

河南省档案馆照片档案数字化技术规范

1 范围

本标准规定了河南省档案馆照片档案数字化的主要技术要求。

本标准适用于河南省档案馆开展照片档案数字化工作，适用于省直向省档案馆移交照片档案数字化成果的单位。各市、县级综合档案馆参照执行。

2 规范性引用文件

下列文件中的条款通过本标准的引用而成为本标准的条款。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 11821-2002 照片档案管理规范

DA/T 18-1999 档案著录规则

DA/T 31-2017 纸质档案数字化技术规范

DA/T 54-2014 照片类电子档案元数据方案

3 术语和定义

下列术语和定义适用于本标准。

3.1 照片档案

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影像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、照片和说明三部分。

3.2 数字化

用计算机技术将模拟信息转换为数字信息的处理过程。

3.3 照片档案数字化

采用扫描仪、数码相机等设备对照片档案进行数字化加工，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。

3.4 数字化文件

照片档案经数字化加工后形成的数字图像文件。

3.5 图像分辨率

单位长度内图像包含的点数或像素数，一般用每英寸点数字（dpi）表示。

3.6 TIFF (Tagged Image File Format) 格式

TIFF（标记图像文件格式）是一种基于标记的不丢失信息的无损压缩格式，用于在应用程序之间和计算机平台之间交换文件。TIFF 是一种灵活的位图图像格式，存储的图像细微层次的信息较多，图像质量较高，适合于色彩单一图像的压缩存储。

3.7 JPEG（Joint Photographic Expert Group）格式

JPEG 是 ISO/IEC 联合图像专家组制定的静止图像压缩标准，是适用于连续色调（包括灰度和彩色）静止图像压缩算法的国际标准。JPEG 格式是目前所有格式中压缩率最高的格式，图像质量多数情况下均可接受，尤其适用于屏幕和打印显示，支持所有主要的计算机平台和 Web 浏览器。从储存空间及传输效率考虑，适合于色彩复杂图像的压缩储存。

3.8 RGB（全彩）模式

RGB 模式由红（Red）、绿（Green）和蓝（Blue）三种原色组合而成。在 RGB 模式下，每一个像素由 24 位的数据表示，因此每一种原色都可以表现出 256 种色调，三种原色混合起来可以生成 1677 万种颜色，即真彩色，适用于表现色彩丰富的图像。

4 基本要求

4.1 基本原则

照片档案数字化的格式与技术参数应根据照片图像保存与使用的不同需求做出适当选择。

4.2 安全原则

应加强照片档案数字化各环节的安全保密管理机制，确保档案原件和数字化档案信息的安全。照片档案数字化转换过程中所涉及的任何设备均不得以任何方式与互联网连接。

4.3 基本环节

照片档案数字化的基本环节包括：档案整理、目录建库、图像采集、图像处理、数据质检、数据挂接、数据验收、数据备份与成果管理等。

4.4 过程管理

照片档案数字化的过程应做详细记录，并及时整理、汇总、装订成册，在数字化工作完成的同时建立起完整、规范的照片档案数字化台帐。

5 档案整理

在数字化转换之前,根据照片档案管理情况,按下述步骤对照片档案进行整理,确保照片档案数字化质量。

5.1 照片检查

5.1.1 照片档案数字化转换前,按顺序将照片从相册中逐张取出,剔除重份的照片,去除影响照片扫描质量的保护模、封套等。

5.1.2 检查照片保存状况,对照片破损、变形、污渍、霉斑、褪色等情况进行登记。

5.1.3 在扫描前对照片的破损、变形、污渍、霉斑、褪色等情况不做任何技术复原,但在确保不会损伤照片的前提下,应对照片进行简单的清洁处理,如去除灰尘、污渍或霉点等,以提高扫描图像质量。

5.1.4 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。

5.2 装册还原

图像采集完成后,取出的照片要及时装册还原,确保照片原件的安全、位置准确、无遗漏。

6 目录建库

6.1 建立目录数据库

按照《河南省档案馆照片档案目录数据库结构规范》,确定照片档案的著录项,建立照片档案目录数据库。

6.2 目录数据质量要求

应采用人工校对或软件自动校对的方式,对目录数据库的建库质量进行检查,确保著录项目完整,著录内容准确、规范,发现不合格的数据应及时进行修改或重录。

7 图像采集

7.1 采集方式

照片图像的采集方式应根据照片档案的实际情况及使用需求选择。

7.1.1 底片扫描

扫描对象为底片。扫描的图像品质最佳，适用于有底片的、对图像品质要求较高的照片扫描。可采用底片扫描仪（胶片扫描仪）直接扫描底片转化为图像。

7.1.2 照片扫描

扫描对象为照片。根据照片幅面的大小选择相应规格的扫描仪进行扫描。如照片幅面过大，可采用小幅面扫描后进行图像拼接的方式处理。

7.1.3 数码相机拍摄

拍摄对象为照片。可能造成图像失真及图像清晰度低，仅适用于没有底片而照片又受幅面和保管条件制约、无法进行扫描的情况。

7.2 色彩模式

7.2.1 采用底片或照片扫描方式时，黑白和彩色底片（照片）均应采用 RGB 全彩模式。

7.2.2 采用数码相机拍摄方式时，应采用 RGB 全彩模式。

7.3 分辨率

7.3.1 采用底片扫描方式，底片小于或等于 120 毫米时，分辨率应不小于 2400dpi；底片大于 120 毫米时，分辨率应不小于 1200dpi。

7.3.2 采用照片扫描方式，照片小于或等于 5 英寸时，分辨率应不小于 600dpi；照片大于 5 英寸时，分辨率应不小于 300dpi。

7.3.3 采用数码相机拍摄方式，应不小于 1000 万像素。

7.4 图像存储

7.4.1 存储格式

扫描后的图像文件宜存储为 TIFF 格式或 JPEG 格式（压缩率为 100%）。

7.4.2 缩略图

扫描后的图像文件应同时生成缩略图。

7.4.3 文件命名

扫描后的图像文件名称应对应照片档案的照片号保持一致。缩略图的命名应在对应的名称后加“s”。

8 图像处理

8.1 方向调整

应对方向不正确的图像应进行旋转还原。

8.2 纠偏

应对偏斜的图像进行纠偏处理,以达到视觉上感觉不到偏斜为准,偏斜不宜超过3度。

8.3 裁边

应对在扫描过程中产生的多余的白边进行剪裁。剪裁时应确保档案信息的完整。

8.4 图像拼接

对大幅面的照片档案进行分幅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。拼接时应确保拼接处平滑地融合,拼接后整幅图像无明显拼接痕迹。

9 数据质检

9.1 目录数据质量检查

应对目录数据库著录项目的完整性、著录内容的规范性进行检查,对质量不合格的目录数据进行修改或重新著录。

9.2 图像质量检查

应对扫描处理后的图像是否清晰、完整进行检查。遗漏的应进行补扫;不符合质量要求的,应重新进行处理或重扫。

10 数据挂接

10.1 数据汇总

照片档案数字化过程中分散存储在多台计算机中的目录数据和图像数据,经过质量检查后,应及时汇总到数据服务器端。

10.2 数据关联

汇总后的数据应通过编制程序或借助档案管理软件实现目录数据与图像数据的挂接。

11 数据验收

11.1 数据抽检

以抽检方式检查目录数据库、图像文件、数据挂接的质量以及数字化工作记录是否完整等。一个批次数据抽检的比率应达到数字化总页数的10%以上(含10%)。

11.2 验收指标

11.2.1 当目录数据录入有错误或图像文件不完整、不清晰或目录数据与图像文件挂接错误时，确定为“不合格”。

11.2.2 统计“不合格”的页数。抽检合格的页数=抽检页总数-抽检不合格的页数，合格率=抽检合格的页数/抽检页总数×100%。

11.2.3 合格率应达到99.7%以上（含99.7%）。

11.2.4 数字化工作记录应完整。

12 数据备份

12.1 备份范围

对验收合格的照片档案目录数据库和数字化文件及时进行备份。

12.2 备份方式

为保证数据安全，备份载体的选择应多样化，可采用在线、离线相结合的方式实现多套备份，并做好异地保存。

12.3 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。

12.4 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。

12.5 备份登记

对每次备份的内容、数据格式、数据大小、文件数量、数据运行平台、备份时间、备份介质、备份套数、备份人及数据存放位置等进行记录。

13 数字化成果管理

13.1 应加强对照片档案数字化成果的管理，确保其安全、完整和长期可用。

13.2 照片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、图像文件的真伪鉴别、产品标志和秘密通信等作用。