德阳市城建档案馆

工程电子档案的编制要求

## **工程电子档案的收集**

## **1、** 收集建设过程中产生的各类电子文件，通用格式应符合表1的要求。

|  |  |
| --- | --- |
| 文本档案 | 双层PDF |
| 图纸档案 | 单层PDF、DWG |
| 照片档案 | JPEG |
| 视频录像档案 | AVI、MPEG |
| 声音档案 | MP3、WAV |

表 1 各类电子文件通用格式

**2、** 电子文件的内容必须真实，准确，完整，且与相应的纸质文件、声像档案内容一一对应。

**3、** 非通用文件格式电子文件，收集时候应将其转换成通用格式。

4、 在无法获取纸质对应电子档案的情况下，应对该纸质文件进行扫描，并收集扫描后的电子档案。

（1）对纸质档案进行数字化加工过程的所有环节，均应符合《纸质档案数字化技术规范》（DA/T31-2005）和《四川省档案资料数字化标准》的要求，确保建设电子档案的真实性、准确性。

（2）数字化加工应在纸质档案整理完毕、并经检查审核符合归档要求后进行且与纸质档案完全一致。纸质档案数字化加工工作的环节主要包括：档案整理、档案扫描、图像处理、图像存储等。

（3）存储方式以JPEG数字压缩编码格式，档案文件彩色扫描，扫描的数字图像采用JPEG数字压缩编码存储。对于通过扫描仪获取的电子档案，文本档案的分辨率不低于300DPI，图纸档案的分辨率不低于300DPI，彩色深度24位。如原文件质量较差且尺寸较小，可作分辨率、对比度等调整处理，以扫描后图像按尺寸显示为准。照片档案的分辨率不低于400DPI，并对扫描好的文件进行加工处理。（扫描加工要求见附录）

（4）数字化加工后形成的建设工程电子文件和存放建设工程电子文件的文件夹命名规则如下：

按“工程”、“案卷”的次序，分别建立相应的目录结构。

建立工程级文件夹。工程级文件夹的命名规则是：工程编号。

在“工程级文件夹”之下建立案卷级文件夹。案卷级文件夹的命名规则是：序号。序号采用三位半角阿拉伯数字格式，不足三位时用前置“0”补足三位。

在“案卷级文件夹”下存放扫描文件。扫描文件的命名规则是：案卷序号-页码。页码采用三位半角阿拉伯数字格式，不足三位时用前置“0”补足三位，以数字“001”开始至最后文件止号结束。例如：01-001……001-N；02-001……002-N。

（5）扫描工作要有严格的登记、核检、审查制度，扫描过程要有完整、详细、规范的记录，确保档案原件和数字化档案信息的安全。扫描后的数字图像要清晰、完整，符合正常浏览方向，内容不得倒置；图像页面中出现因扫描而产生的影响图像质量的杂质，如黑点、黑线、黑框、黑边等要进行适当的去污处理，处理过程要遵循在不影响可读度的前提下展现档案原貌的原则；出现偏斜的图像应进行纠偏处理，图像偏度不得大于1度；发现不符合图像质量要求时，应重新进行扫描及图像处理。同时，发现遗漏及时补扫并插入对应位置。

**5、** 对于一份电子档案可能含有多页的，应将多页合成一个多页pdf文件，如果是文字类型的材料，要求pdf文件为OCR处理过的双层pdf，识别率应在95%以上。

**6、** 如果一份电子档案同时含有图纸和文本，应将这份档案分两份件存储，一份存图纸档案，一份存文本档案。

**7、** 对于声像档案，如果采用磁盘磁带存储的，将磁盘磁带内容采集为对应的通用格式电子文件，每份电子文件存成一个单独的电子档案。

**8、** 与电子档案的真实性、有效性、完整性相关的具有法律效力的数字签名应一同收集。

**9、** 收集好的电子文件经过数字化加工处理和数字签名后，采用AES加密算法对文件进行加密，同时将密钥写入对应的元数据文件，形成电子档案，移交城建档案管理机构管理。元数据文件格式和内容参见本规范。

## **二、 工程电子档案元数据的收集**

1、元数据收集分两类，一类为纸质电子档案，另一类为声像电子档案，元数据描述文件统一采用xml格式文件展示。

**2、**纸质电子档案的元数据信息需要同时包含案卷信息和文件信息，内容如下表：

案卷元数据信息：

|  |  |  |  |
| --- | --- | --- | --- |
| 元数据名称 | 描述 | 类型 | 限定条件 |
| 案卷题名 |  | 字符型 | 必填 |
| 案卷序号 | 案卷在当前工程下的顺序号 | 字符型 | 必填(长度为3，从001开始依次递增) |
| 编制单位 |  | 字符型 | 必填 |
| 责任者 |  | 字符型 | 选填 |
| 规格 | 案卷的背脊规格，单位为mm | 数字型 | 40 |
| 保管期限 |  | 字符型 | 长期、永久、短期 |
| 密级 |  | 字符型 | 秘密，机密，绝密 |
| 附注 | 案卷补充说明内容 | 字符型 | 选填 |

表2 案卷元数据表

3、文件的元数据信息：

|  |  |  |  |
| --- | --- | --- | --- |
| 元数据名称 | 描述 | 类型 | 限定条件 |
| 文件题名 |  | 字符型 | 必填 |
| 电子文件名称 | 电子文件存储名称 | 字符型 | 必填 |
| 文件序号 | 文件在案卷内的顺序号 | 字符型 | 必填(长度为3，从001开始依次递增) |
| 数量 | 文件包含页面的数量 | 数字型 | 必填 |
| 规格 | 文件的版式 | 字符型 | A4，A3，A2，A1，A0, A0+ |
| 形成时间 | 文件的形成时间 | 日期型 | 必填，格式：  YYYY-MM-DD |
| 文件格式 | 文件的扩展名 | 字符型 | 必填 |
| 文图号 | 文件的编号 | 字符型 | 选填 |
| 责任者 | 文件编制的所属责任者 | 字符型 | 选填 |
| 载体类型 |  | 字符型 | 文字（张），图纸（张） |
| 保管期限 |  | 字符型 | 短期，长期，永久 |
| 密级 |  | 字符型 | 秘密，机密，绝密 |
| 分辨率 |  | 字符型 | 必填 |
| 色彩值 |  | 字符型 | 彩色，灰度，黑白 |
| 加密算法标识 | 电子文件加密所采用的算法名 | 字符型 | 限定算法为AES算法 |
| 解密密钥 | 用来解密加密文件的密钥 | 字符型 | 必填。 |
| 附注 | 文件补充说明内容 | 字符型 | 选填 |

表 3 文件元数据表

**4、** 纸质电子档案的元数据格式如下：

<?xml version="1.0" encoding="gb2312"?>

<工程 工程名称="">

<案卷 案卷序号="" 案卷题名="" 编制单位="" 责任者="" 规格="" 保管期限="" 密级="" 附注="">

<文件 文件序号="" 文件题名=""电子文件名称="" 数量="" 规格="" 形成时间="" 文件格式="" 文图号="" 责任者="" 载体类型="" 保管期限="" 密级="" 分辨率="" 色彩值="" 加密算法标识="" 解密密钥="" />

</案卷>

</工程>

一个工程可以有多个案卷标记，每个案卷标记可以包含多个文件标记。

**5、**声像电子文件元数据信息如下表：

|  |  |  |  |
| --- | --- | --- | --- |
| 元数据名称 | 描述 | 类型 | 限定条件 |
| 文件题名 |  | 字符型 | 必填 |
| 文件序号 |  | 字符型 | 必填(长度为3，从001开始依次递增) |
| 文件类型 | 声像文件的类型 | 字符型 | 必填，可选值：照片、音频、视频 |
| 拍摄时间 | 文件包含页面的数量 | 日期型 | 必填，格式：  YYYY-MM-DD |
| 拍摄地点 | 文件拍摄的地点 | 字符型 | 必填 |
| 摄影者 |  | 字符型 | 必填 |
| 事由 |  | 字符型 | 必填 |
| 人物 |  |  | 必填 |
| 背景 |  |  | 必填 |
| 文件格式 | 文件的扩展名 | 字符型 | 必填 |
| 加密算法标识 | 电子文件加密所采用的算法名 | 字符型 | 选填，限定算法为AES算法 |
| 解密密钥 | 用来解密加密文件的密钥 | 字符型 | 如果加密算法标识填写，则该项必填。 |
| 附注 | 声像文件补充说明内容 | 字符型 | 选填 |

表4 声像文件元数据表

6、声像档案的元数据格式如下：

<?xml version="1.0" encoding="gb2312"?>

<工程 工程名称="">

<文件 文件序号="" 文件题名="" 文件类型="" 拍摄时间="" 拍摄地点="" 摄影者="" 事由="" 人物="" 背景="" 文件格式="" 加密算法标识="" 解密密钥="" 附注="" />

</工程>

一个工程可以包含多个文件标记。

## **三、工程电子档案的整理和移交**

**1、** 收集好的电子档案，以单位工程为单位刻盘，如文件过多可采用DVD光盘刻录，光盘目录结构如图1所示：

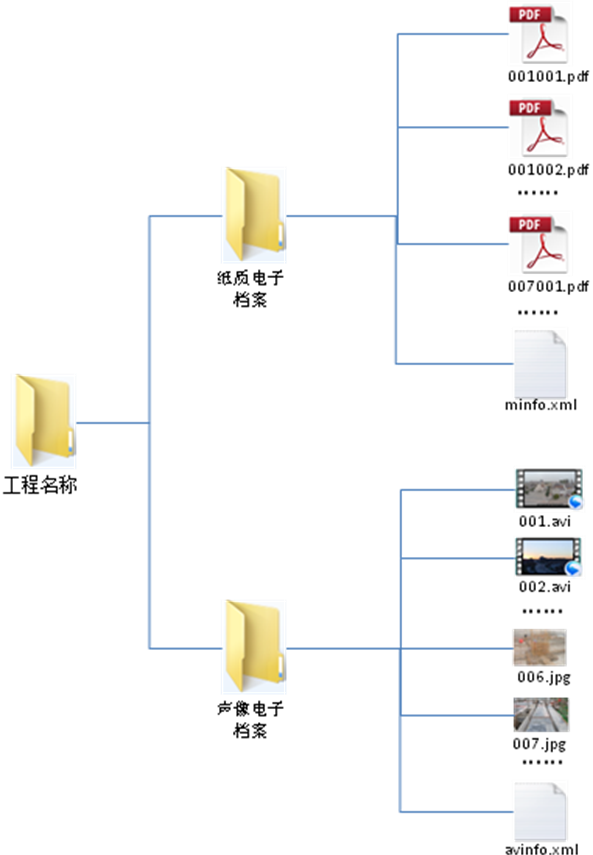


图1 电子档案移交结构图

注：（1）minfo.xml为固定命名，为电子档案元数据文件，以案卷顺序号在项目下创建案卷文件夹，文件夹内的文件以卷内目录的顺序号命名；

（2）avinfo.xml为固定命名，表示为声像文件元数据描述，声像电子文件以声像档案的顺序命名。

附录 建设工程档案数字化规范

扫描加工的程序与要求

**1.1.1** 选择扫描仪

根据文件、图纸幅面的大小选择相应规格的扫描仪或专业扫描仪进行扫描。A3，A4规格，纸质较好，可拆卸，可选择高速扫描仪来提高扫描效率，工程图纸可采用宽幅面扫描仪器进行扫描。

纸张状况较差，以及过薄、过软或超厚的档案，应采用平板扫描方式。

**1.1.2** 扫描色彩模式选择

纸质文件、图纸、照片扫描一般采用黑白二值、灰度和彩色三种方式。

一般文件、图纸，页面为黑白两色，并且字迹清晰、不带插图的档案，采用黑白二值模式进行扫描，红头文件或者页面有红章的档案采用彩色模式扫描。

黑白照片采用灰度模式扫描，灰度方式一般采用256灰阶。彩色照片采用彩色模式扫描。

**1.1.3** 扫描分辨率

扫描分辨率参数大小的选择，原则上以扫描后的图像清晰、完整、不影响图像的利用效果为准。

一般文件扫描分辨率采用300DPI，工程图纸文件采用分辨率300DPI，照片文件分辨率采用400DPI。

**1.1.4**  图像处理

（**1**）图像数据质量检验

对图像偏斜度、清晰度、失真度等进行检查。发现不符合图像质量要求时，应重新进行图像的处理。

由于操作不当，造成扫描的图像文件不完整或无法清晰识别时，应重新扫描。

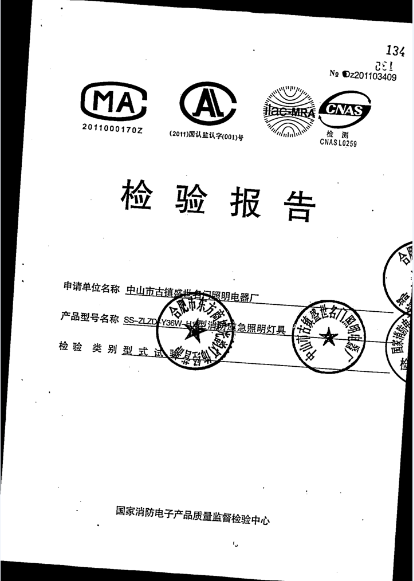
发现文件漏扫时，应及时补扫并正确插入图像。

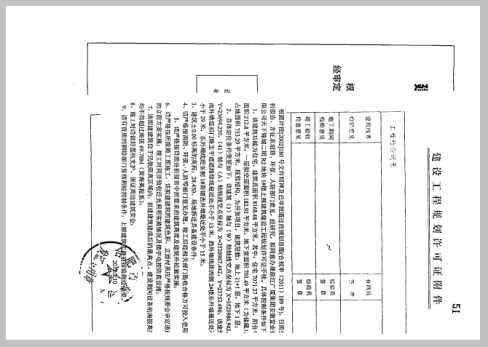
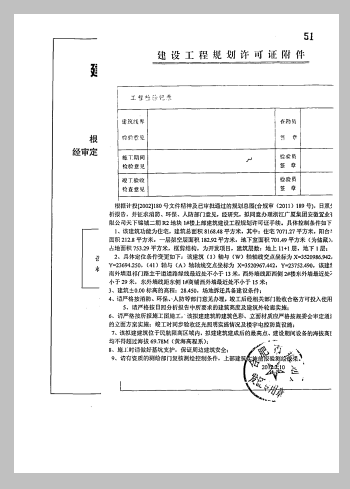
发现扫描图像的排列顺序与档案原件不一致时，应及时进行调整。

（**2**）纠偏

对出现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应进行旋转还原，以符合阅读习惯。

如下图：

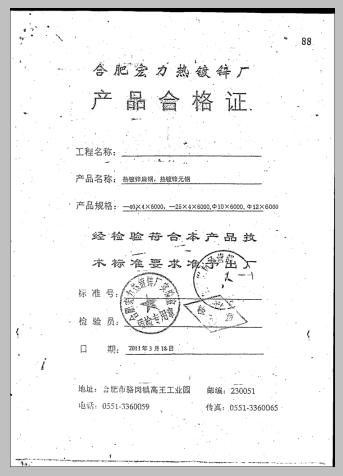
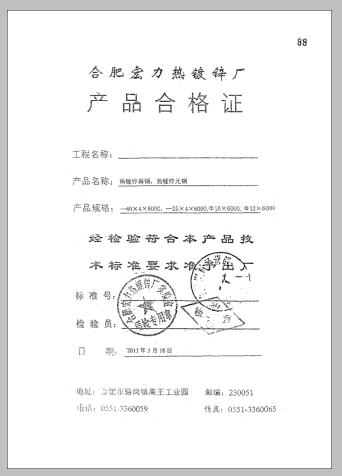
 

（**3**）去污

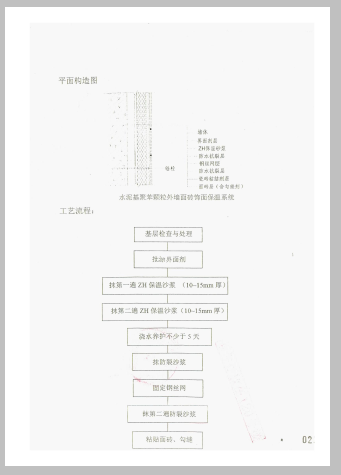
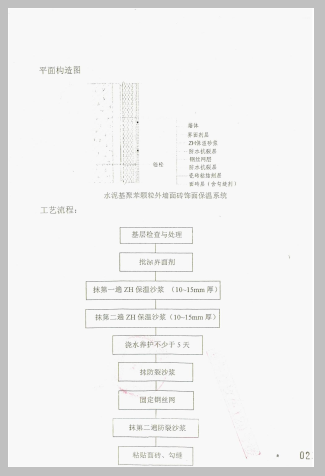
对图像页面中出现的影响图像质量的杂质如黑点、黑线、黑框、黑边等应进行去污处理。

处理过程中应遵循在不影响可懂度的前提下展现档案原貌的原则。如下图：

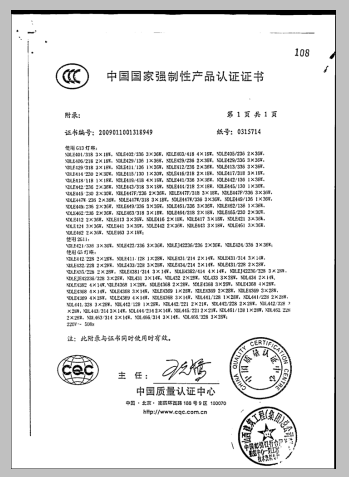
（**4**）裁边处理

扫描的图像应进行裁边处理，去除多余的白边，以有效缩小图像文件的容量，节省存储空间。如下图：

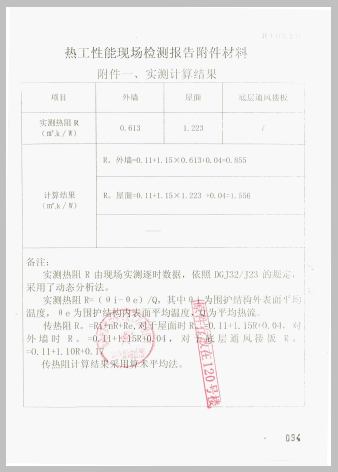
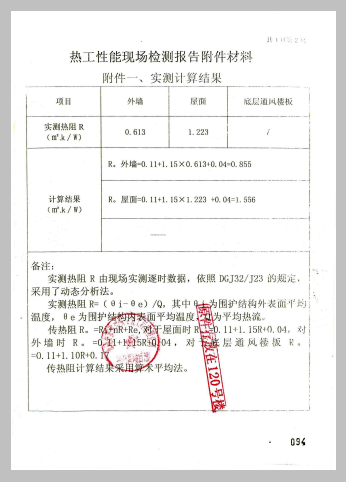
（**5**）去黑边

对有黑边的图像，需要去除黑边，同时不得改变图像的长、宽，如下图：

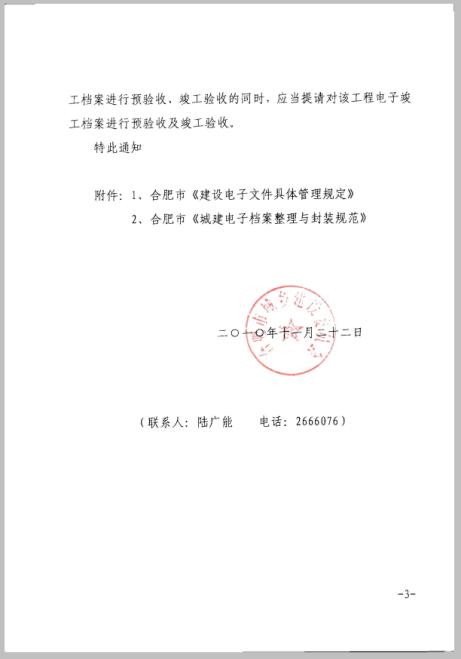
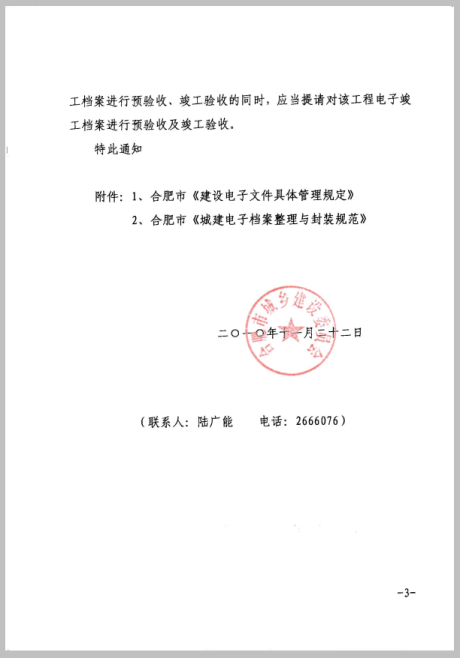
（6）亮度对比度调节

对扫描的图像显示模糊，与实际页面效果有差距的，需要进行亮度、对比度的调整，以增加图片的浏览效果，如下图：

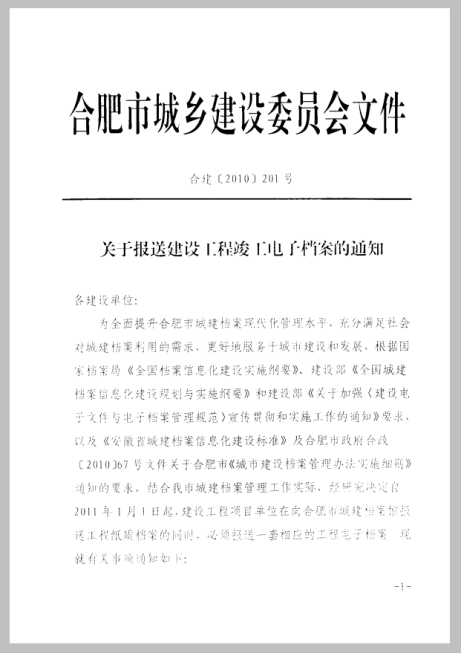
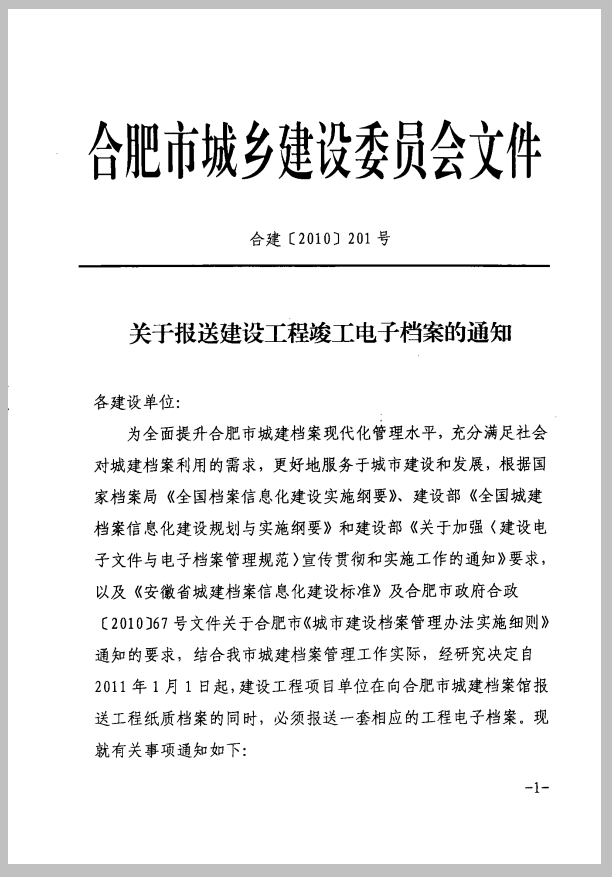
（**7**）红章优化

扫描图像红章不清晰的，需要进行优化操作，以红章内容清晰为准，如下图：

（8）文字效果调整

对图像文字模糊不清的图像，需要对文字效果进行调整，以图像文字清晰度不低于原纸张效果为准，如下图：

（9）红头文件优化

对于红头文件档案图像，页面不清晰、失真的，优化红色部分，以红色部分正常显示为准，如下图：

**1.1.5** 图像存储

采用黑白二值模式或者彩色模式扫描的一般文件，采用PDF格式存储，单份多页文件存储成多页PDF。

采用黑白二值模式扫描的工程图纸，采用TIFF，TIFF（G4）格式存储，可以采用 LZW 无损压缩，单份多页文件存储成多页TIFF。

采用灰度模式和彩色模式扫描的照片文件，一般采用JPEG格式存储。存储时的压缩率的选择，应以保证扫描的图像清晰可读的前提下，尽量减小存储容量为准则。

图像文件不是PDF的最终转换为多页双层PDF文件。

**1.1.6** 检查图形文件质量

对存入临时文件夹的图形文件进行浏览检查，检查文件格式是否正确。

**1.1.7** 电子文件导出

电子文件的命名及导出参见正文第二条。