	数据采集	15
5.5.4	图像存储	15
6	数字化文件存储	15
6.1	数字化文件存储级别和格式	16
6.2	图像文件的命名	16
6.3	数据关联	16
6.4	数据质检	16
6.5	数据挂接	16
6.5.1	数据挂接	17
6.5.2	检查挂接	17
6.5.3	挂接登记	17
7	数据验收	17
7.1	数据抽检	17
7.2	验收指标	17
7.3	验收登记	17
8	数据备份	17
8.1	备份范围	17
8.2	备份方式	18
8.3	备份数据校验	18
8.4	备份标签	18
9	数字化成果管理	18
10	安全保密要求	18
10.1	档案数字化承建机构	18
10.2	单位保密管理制度	18
10.3	人员管理	18
10.4	数字化加工设备的管理	19
10.5	加工场所的管理	19

10.6 档案安全	19
10.7 数据安全 管理	19
附录 A（规范性附录）档案数字化转换流程图	20
附录 B（资料性附录）档案数字转换用表	25

前 言

本标准依据GB/T 1.1-2009 《标准化工作导则 第1部分：标准的结构和编写》的要求进行编写。

本标准由江苏省档案局提出并归口。

本标准主要起草单位：江苏省档案局、江苏省标准化研究院。

本标准主要起草人：谢波、齐丽华、汪东华、钱耀明、杨红静、许建智、张姬雯、顾长青、吴晓、张宁芳。

档案数字化转换操作规程

1 范围

本标准规定了对纸质档案、录音录像档案、照片档案、缩微胶片档案、实物档案数字化转换的主要技术要求。

本标准适用于江苏省各类档案的数字化加工处理及数字化成果转换的操作要求与管理。

2 规范性引用文件

下列文件对于本文件的应用是必不可少的。凡是注日期的引用文件，仅所注日期的版本适用于本文件。凡是不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

- GB/T 6159.1-2003 缩微摄影技术 词汇 第1部分：一般术语
- GB/T 6159.4-2003 缩微摄影技术 词汇 第4部分：材料和包装物
- GB/T 6159.5-2000 缩微摄影技术 词汇 第5部分：影像的质量、可读性和检查
- GB/T 6159.7-2000 缩微摄影技术 词汇 第7部分：计算机缩微摄影技术
- GB/T 17235.1-1998 信息技术 连续色彩静态图像的数字压缩编码 第1部分：要求和指南
- GB/T 17235.2-1998 信息技术 连续色彩静态图像的数字压缩编码 第2部分：一致性测试
- GB/T 11821-2002 照片档案管理规范
- GB/T 18894-2002 电子文件归档与管理规范
- GB/T 20163-2006 中国档案机读目录格式
- GB/T 20225-2006 电子成像 词汇
- DA/T 15-1995 磁性载体档案管理与保护规范
- DA/T 18-1999 档案著录规则
- DA/T 31-2005 纸质档案数字化技术规范
- DA/T 38-2008 电子文件归档光盘技术要求和应用规范
- DA/T 43-2009 缩微胶片档案数字化技术规范
- DB32/ 505-2002 文书档案文件及目录数据库结构与数据交换格式

3 术语与定义

3.1

纸质档案 paper-based archives

国家机构、社会组织或个人在社会活动中直接形成的有价值的以纸质为载体的历史记录。

3.2

照片档案 photographic archives

国家机构、社会组织或个人在社会活动中直接形成的以静止摄影为主要反映方式的有保存价值的历史记录。照片档案一般包括底片、照片和说明三个部分。

3.3

录音、录像档案 audio and video tape archives

国家机构、社会组织或个人在社会活动中直接形成的记载在物理载体上的影像或声音为主要反映方式的有保存价值的历史记录。

3.4

缩微胶片 microfilm

一种能够记录缩微影像的透明而柔软的胶片，或者已摄有原件缩微影像的缩微品，一般分为16mm和35mm卷片和平片。

3.5

实物档案 object archives

在管理、生产、经营、科研及其他社会活动中形成的，具有保存价值的，按归档要求整理归档的实物。

3.6

数字化 digitization

利用计算机技术将模拟信息转换为数字信息的处理过程。

3.7

纸质档案、照片档案数字化 digitization of paper-based and photographic archives

采用扫描仪或数码相机等数码设备对纸质档案、照片档案进行数字加工，将其转化为存储在磁带、磁盘、光盘等载体上并能被计算机识别的数字图像或数字文本的处理过程。

3.8

录音、录像档案数字化 digitization of audio and video tape archives

采用数字化转换设备对录音、录像档案进行数字化转换，将影像、声音的模拟信息转换成可以被计算机存储和处理的数字信息的处理过程。

3.9

缩微胶片档案数字化 digitization of microfilm

用缩微影像扫描器等设备将缩微胶片上的影像转换成存储在磁盘、磁带、光盘等载体上并能被计算机识别的数字图像或数字文本的处理过程。

3.10

实物档案数字化 digitization of object archives

采用数码拍摄的方式，将档案实物影像记录成计算机识别的数字图像或数字文本的处理过程。

3.11

数字图像 digital image

表示实物图像的整数阵列。一个二维或更多维的采样并量化的函数，由相同维数的连续图像产生。在矩阵（或其他）网络上采样——连续函数，并在采样点将值最小化后的阵列。

3.12

黑白二值图像 binary image

只有黑白两级灰度的数字图像，它对应于黑白两种的文字稿，线条图等。

3.13

连续色调静态图像 continuous-tone still image

以多于两级灰度的不同浓淡层次或以不同颜色通道组合成的静态数字图像。

3.14

分辨率 resolution

单位长度内图像包含的点数或像素数，一般用每英寸点数（dpi）表示。

3.15

失真度 distortion

对档案进行数字化转换后，数字图像与档案原件在色彩、几何等方面的偏离程度。

3.16

可懂度 intelligibility

数字图像向人或机器提供信息的能力。

3.17

图像压缩 image compression

消除图像冗余或对图像近似的任一种过程。

3.18

密度 density

在缩微摄影技术中衡量感光材料曝光和显影后变黑程度的物理量，也可用透射密度（DT）来表示，定义为不透明度的基10对数。

3.19

对比度 contrast

表示影像中明暗区域相互关系，是影像中最亮与最暗的色调之间差异范围的指数。

3.20

曝光亮度 exposure brightness

决定数字图像明暗强度的指数。

3.21

采样率 sampling ratio

模拟信号采成数字信号时的取样频率。

3.22

量化级 quantifying level

描述声音波形的数据是多少倍的二进制数据，通常用bit做单位。

3.23

通道数 number of channels

是一个记录产生一个波形（单声道）或两个波形（双声道立体声）的信号数。

3.24

数码率 digital rate

系统在单位时间内传送的数据量，在实时传输情况下，数码率等于取样与量化比特数的乘积，数码率的单位通常为Mbit/s，即兆比特/秒。

3.25

电子文件类别代码 code of electronic records classification

区分电子文件类别的代码，可在著录和建立数据库时使用。T-文本文件，I-图像文件，G-图形文件，V-影像文件，A-声音文件，O-超媒体链接文件，P-程序文件，D-数据文件。

4 总则

4.1 目标

以先进的数字技术为依托，通过档案数字化的方式，将各类传统载体档案信息转换为各类数字档案信息，建设和整合各级综合档案馆、专业档案馆和档案室的档案信息资源，建立规范化和共享的各类档案数据库，开发利用档案信息资源，发挥档案的社会效益和经济效益，实现档案信息资源的社会共享，满足国家机构、社会组织或个人日益增长的档案信息需要。

4.2 原则

4.2.1 基本原则

档案数字化应遵循“开放性、抢救性、重点性、共享性”原则，使档案信息资源真实、准确、安全、方便、快捷共享和提供利用，确保实体档案的安全，实现电子档案备份。

4.2.2 数字化对象确定原则

应对所有要进行数字化的档案对象按照规定的原则和利用方法进行确认，只有内容和自身条件符合数字化转换基本要求的档案才能进行数字化。

4.2.3 守法原则

所有档案数字化转换活动应严格遵守国家、地方和行业相应的法律、法规制度，符合国家档案整理、开放等各项规定。

4.2.4 保密安全原则

档案数字化全过程应遵守保密安全要求，对数字化对象进行严格划控、鉴定和分级。涉及国家秘密或不宜在网上利用的档案一般不予数字化。档案数字化工程依法遵循档案行政管理部门备案制度。

4.2.5 价值性原则

属于归档和保藏范围，且应永久或长期保存的，或社会利用价值高的档案可列入数字化加工范围。可按照扫描鉴定从宽、发布鉴定从严的原则处理，尽量保持已有档案全宗或者目录数字化完整性。

4.2.6 质量优先原则

档案数字化的组织工作应由档案馆或档案室负责全过程管理，确保数字化成果标准可用。注重成本和利用实效。

4.3 档案数字化转换对象的范围

4.3.1 基本范围

档案数字化转换对象应是符合国家有关档案开放、档案抢救与保护规定，永久或定期保存，且社会利用价值高、涉及民生和民众的档案。

4.3.2 文书（管理）档案

各单位履行主要职能时制定或形成的各类文书和业务类文件、召开重要会议与举办重大活动等形成的主要文件及组织机构演变与从事任免表彰文件等材料。

4.3.3 基建科技档案

新建、扩建、改建和技术改造等基建项目总体规划、建设规划、项目建议书，可行性研究、评估报告、设计基础材料、设计研究文件，统计报表，各类合同、协议书、房产土地等证件，基建财务、水电气管道、通讯分布图、工程管理文件、施工竣工文件、生产、工艺、设备文件等。

4.3.4 照片、录音录像、实物档案

各单位拍摄录制有关本单位工作的或媒体拍摄录制的各类声像记录、报道、奖状等荣誉、书画或特色实物档案。

4.4 基本环节

档案数字化转换分为：档案准备、数据采集、数据存储和数据备份四个阶段。基本环节主要包括：档案前处理、档案数字化转换、图像处理、图像存储、目录建库、数据挂接、数据验收、数据备份、成果管理等。档案数字化转换操作流程见附录A，档案数字化转换用表格可参考附录B。

5 操作规程

5.1 纸质档案数字化转换

5.1.1 档案准备

档案准备包括档案领取、前处理和归卷环节。具体要求见表1。

表1 纸质档案数字化档案准备工作要求

项目	具体要求
规范目录数据	按照DA/T 18-1999要求，以全宗为单位，对照档案实体和目录，逐卷逐件进行检查核对、纠错与补漏，达到案卷级，文件级目录与档案实体内容三者的准确一致，确保文档与目录一一对应，发现错误应以档案实体内容和规范为依据，进行统一规范化修改、完善。没有进行划控开放鉴定的，必须逐件进行鉴定并进行开放或控制数据项著录。
检查项目	未编制张页号或编号有误的，应进行补编和更正，文件没有标题或标题不能揭示文件内容的，应重新拟制题名。对民国文体格式的档案进行著录时，题名应概括“事由”和“文种”两部分内容。对同一事由形成的若干份文件，编制组合条目。
区分扫描件和非扫描件	按要求把同一案卷中的扫描件和非扫描件区分开。普发性文件区分的原则是：重份档案取出或做出标记及不进行扫描的说明，有正式件的可以不扫描原稿。区分原件与复印件，只有复印件的可扫描复印件。
页面修整	破损严重、无法直接进行扫描的档案，应予以注明并进行技术修复，破损程度影响档案内容，单面档案进行托裱，双面有字档案进行丝网加固处理；严重折皱不平影响扫描质量的原件可先进行压平或低于50℃熨平等相应处理再进行扫描。
整理登记	制作并填写纸质档案数字化加工过程交接登记表，详细记录档案整理后每份文件的起始页号和页数。每一份文件的形式、数量等内容。根据纸张质量，区分使用高速、平板、无边距或书刊、工程扫描方式。
装订	扫描工作完成后，拆除过装订物的档案应按档案保管的要求重新装订，恢复装订时，应注意保持档案的排列顺序不变，做到安全、准确、无遗漏。

5.1.2 目录数据库建设

纸质档案数字化包括目录数据库和全文数据库建设两项内容。目录数据库是全文数据库建设的基础和前提，应准确、完备。目录数据可在数字化过程中以EXCEL形式从数字化软件中导出，并导入档案管理系统数据库。

5.1.2.1 目录数据结构

按照DB32/ 505-2002的要求，档号、顺序号、责任者、文号、文件题名、形成时间、密级、页号控制开放标识、备注等著录项齐全。字段长度和内容应符合要求。

5.1.2.2 目录数据交换格式

按照DB32/ 505-2002的规定执行，可选择DBF或TXT数据交换格式。

5.1.3 全文数据库采集

5.1.3.1 纸质档案数字化采集方式和技术要求

纸质档案按档案的幅面大小，选择相应规格的扫描仪或专业扫描设备进行扫描。大幅面档案可采用小幅面扫描后图像拼接方式处理。

纸质档案数字化采集方式和技术要求见表2。

表2 纸质档案数字化采集方式和技术要求

采集方式	档案质量	扫描色彩模式选择			分辨率	扫描仪规格
		黑白	灰度	彩色		
平板扫描仪 无边距扫描仪	纸张状况较差，过薄、过软或超厚的档案	档案页面为黑白两色，字迹清晰、不带插图	档案页面为黑白两色，字迹清晰度差或带有插图的档案，页面为多色文字	档案页面中有红头、印章或插有黑白照片、彩色照片、彩色插图	≥200dpi ≥300dpi (OCR 汉字识别)	根据档案幅面的大小(A4、A3、A0等)选择相应规格的扫描仪
高速扫描仪	纸张状况好的档案				≥200dpi ≥300dpi (OCR 汉字识别)	
工程扫描仪	工程图纸或大幅面照片、报纸等档案				≥600dpi	
书刊扫描仪	不可拆装的书籍刊物等档案材料				≥300dpi	

5.1.3.2 图像处理

图像处理应运用数字化设备的配套软件或专业的图像处理软件对采集的图像进行技术处理，以保存为目的的数字化产品应尽可能不做或少做图像处理，保持档案原件的特性。

纸质档案扫描图像的处理要求见表3。

表3 纸质档案扫描图像的处理要求

处理方式	作用
方向调整	对方向不正确的图像应进行旋转还原，以符合数字图像的浏览习惯。
纠偏	对出现偏斜的图像应进行纠偏处理。
去污	对图像页面中出现的影影响图像质量的杂质如黑点、黑线、黑框、黑边等应进行去污处理。处理过程中应遵循在不影响可读度的前提下展现档案原貌的原则。
图像拼接	对大幅面档案进行分区扫描形成的多幅图像，应进行拼接处理，合并为一个完整的图像，以保证档案数字化图像的整体性。对大幅面的档案进行分区扫描时，应使各分区的大小尺寸相近、分辨率和曝光量一致，相邻图像之间必须有部分重叠，重叠面积不少于各分区面积的1/20。扫描后形成的多幅图像，应进行拼接处理，合并为一个完整的图像，以保证档案数字图像的完整性。经图像拼接处理的照片档案应做好记录，并在著录目录数据库时，在附注项中加注“图像拼接”。

表 3 纸质档案扫描图像的处理要求（续）

处理方式	作用
裁边处理	对扫描或转换后的图像应进行裁边处理, 去除多余的白边, 有效地减少数字图像文件的容量, 节约存储空间。裁边后的图像大小保持一致。
图像修复	修复的效果以图像不失真、效果清晰完整为准。经图像修复处理的照片应做好记录, 并在著录目录数据库时, 在附注项中加注“图像修复”, 在必要时须经批准后才能进行图像的修复。

5.1.3.3 全文数据格式

档案数字化的全文应提供完整的三种格式, 即单页的JPEG、多页的TIFF和双层PDF/CEB, 支持选择、复印、全文检索等功能。

TIFF、PDF格式的全文数据文件与档案实体及电子目录一一对应, 并以其档号命名, JPEG格式的文件以“档号—页号”的形式命名。档号的构成和形式要符合DB32/ 505-2002的规定。全文数据存储在以全宗号、档号命名的文件夹中。

5.1.3.4 数据检查

目录数据质量检查: 用人工校对或软件自动校对的方式, 对目录数据库著录项目的完整性、一致性、规范性等方面进行全面的质量检查, 对质量不合格的数据提出修改或重新著录要求。

图像质量检查: 对原始档案数字化转换形成数字图像, 由转换操作人员对图像进行逐一质量检查, 对方向不正确、偏斜、清晰度、失真度等不合格的图像质量要求的数字化文件, 提出重新处理或重新采集要求。

数据采集有遗漏应及时补充。

5.2 录音录像档案数字化转换

5.2.1 档案数字化前处理

5.2.1.1 检查录音录像带状况

档案准备时应严格检查录音录像带是否有断带、过紧、过松情况。发现有断带情况, 用胶带粘连等方式进行续接, 过紧或过松的, 用小棒转紧或用手轻轻拍松, 使之能正常播放。为保证播放效果, 在进行数字化转换前, 对存放过久不常使用的录音录像带, 应进行倒进带处理, 防止磁带粘连等情况发生, 同时检查是否规整、有无损坏, 是否需要清洁等情况, 如有不规整或损坏应返还有关部门进行技术修复, 需要清洁的, 应使用清洁器材进行清洁。

5.2.1.2 检查录音录像带播放画面、音质情况

在进行检查和倒进处理后, 应播放录音录像片段, 检查画面、音质情况, 并做好记录, 以便对数字化后的效果进行比较。

5.2.1.3 环境检查

音频视频档案数字化时应远离强磁场等外部干扰。在前处理过程中, 录音录像盘/带与磁场源之间的距离不得少于76mm, 操作人员应戴非棉制手套进行操作。

5.2.1.4 检查记录

对录音录像带领取情况、检查情况进行登记，为数字化转换前后进行对比做依据。数字化完成后及时整理归还。

5.2.2 数字化转换

5.2.2.1 转换方式

依据档案原有的播放要求，通过各种型号录像机、盒式录放机、开盘式放音机等播放设备进行播放，并和计算机进行连接，通过计算机的音频视频采集卡、应用接收模拟信号，对录音录像带播放出来的声音、画面进行录制采集，将采集到的模拟信号量化成数字信号，然后压缩编码成为数字文件。

5.2.2.2 录音录像数字化转换硬件设备

录音录像档案数字化转换硬件设备要求见表4。

5.2.2.3 录音录像档案数字化转换格式

录音录像档案采集或文件格式转换技术参数见表5。

5.2.3 音频视频数字文件存储

5.2.3.1 音频视频数字文件存储格式

音频视频数字文件存储格式为无损WAV文件，以尽量保持音质不变。提供网络查询的音频数字文件，可把原WAV文件转换为MP3、WMA或其他便于网络传输的格式。

表4 录音录像数字化转换硬件设备要求

硬件设备	作用
播放设备	录像机、盒式录放机、开盘式放音机等。
采集卡	1) 计算机通过采集卡接收来自输入端的模拟信号，将采集到的模拟信号量化成数字信号，然后压缩编码成为计算机的数字文件。 2) 视频采集卡应支持 AVI 和 MPEG 格式，支持多种影像源输入。 3) 音频采集卡应至少支持 44.1kHz（最好是 48kHz）采样频率，支持 16 位量化级，支持双声道以上并支持多种声源输入。
计算机	应配备性能稳定、配置较高并配有采集卡的计算机作为录音、录像档案数字化转换设备。
存储设备	数字化文件除存储在硬盘中随机使用外，还应备份在磁带或光盘上。

表5 录音录像档案采集或文件格式转换技术参数

档案类别	主要参数	
	名称	设置
录音档案	采样频率	48kHz，不低于 44.1kHz
	量化级	24bit
	通道数	以单声道录制的，以单声道（MONO）采样； 以多声道录制的，以立体声（STEREO）采样
	文件格式	WAV、MP3

表5 录音录像档案采集或文件格式转换技术参数（续）

档案类别	主要参数	
	名称	设置
录像档案	分辨率（像素）	720*576
	帧数	25 帧/秒
	视频数据速率	4Mbps
	音频位速率	224Kbps
	音频采样频率	48kHz
	音频设定	立体声（STEREO）
	视频制式	PAL
	文件格式	AVI、MPEG-4
采集软件应与采集卡相配套，并根据需求进行参数设置。		
采集过程中，操作人员应全程监视档案信号采集的情况，如发生信号中断或其它意外情况，应重新进行采集。		

5.2.3.2 视频文件存储格式

录像档案采集时，以无损AVI格式采集数据。把无损AVI格式文件转换为MPEG2格式文件，作为常规存储。把MPEG2文件格式转换为RM、MPG、ASF、WMV或其他格式，作为第二套存储。

5.2.3.3 存储方式

数字转换时实时保存进入服务器或磁盘阵列等大容量存储设备。

5.2.4 音视频数字文件处理

5.2.4.1 音频数字文件

对录音带转换的音频数字文件进行振幅标准化处理，以符合正常的声音播放频率。

原录音带噪声过大的，应对其音频数字文件进行适当的降噪处理，使声音更清晰、干净。

音频数字文件首尾空白无内容、且时间过长的部分，应进行适当剪切，使声音开始前和结束后有5秒左右的无声效果即可。

5.2.4.2 视频数字文件

对录像带转换的视频数字文件进行检查，裁掉视频前后无内容的部分。

每个视频数字文件开始时应有5秒左右的无声画面，应根据原录像档案情况适当进行编辑、处理，如截取有录像档案名称的画面，按期编辑的内部参考资料的期数画面等。

5.2.5 数据管理

5.2.5.1 文件命名

音视频数字文件应以录音录像档案档号来命名。为了便于管理和数据挂接，音频数字文件采用的档号格式应与其目录数据库的格式一致。由于录音数字化是其A、B面分别进行采样，其数字文件命名时应标注A、B面，和录音带相对应。便于音频数字化文件的管理和利用，应建立相应的文件夹，即同一全宗每个目录或年度建立一个文件夹。

5.2.5.2 数据格式

应选择通用数据格式，主要为DBF、TXT，能直接或间接地通过XML文档进行交换。

5.2.5.3 数据校验

目录数据质量检查采用人工抽查方式，核对著录项是否完整、著录内容是否规范、准确。为了保证目录数据与音频数字文件能够正确挂接，重点检查档号是否准确，发现数据错录、漏录应及时进行修改、补录。

5.3 照片底片档案数字化转换

5.3.1 档案准备

档案准备包括档案领取、前处理环节。照片档案数字化转换前，按顺序将照片从相册中逐张取出，放回重份的照片，除下影响扫描质量的保护膜、封套等。当受照片保护或保管条件制约无法取出照片时，可采用底片扫描或数码相机拍摄等方式进行数字化转换。

5.3.2 档案还原

照片档案在图像采集完成的同时，取出的照片应及时装册还原，确保照片原件的安全、位置准确、无遗漏。

5.3.3 目录建库

检查照片底片档案目录，确认照片号、底片号、题号、拍摄时间、摄影者、文字说明、密级、保管期限、照片规格、照片类型、照片数量、全文标识等著录项齐全、准确，不齐全的在数字化过程中补齐。漏号、错号的予以纠正，并在原件上标识纠正后的数据项。照片底片档案目录数据库还应包括数字化时间、扫描分辨率、扫描色彩模式、图像压缩方式等著录项。

5.3.4 数据采集

照片档案依据照片品质、幅面、保管条件制约等情况，采用扫描照片、扫描底片、数码相机拍摄方式进行数字化转化。扫描后的照片应能够按原件1:1比例还原，可满足出版、展览等需求，并能真实反映照片、底片原貌（清晰度、色彩、黑白对比度等）。

照片底片数字化硬件设备要求见表6。

照片底片采集或文件格式转换技术参数见表7。

5.3.5 图像修复

对照片档案进行数字化转换时，应经常对数字化设备执行色彩校正程序，以调整数字设备色彩的准确性，提高数字化图像的质量。

由于原件本身的质量问题导致数字化图像的效果不佳，需要做滤镜处理时，可在扫描（或拍摄）时直接使用数字化设备所提供的滤镜功能（如去网纹等），以便达到优于事后使用图像处理软件进行修正的效果。

表6 照片底片数字化硬件设备要求

采集方式	图像品质	适用档案
用平板扫描仪扫描照片	图像品质较好，满足大多数应用。	除少数受幅面和保管条件制约，无法进行扫描的照片。

表6 照片底片数字化硬件设备要求(续)

采集方式	图像品质	适用档案
用带透扫功能的平板扫描仪扫描底片	图像品质高于照片扫描, 低于专业底片扫描。	有底片的、对图像品质要求较高的照片。
用专业底片扫描仪扫描底片	图像品质最佳。	有底片的、对图像品质要求很高的照片。
用数码相机拍摄	与照片原件相比, 会造成图像的失真以及图像清晰度低等不可避免的损失。	少数受幅面和保管条件制约, 无法进行扫描的照片。

表7 照片底片图像采集或文件格式转换技术参数

采集方式	载体规格	主要参数设定			压缩算法与压缩率
		分辨率	色彩模式	文件格式	
照片扫描	≤10 英寸	600dpi	黑白照片采用灰度模式/彩色照片采用 RGB 全彩模式。	TIFF、JPEG	TIFF 采用无损压缩方式
	>10 英寸	300dpi			
底片扫描	≤120mm	2400dpi	黑白照片采用灰度模式/彩色照片采用 RGB 全彩模式。	TIFF、JPEG	TIFF 不压缩
	>120mm	1200dpi			
数码相机翻拍	大幅或超大幅照片	≥1000 万像素	采用 RGB 全彩模式。	TIFF、JPEG	TIFF 不压缩

5.3.6 图像拼接

大幅面档案进行分区扫描形成的多幅图像, 应进行拼接处理, 合并为一个完整的图像, 以保证档案数字化图像的整体性。以图像拼接处理的照片档案应做好记录, 并在著录目录数据时, 在附注项中加注“图像拼接”。

5.3.7 图像处理

对扫描图像应进行裁边处理时, 应去除多余的边框或白边, 以便有效减少数字图像文件的容量, 节约存储空间。

对扫描图像偏斜度、清晰度、失真度、完整情况等进行检查, 发现问题及时处理或重新扫描。

发现照片底片漏扫时, 应及时补扫, 并填写图像处理情况表单, 记录检查结果和处理意见。

对曝光不足或曝光过度的照片和底片, 应采取增加或减少亮度等措施进行补救处理。

5.3.8 图像存储

扫描图像以多页TIFF或JPEG格式存储。对于重要照片和扫描质量要求较高的照片，以TIFF（LZW）无损压缩格式优先存入统一的照片档案数据库，以便进行缩微备份转换。

所有扫描图像应按“照片号”命名并建立文件夹存储，应将同一题名的若干扫描图像文件存放在一个文件夹中，按一级类目、二级类目、文件分级建文件夹存储。命名的格式应与目录数据库中“照片号”的著录格式一致。

5.3.9 版权保护

为长期保管和版权保护需要，档案馆馆藏照片底片档案扫描获得的图像应加添“×××档案馆馆藏”数字水印，或通过CEB/X或PDF/A长期保管格式工具转换长期保管格式并批量添加数字水印。

5.4 缩微胶片档案数字化

5.4.1 数字化范围

纸质和缩微胶片并存的档案，采用缩微胶片转换的方式，不再对档案实体原件进行扫描数字化。宜选择第二代或第三代缩微拷贝胶片进行数字化。

5.4.2 检查整理

扫描之前，应检查缩微胶片物理形态无卷曲、变形、脆裂、粘连、乳剂层脱落等情况；无可见性微斑、变色、生霉等情况；密度、解像力等技术指标无明显变化；无影响缩微胶片影像可读性的其他情况。如不满足上述要求，可先对缩微胶片进行处理。必要时调用档案原件进行扫描，以确保数字化质量。

5.4.3 内容检查

应严格检查档案的完整性，对存在的问题加以记录和说明；检查档案有无漏拍、补拍、分幅、合幅、双幅、重复拍照等情况并进行记录，以便提示工作人员在扫描时按要求进行处理；检查档案的页号顺序和页数，对照档案目录逐条记录页号、页数。

5.4.4 胶片扫描

5.4.4.1 调用与归卷

从库房调用缩微胶片时应按有关规定进行温、湿度平衡调整；扫描过程中，应对缩微胶片进行妥善的保管和保护，工作人员应戴洁净的棉质薄手套，轻拿缩微胶片的边缘。

扫描工作完成后，再次整理缩微胶片，应保持原排列顺序不变，做到齐全、准确、无遗漏。

5.4.4.2 扫描方式

根据缩微胶片扫描设备的型号和图像质量，选择自动扫描和手动扫描。一盘（张）缩微胶片里影像的密度、解像力、幅面尺寸基本一致时，可选择自动扫描方式；不一致时，可对对比度、曝光亮度、画幅大小进行调整，选择手动扫描方式。

16或35mm卷片档案缩微胶片，一般采用“对中扫描”，以扫描后的图像清晰、完整、不影响利用效果为准，以最大程度获取影像信息为宜。

缩微胶片档案图像采集技术参数见表8。

表8 缩微胶片档案图像采集技术参数

扫描对象	倍率	分辨率	压缩方式	色彩模式	对比度
35mm 缩微胶片	12	300dpi	G4-2D, 压缩率应以图像清晰、可读、完整为前提。	字迹清晰的影像采用黑白二值模式, 字迹清晰度差或带有插图的影像, 可采用灰度模式扫描。	根据缩微胶片影像的密度、解像力进行调整和设定, 一般对比度为60~70。
16mm 缩微胶片	24				
平片缩微胶片	24				
A4 幅面档案原件低缩率拍摄/OCR 识别	/	200dpi/300dpi	/	/	/
幅面较大、采用高缩率拍摄		600dpi			

按照档案目录逐条核对实际扫描的页号、页数与档案整理时的页号、页数是否一致, 不一致时应注明原因和处理办法; 填写扫描登记表, 记录工作种类、缩微胶片盘(张)号、胶片种类(银盐、重氮)、代数、扫描时间、设备型号、技术参数、出现的问题、处理情况以及责任人等。

5.4.4.3 图像处理

5.4.4.3.1 补扫

扫描后对图像完整性、清晰度、失真度等进行检查, 对漏扫的文件和不符合质量要求的应进行补扫, 并插入正确位置。

5.4.4.3.2 纠偏

对偏斜的图像应进行纠偏处理, 以达到视觉上基本不感觉偏斜为准。对方向不正确的图像进行旋转处理, 以符合阅读习惯。

5.4.4.3.3 去污

在去污处理过程中, 应保持档案原貌, 维护档案的真实性。对扫描中产生的影响图像质量的黑点、黑线、黑框、黑边等进行去污处理。

5.4.4.3.4 拼接

为保证图像内容的完整性, 应对分幅图像进行拼接。

5.4.4.3.5 记录

填写相关图像处理登记表, 记录工作种类、图像出现的问题、处理情况以及责任人等。

5.4.4.4 图像存储

采用黑白二值模式扫描的图像文件, 一般采用TIFF(G4)格式存储; 采用灰度模式扫描的文件, 一般采用JPEG格式存储;

提供网上检索利用的图像文件, 也可存储为CEB、PDF等格式。

5.4.4.5 目录建库

目录数据准备与输入，按照 DA/T 18-1999、GB/T 20163-2006 和 DB32/ 505-2002 的要求，规范档案中的目录内容，确定档案目录的著录项目、字段长度和内容要求，建立目录数据库。

图像文件命名，宜用档案目录中的档号作为文件夹名，文件夹内的图像文件，按缩微胶片的盒(片)号及页码顺序命名。

5.4.5 数据校验

每一份图像文件的名称、页数与目录数据中的档号、页数应一致。图像文件的名称与目录数据建立起一一对应的关联关系。

采用人工校对或软件自动校对的方式，对目录数据的质量进行检查。查检著录项目是否完整、著录内容是否规范、准确，对不符合要求的数据进行修改。

应选择通用的DBF或TXT数据交换格式，所选定的数据格式应能通过XML文档进行数据交换。

5.5 实物档案数字化转换

5.5.1 档案整理

对拟数字化的实物档案进行目录数据著录，并确保数据正确。著录项一般包括：档号、形成时间、责任者、保管期限、载体类型、载体数量、载体规格和载体单位。

对实物档案内容和载体质量进行检查，有残缺的应进行修补，并做好记录。

对载体进行全面彻底的卫生清理，确保载体的清洁，表面无反光或反光点。

5.5.2 数字化准备

设置拍摄场地，提供侧主光、底光、顶光、背光等柔性照明，采用单一背景烘托实物。

宜采用单反数码相机和标准、变焦镜头进行彩色拍照，真实反映实物的全貌。

5.5.3 数据采集

拍摄时被拍摄物的图像一般可占取景框的2/3。珍贵的实物采取多角度拍摄，以便今后制取立体影像。

图像质量可选择“标准”，图像尺寸可选择1280×960；需制作大幅面照片或展览图片的拍摄的图像质量可选择“精细”，图像尺寸可选择2272×1704。

字画拍摄可采用分段分区方式，各分区的大小尺寸相近、分辨率和曝光量一致，相邻图像之间应有部分重叠，重叠面积不少于分区面积的1/20。通过PHOTOSHOP等专业软件进行无缝拼接，合并为一个完整的图像，以保证档案数字化图像的整体性。在著录目录数据库时，在附注中加注“图像拼接”。拍摄的图像应曝光正确、反差适中，光线采用侧立光或反射光。

5.5.4 图像存储

根据保存目的，图像数据可选择TIFF或JPEG格式存储。需要制作大幅面照片的，宜选择TIFF格式。

目录数据按一线类目、二级类目文件分级建立文件夹存储。各层文件简介的命名由相应档案类别的分类码组成。

6 数字化文件存储

6.1 数字化文件存储级别和格式

根据保存和使用的不同需求，数字化文件的存储可分为以下三种级别，见表9。

表9 数字化文件存储及格式

存储级别	生成	作用	格式	字母
档案典藏级	由档案数字化文件无损压缩存储生成。	档案典藏级图像用于出版和印刷可作为档案典藏以及格式转换或复制的母本。	纸质 TIFF（黑白二值模式）/ CEBX / PDF； 纸质 JPEG（灰度模式和彩色模式）CEBX / PDF 照片 TIFF/ CEBX / PDF 音频 WAV 视频 AVI、MPEG-4	S（STORAGE）
复制加工级	复制加工级由档案典藏级 TIFF 文件 1:1 转换生成。	加工复制各种精度、大小的屏幕浏览图像的母本文件，供专家、合作伙伴及专门组织成员网上有条件权限的访问。	纸质 JPEG 照片 JPEG	C（COPY）
浏览利用级	纸质、照片档案浏览利用级由复制加工级 JPEG 文件批量降低分辨率转换生成。 录音录像由浏览利用级由档案典藏级转换而成。	提供政务网或因特网用户的访问利用。	纸质 JPEG 照片 JPEG 音频 MP3、WMA 视频 WMV、RMVB	B（BROWSE）

6.2 图像文件的命名

每份纸质档案、照片档案、录音录像档案都有一个与之相应的唯一的档号，以档号为这份档案转换后数字化文件的命名。

6.3 数据关联

通过目录数据库的档号与数字化文件的一致性和唯一性，建立起一一对应的关联关系，以实现目录数据库与数字图像文件的批量挂接，实现与档案实体的一一对应。

6.4 数据质检

数字化文件入库存储，应采用人工校对或软件自动校对的方式进行目录总体质量检查，对目录数据库的建库质量进行检查。对目录数据库著录项目的完整性、著录内容的规范性等进行检查，对质量不合格的数据提出修改或重新著录要求。图像质量检查要对偏斜度、清晰度、失真度等不符合图像质量要求的数字化文件，提出重新处理要求。播放录音录像档案，并同时在计算机设备中播放与之对应的数字化文件，通过比对二者信号是否相同来进行校对，如数字化文件存在图像偏差、声音延迟等差错，不符合视频。音频质量要求时，应重新采集。数据采集有遗漏时应及时补充。

6.5 数据挂接

档案目录数据库和数字化文件,通过质检环节确认为合格后,通过网络或加密离线载体及时加载到数据服务器端汇总。通过相应软件实现目录数据对电子文件全文的自动关联,加入对应的电子地址信息等,实现批量、快速挂接。

6.5.1 数据挂接

档案数字化后形成的数字文件,经过质量检查确认合格后,通过相应的软件与对应的档案目录数据库进行批量挂接。

6.5.2 检查挂接

数据挂接后,以档案目录数据库为依据,对挂接的数字文件进行检查,核对数字文件的档号与档案目录的档号是否一致,发现问题及时修正。

6.5.3 挂接登记

对汇总挂接及检查挂接的情况进行登记,对挂接错误的注明错误原因,处理办法及修正结果。

7 数据验收

7.1 数据抽检

以抽检方式检查目录数据、图像数据的质量;一个批次数据抽检的比率不得低于5%。目录数据、图像数据有不完整、不清晰等质量问题时,抽检标记为“不合格”,不合格的应予以改正;一个批次数据质量抽检的合格率达到99.5%以上(含99.5%)时,验收予以通过。

抽检合格的文件数=抽检文件总数-抽检不合格的文件数。

合格率 = (抽检合格的文件数 / 抽检文件总数) × 100%

验收登记,填写数据验收登记表,记录验收结果,验收结论应经审核、签字后方有效。

以抽检的方式检查已完成数字化转换的所有数据质量情况,包括目录数据、数字化文件及数据挂接的总体质量检验。抽检数据以全宗为单位进行,抽检率不得低于10%。

抽检率=抽检文件数/待验收全宗的文件总数×100%。

7.2 验收指标

抽检的目录数据库、数字化文件不符合质量要求或目录数据库与数字化文件挂接错误,出现不完整、不清晰、图像偏差、声音延迟、有错误等质量问题时抽检标记为“不合格”。以全宗为单位,抽检数据的合格率纸质档案为99%,照片、缩微胶片、实体档案、录音录像档案100%。没有通过验收的数据退回质检环节进行检查、整改。

合格率=抽检合格的文件数/抽检文件总数×100%。

7.3 验收登记

验收完成后填写验收相关表单,完成验收报告,验收表单及验收报告应经过审核,签字确认。验收组成员应符合档案管理相关制度规定的人员组成。

8 数据备份

8.1 备份范围

对验收合格的各类数字化文件的档案目录数据库、全文和多媒体数据库及时进行备份。

8.2 备份方式

备份载体的选择应多样化,至少两种以上备份物理拷贝介质,其中一套应为可移动的、离线的永久备份介质。档案数据可采用服务器硬磁盘、光盘和缩微胶片三级异质备份方式。

可采用在线、离线相结合的方式实现数据传输和备份。离线载体依次推荐为缩微胶片、大容量蓝光数据存储光盘、加密移动硬盘,一式三套,一套封存保管、一套提供利用、一套异地保存。

8.3 备份数据校验

备份数据校验内容主要包括:备份数据能否打开、数据信息是否完整、文件数量是否准确等。

8.4 备份标签

数据备份后应在相应的备份介质上做好标签,以便查找和管理。标签内注明备份载体编号、载体类型、套别、密级、保管期限、文件数量、文件内容、制作人员、制作时间、内容索引、档案保管单位数据备份单位、数据备份标准、阅读备份数量所需要的硬件环境、软件环境、转储更新日期等。载体编号由数字化文件类型代码、档案原件类别和载体序号组成。

9 数字化成果管理

应加强对纸质档案、照片档案、视频、音频档案数字化成果的管理,制定相应的管理制度,对备份数据定期检查,确保数据的安全、完整和长期可用。

档案数字化成果应作为档案信息资源的重要组成部分,由档案管理系统进行管理、保存和利用。具有永久和长期保存价值的成果,应作为电子文件移交同级国家档案馆。

纸质档案、照片档案、视频、音频档案数字化成果在提供出版印刷或网上传输、发布时,应采用相应的防伪措施(如数字水印、制作单位的电子标识等),以起到版权保护、数字化文件的真伪鉴别、产品标志和秘密通信等作用。

10 安全保密要求

10.1 档案数字化承建机构

需进行档案数字化工程的各级各类档案馆、档案室,应按相关档案及保密有关的法律、法规、制度、标准要求,选择具有一定规模、管理制度健全、经验丰富、具备相关资质要求的承建单位,选择的承建单位应先经过行政管理部门和保密工作部门的保密审查,接受指导监督。应与承担档案数字化工程单位签订保密协议。

境外机构或人员不得承建档案数字化工程。

10.2 保密管理制度

档案数字化过程中应实行保密管理领导责任制,制定与信息涉密等级相应的管理制度,定期对制度实施情况进行检查。

10.3 人员管理

档案数字化工作人员上岗前应进行保密制度审查与考核，档案数据化人员应保持相对稳定。外部承建单位所有参加数字化工作的人员应通过安全保密审查，签订保密协议。

10.4 数字化加工设备的管理

档案数字化生产加工过程中所使用的扫描仪、声频（视频）采集卡等设备，应符合有关保密技术标准和保密要求，并按涉密设备进行管理。不应使用无保密保障的扫描设备。根据所处理档案的最高密级，确定计算机及其存储介质的密级，采取相应的保密技术防护措施。如果数字规模较大、时间较长，可安装外联监管软件和移动存储介质管理系统，并登记报本级保密部门备案。

数字化加工使用的计算机应加装封条，并与相关设备组成专用网络，与其他网络实行物理隔离，禁止接入互联网或单位内部局域网。不应安装具有无线互联功能的硬件模块和无线鼠标、键盘等无线互联的外围设备。任务完成后，应当继续按涉密计算机使用和管理，或者拆除硬盘。使用的硬盘、移动硬盘、U盘、光盘等存储介质，按涉密载体进行管理，不应在非涉密计算机及其网络上使用。

10.5 加工场所的管理

数字化加工场所一般设在单位内部，按临时保密要害部位进行管理，应在生产加工场所安装监控摄像头和录像设备，全程摄录加工现场情况并归档备查。

生产加工场所应有专门的休息室或储物柜，便于工作人员存入私人物品。严禁将手机、数码相机、数码摄像机、私有移动存储介质等设备和水瓶、水杯、打火机等物品带入生产加工场所。

应制定书面消防预案，加工场地应配备一定量手携消防器材，定期、专人检查。设置兼职消防安全员，每五人配备一名，实施消防安全员值日制度，定期检查，发现异常情况立即报告并采取适当措施。

10.6 档案安全

档案在数字化过程中未经批准，任何人不得查阅和借出档案。

档案调卷、还卷、作业中各道手续应详细登记，登记内容应包括：前处理拆卷人、录入人、扫描人、质量检验人、后整理装订人等，各登记记录及时整理、汇总、装订成册。

档案出入库运输中不得人档分离。运输过程应对档案进行防护，防止档案受潮、受污、受损、丢失等异常情况发生。数字化作业完毕的档案应及时归还库房，暂存数字化加工场所的档案应专人管理，对档案进行妥善保存。

10.7 数据安全

应对存储数据的光盘和硬盘进行防护，防止由于防护不当造成的弯曲、划伤、破损等现象发生。

数据备份介质保管环境温度为17℃～20℃，相对湿度应35%～45%。

备份数据按存贮介质要求按一定周期进行有效性、安全性检查，并重新刻录或备份数据。

当数据运行的软、硬件平台发生改变，应及时对数据进行转换、迁移，并做好备份。

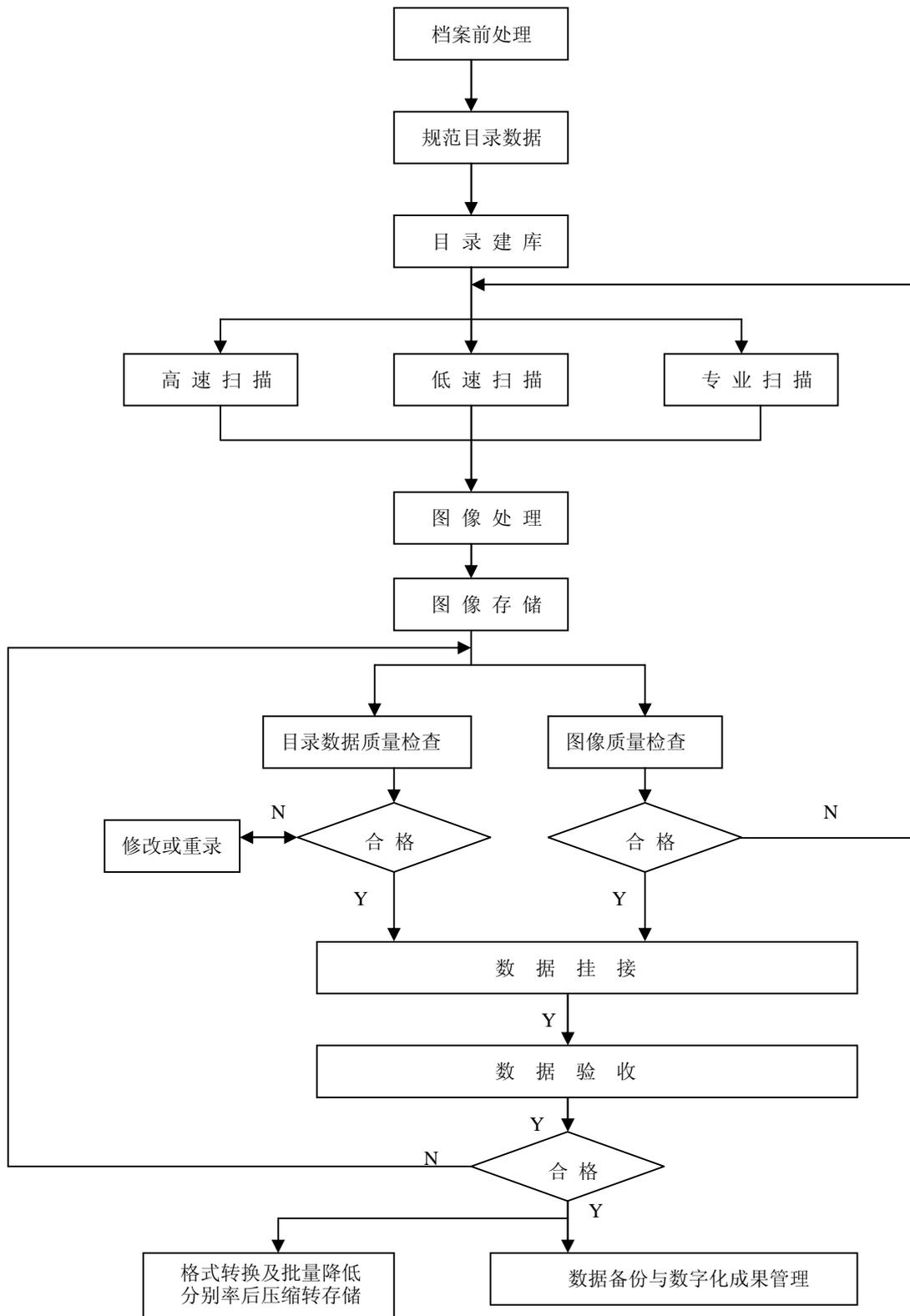
数据检测与维护过程应进行登记，数据检测时说明检测方法、结果及处理措施，数据转换时应说明数据转换的原因及采取的转换措施、转换数据数量等。

A

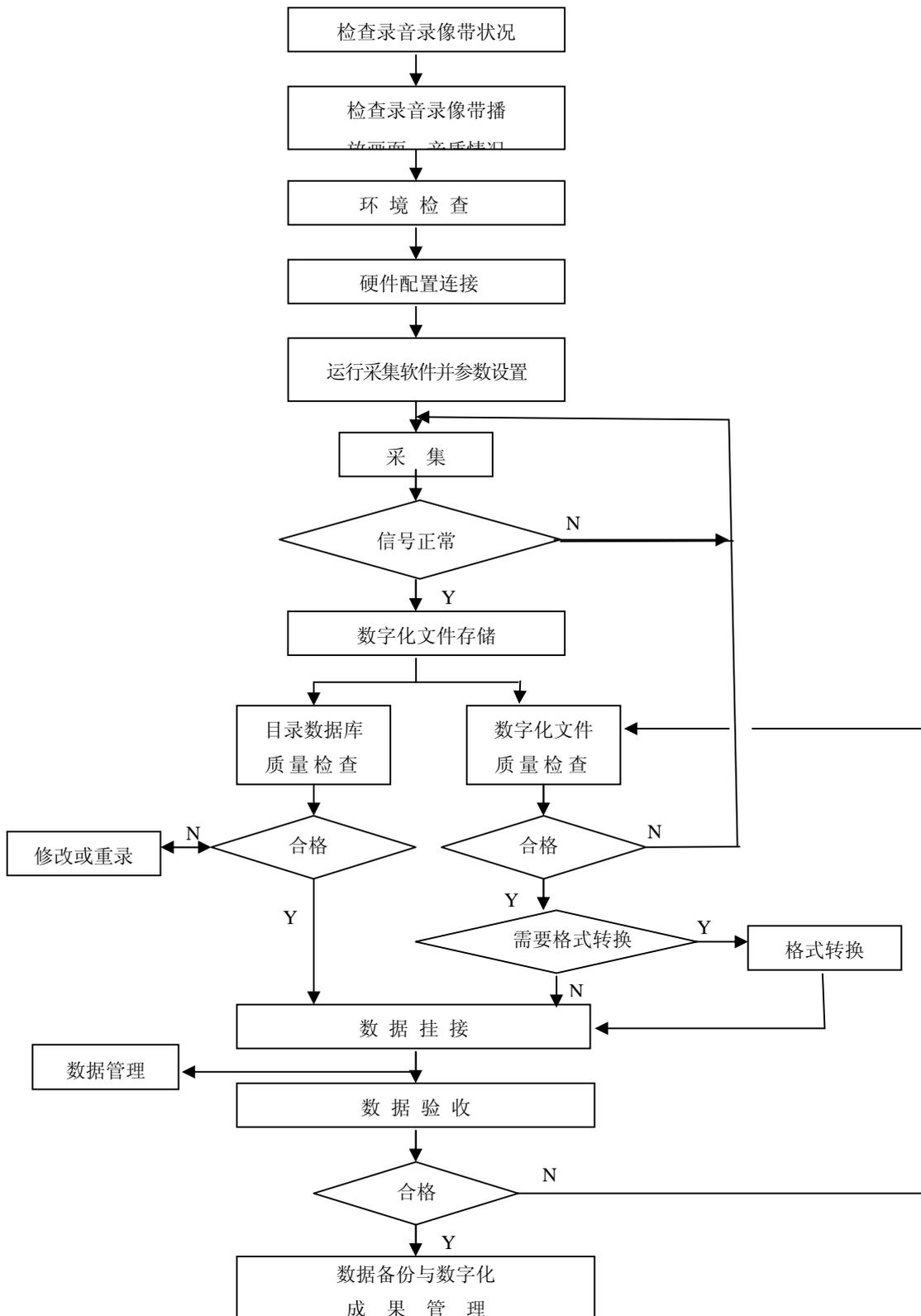
附录 A
(规范性附录)

档案数字化转换流程图

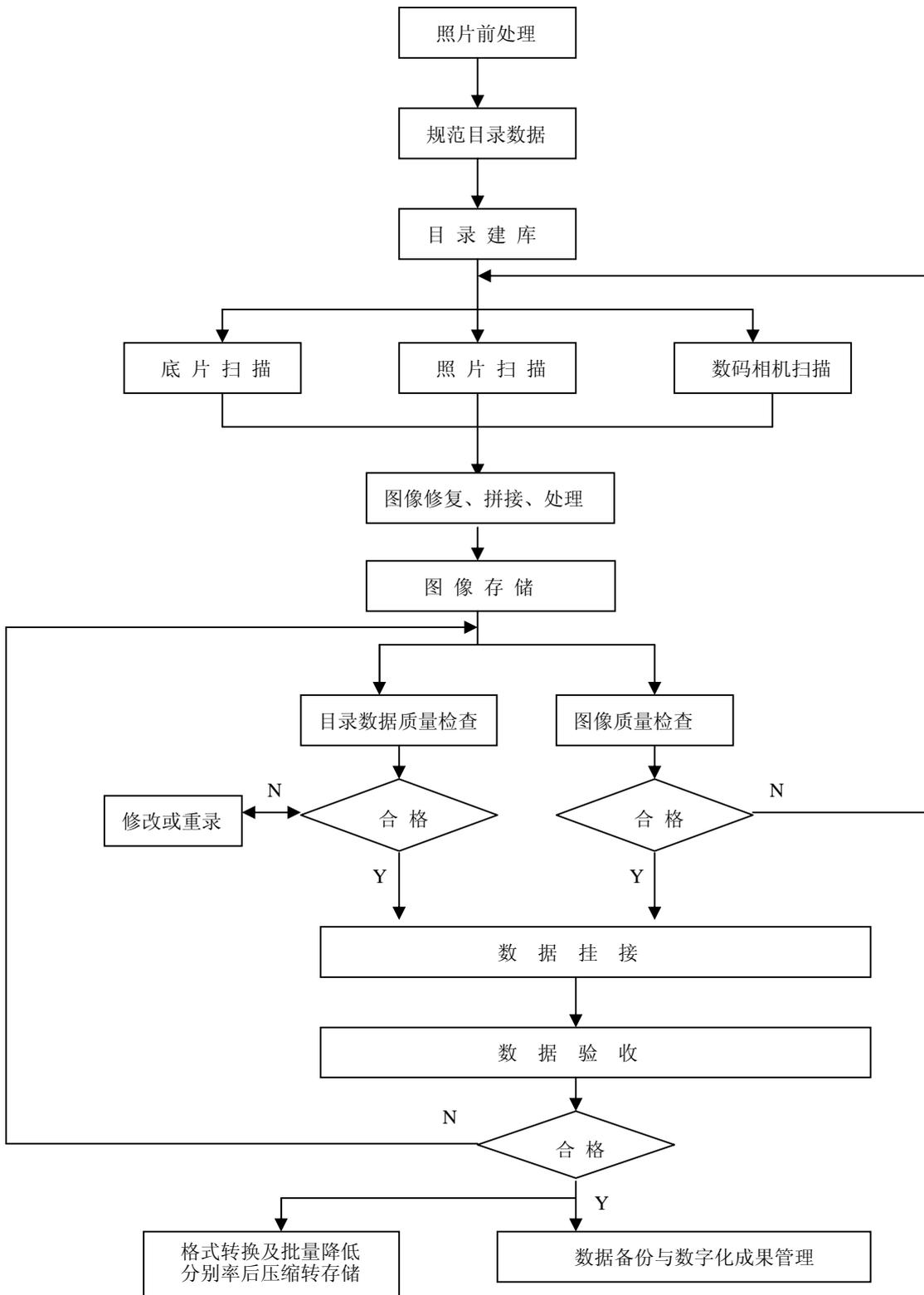
A.1 纸质档案数字化转换操作流程



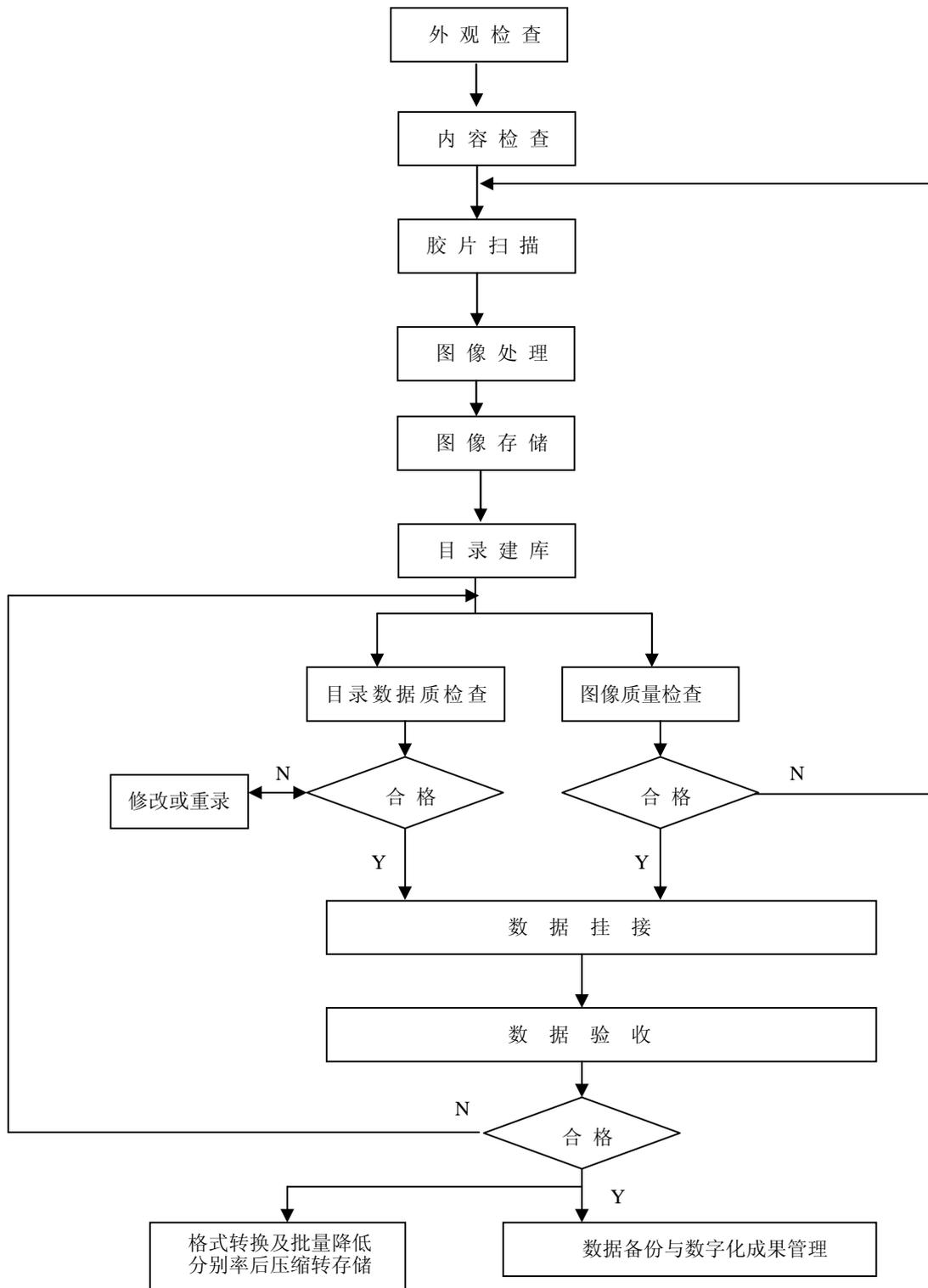
A.2 录音录像档案数字化转换流程图



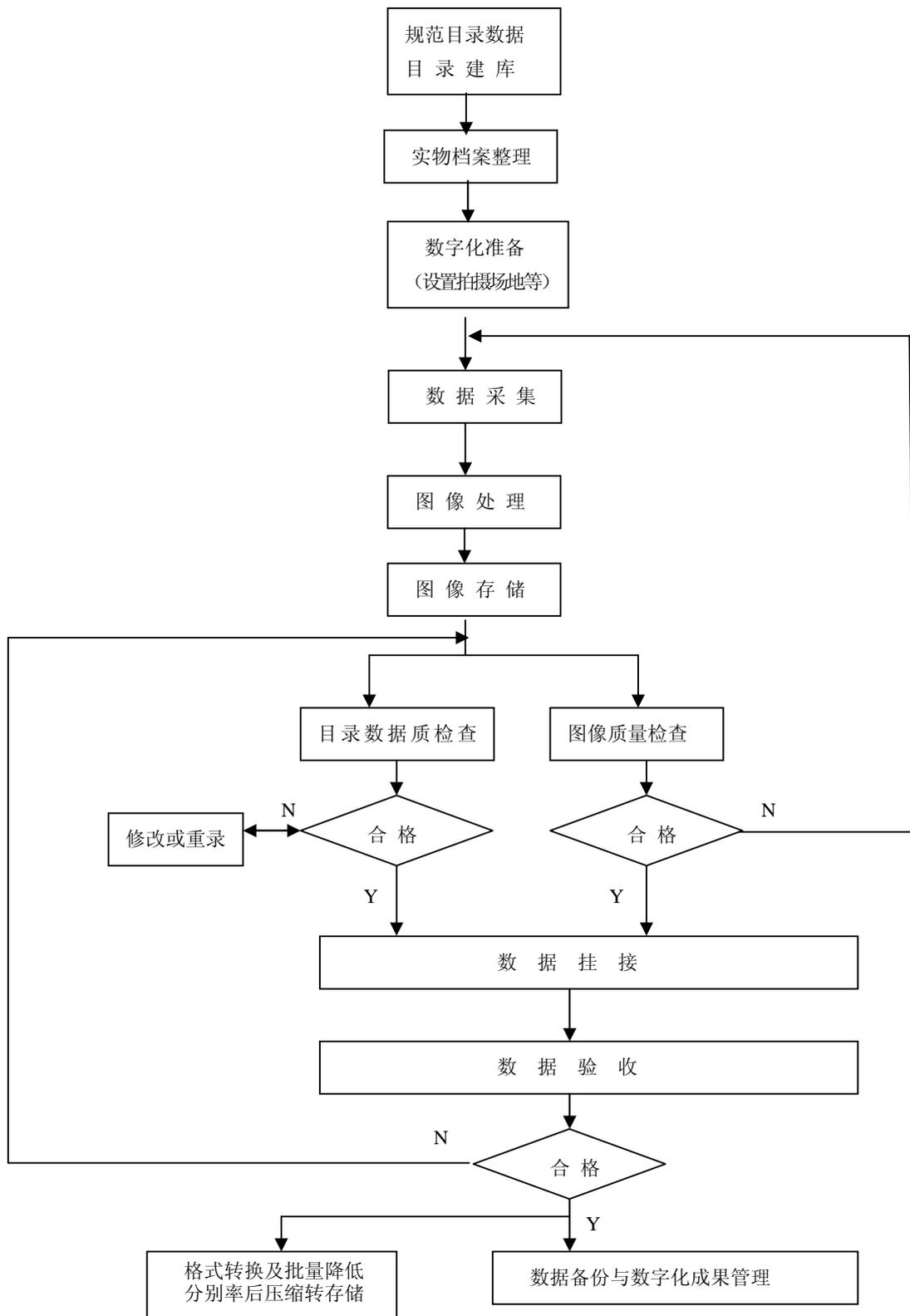
A.3 照片档案数字化转换流程图



A.4 缩微胶片档案数字化转换流程图



A.5 实物档案数字化转换流程图



附 录 B
(资料性附录)
档案数字化转换用表格

表B.1 纸质档案领取登记表

序号	全宗号	全宗名称	案卷号	件数	页数	备注
移交单位				领取单位		
移交人				领取人		
移交时间				领取时间		

表B.2 纸质档案数字化流程交接登记表

全宗号:		目录号:		案卷号:		文件数量(总计):	
档案整理			档案扫描		图像处理		备注
序号	件号	档案整理情况登记	件号	扫描情况	件号	处理方式	
操作人员:			操作人员:		操 作 人:		
操作时间:			操作时间:		操作时间:		
色彩模式:			分辨率:		存储格式:		

表B.4 录音录像档案领取登记表

序号	全宗号	全宗名称	档号	盘片号	载体类型、规格	备注
移交单位				领取单位		
移交人				领取人		
移交时间				领取时间		

表B.5 录音录像档案整理情况交接登记表

盘片号		档号	
题名			
讲话人姓名 (录音带)		内容制式 (录像带)	
载体类型		载体规格	
时间长度		责任者	
<p>原件整理情况登记:</p> <p>操作人:</p> <p>操作时间:</p>			

表B.8 照片档案数字化流程交接登记表

全宗号:		册号:		照片数量(总计):			
照片整理			照片扫描		图像处理		备注
序号	张号	照片整理情况登记	张号	扫描情况	张号	处理方式	
操作人员: 操作时间:			操作人员: 操作时间:		操 作 人: 操作时间:		
色彩模式:		分辨率:		存储格式:			

表B.9 档案数字化质检登记表

全宗号:		质检人:		质检日期:		
工作环节	工作机号	工作人员姓名	档号	质量问题原因	处理结果	质检结论

表B.11 档案数字化转换验收登记表

全宗号:		全宗名称:		文件数量:	
抽检 序号	档号	是否 合格	不合格原因	处理结果	
合格率		验收结论			
验收人		验收时间			
审批人		审批时间			

备注：1、合格打“√”不合格打“×”。

2、不够填可将多张表格粘贴在一起，只在末页填写合格率、验收结论，并签上姓名、日期。

3、验收登记表可以全宗为单位汇编成册，形成台帐。