



ICAO

Doc 9303

机读旅行证件 第七版, 2015年

第3部分 所有机读旅行证件的通用规范



经秘书长批准并由其授权出版

国际民用航空组织



| ICAO

Doc 9303

机读旅行证件 第七版, 2015年

第3部分 所有机读旅行证件的通用规范

经秘书长批准并由其授权出版

国际民用航空组织

国际民用航空组织分别以中文、阿拉伯文、英文、法文、俄文和西班牙文版本出版
999 Robert-Bourassa Boulevard, Montréal, Quebec, Canada H3C 5H7

下载文件和获取额外信息，请登录 www.icao.int/Security/FAL/TRIP。

Doc 9303 号文件 — 《机读旅行证件》
第 3 部分 — 所有机读旅行证件的通用规范
ISBN 978-92-9249-956-3

© ICAO 2016

保留所有权利。未经国际民用航空组织事先书面许可，不得将本出版物的任何部分复制、存储于检索系统或以任何形式或手段进行发送。

文件修改记录

Doc 9303 号文件第 3 部分

日期	编号	受影响的章节/页码
1/07/16	1	<p>第 28 页 新的 B 部分 — ISO 3166/MA 预留的其他代码 增加新代码：EUE</p> <p>第 29 页 D 部分 — 其他签发机关的代码 增加新代码：XBA 和 XEC</p>
16/09/16	2	第 2 页 3.1 语文和字符 — 增加“基于拉丁文”和“抄写或”
5/12/16		第 3 页 更新图 1（日期格式）
16/09/16		第 4 页 3.4 持有人姓名的书写惯例 — 增加“抄写或”
16/09/16		第 6 至 8 页 3.8 日期的表示方式 — 旅行者身份识别方案技术咨询组商定的建议
12/07/16		第 29 页 5.代表国籍、出生地、签发国/机构所在地和其他情况的代码 — D 部分 — 其他签发机构的代码 增加新代码：XCE（欧洲理事会）；和 XDC（南部非洲发展共同体）
16/09/16		第 30 至 34、36 和 37 页 6.建议各国使用的转写方法 — 根据统一码对字符重新排序，更正了统一码错误，增加了对乌克兰语的例外，并修订了阿拉伯文字的姓名
16/09/16		第 49 页 增加 ISO 3166/MA 作为参考文献
13/07/17		App A-4 页 更正第 4 行打字错误（所得的积用‘77620’代替先前的‘00000’）
16/09/16		App B-9, App B-12, App B-15 和 App B-16 页 B.5.关于机读区的建议。修订了阿拉伯字母转写的拼写。

本出版物中所用称谓和展示材料之方式，并不代表国际民航组织对任何国家、领土、城市或地区或其当局的法律地位，或就其边境或疆界的划分，表达了任何意见。

目录

1. 范围	1
2. 机读旅行证件的物理特性	2
3. 视读区(VIZ)	2
3.1 语文和字符	2
3.2 字体和字号	3
3.3 说明文字/域	4
3.4 持有人姓名的书写惯例	4
3.5 签发国或签发机构的表示方式	5
3.6 国籍的表示方式	5
3.7 出生地的表示方式	5
3.8 日期的表示方式	6
3.9 显示的持有人识别特征	8
4. 机读区(MRZ)	16
4.1 机读区的用途	16
4.2 机读区的特性	16
4.3 机读区的限制条件	17
4.4 打印规格	17
4.5 机读要求和有效阅读区	18
4.6 持有人姓名的书写惯例	18
4.7 签发国或签发机构和持有人国籍的表示方式	20
4.8 日期的表示方式	20
4.9 机读区内的校验数位	20
4.10 机读区的特征	21
4.11 机读区的质量规范	21
5. 代表国籍、出生地、签发国/机构所在地和其他情况的代码	22
6. 建议各国使用的转写方法	30
7. 偏差	38
7.1 运作经验	38
7.2 偏差列表做法	38

7.3 方法.....	39
7.4 公布.....	44
7.5 技术规范.....	45
8. 参考文献(规范性).....	49
第 3 部分附录 A 校验数位计算方法示例(资料性).....	App A-1
第 3 部分附录 B 机读旅行证件中阿拉伯文字的转写(资料性).....	App B-1
B.1 阿拉伯文字.....	App B-1
B.2 机读旅行证件中的阿拉伯文字.....	App B-1
B.3 对视读区的建议.....	App B-3
B.4 机读区内字符的转写.....	App B-5
B.5 关于机读区的建议.....	App B-6
B.6 机读区中的逆向转写.....	App B-15
B.7 计算机程序.....	App B-17
B.8 参考文献(资料性).....	App B-20

1. 范围

Doc 9303 号文件第七版对国际民航组织关于机读旅行证件的规范进行了结构调整。在 Doc 9303 号文件的这一新版本中并未纳入对规范的大量修改，而是对内容进行了重新编排，形成一套关于 1 型机读官方旅行证件（TD1）、2 型机读官方旅行证件（TD2）和 3 型机读旅行证件（TD3）以及签证的规范。这套规范由多个不同文件组成，其中对一般规范（适用于所有机读旅行证件）以及与特定尺寸的机读旅行证件相关的规范做了分类。

Doc 9303 号文件第 3 部分是以 Doc 9303 号文件第 1 部分第六版第 1 卷 — 《机读护照 — 带有按光学字符识别格式存储的机读数据的护照》（2006 年）以及 Doc 9303 号文件第 3 部分第三版第 1 卷 — 《机读官方旅行证件 — 带有按光学字符识别格式存储的机读数据的机读旅行证件》（2008 年）为基础编制的。

第 3 部分明确了通用于 TD1 型、TD2 型和 TD3 型机读旅行证件（MRTDs）的规范，包括为确保通过视读和机读（光学字符识别）手段实现全球互操作性所需遵守的规范。适用于每种证件类型的具体规范，见 Doc 9303 号文件第 4 部分至第 7 部分。

阅读第 3 部分应结合以下部分进行阅读：

- 第 1 部分 — 引言；
- 第 2 部分 — 机读旅行证件的设计、制作和签发的安全性规范；

以及与特定尺寸旅行证件相关的部分：

- 第 4 部分 — 机读护照（MRPs）和其他 TD3 型机读旅行证件规范；
- 第 5 部分 — TD1 型机读官方旅行证件（MROTds）规范；
- 第 6 部分 — TD2 型机读官方旅行证件（MROTds）规范；和
- 第 7 部分 — 机读签证。

这些规范同时还适用于包含非接触式集成电路的机读旅行证件，即电子机读旅行证件（eMRTDs）。仅适用于电子机读旅行证件的规范载于 Doc 9303 号文件的以下部分：

- 第 9 部分 — 生物特征识别技术的运用和机读旅行证件的电子数据存储；
- 第 10 部分 — 在非接触式集成电路（IC）中存储生物特征和其他数据的逻辑数据结构（LDS）；
- 第 11 部分 — 机读旅行证件的安全机制；和
- 第 12 部分 — 机读旅行证件的公钥基础设施。

2. 机读旅行证件的物理特性

签发国和签发机构可自由选择用于制作旅行证件的材料。但是，所选择的材料不得对机读旅行证件中的任何其他组成部分产生不利的影响，并且机读旅行证件在其整个有效期的正常使用中应满足下列要求：

- **变形** 机读旅行证件所用材料应具有可弯性（不起皱折），即由于正常使用造成的变形通过读取装置可以使其变平，而不至于影响其使用或阅读器的功能。
- **毒性** 根据[ISO/IEC 7810]的规定，在正常使用过程中，机读旅行证件不得出现毒性危险。
- **抗化学品** 机读旅行证件应能抵抗正常处理和使用过程中出现的化学影响，出于安全原因加入的化学敏感性除外。
- **温度稳定性** 机读旅行证件应在-10℃至+50℃（14°F至122°F）的操作温度范围内保持机读性。机读旅行证件在暴露于-35℃至+80℃（-31°F至176°F）的温度下不得失去其功能。
- **湿度** 根据[ISO/IEC 7810]的规定，机读旅行证件在湿球温度最高为25℃（77°F）、空气相对湿度为5%至95%之间，应保持机读性。机读旅行证件储存或暴露在0%至100%（不冷凝）的相对湿度下不应失去其可靠性。
- **光** 根据[ISO/IEC 7810]的规定，机读旅行证件在正常使用中暴露在光下不应损坏。

3. 视读区(VIZ)

机读旅行证件视读区由针对视读设计的强制性和选择性数据元素构成。选择性数据元素以及强制性数据元素可以在保持足够的统一性以确保所有机读旅行证件全球互操作性的同时，满足签发国和签发机构的不同需要。

3.1 语文和字符

视读区的数据应该用拉丁字母字符，即A到Z，和阿拉伯数字，即1234567890表示。可以使用变音符。第6.A节 — “多国基于拉丁语的字符转写”中所列出的基于拉丁语的国家字符，如Ð和β，也可以用于视读区中，无需转写。当强制性数据元素用不使用拉丁字母的本国语文表示时，还应提供抄写或转写。

3.3 说明文字/域

应使用说明文字标明视读区内强制性数据元素使用的所有域，Doc 9303 号文件第 4 部分至第 7 部分中有关每种证件类型的数据元素目录另有规定的除外。

说明文字可以使用签发国的官方语文或签发机构的工作语文。当这些语文使用拉丁字母时，应采用直体字打印说明文字。

如果签发国的官方语文或签发机构的工作语文不是英文、法文或西班牙文，还应在所印的说明文字之后加上一个斜杠字符 (/) 以及说明文字的英文、法文或西班牙文译文。第二种语文应采用斜体字。

如果签发国的官方语文或签发机构的工作语文是英文、法文或西班牙文，则签发国或签发机构应使用另外两种语文中的一种在斜杠字符 (/) 之后打印出说明文字。第二种语文应采用斜体字。

说明文字应采用清晰的直线型字体打印，字号应在 1.0 毫米至 1.8 毫米 (0.04 英寸至 0.07 英寸) 之间。

当某个域未被使用时，不应在旅行证件上显示说明文字。

3.4 持有人姓名的书写惯例

持有人姓名通常用两部分表示：主要标识符和次要标识符。

签发国或签发机构应确定姓名的哪一部分是主要标识符，它可以是姓、婚前姓或夫姓、主名、别名，或者在某种情况下当持有人姓名无法分为两部分时可以是整个姓名。这一部分应登记到视读区的主要标识符域内。建议使用大写字母，但在有诸如“von、”“Mc”或“de la”等前缀的情况下除外。在这些情况下，混合使用大小写是适宜的。

姓名的余下部分为次要标识符，它可以是名、字、教名、首字母或任何其他辅助名。这些名字应写在视读区次要标识符域内。建议全部使用大写字母。

如果姓名是使用单一一个域，那么次要标识符和主要标识符之间应使用一个逗号 (,) 分开。如果是使用多个域，不需要使用逗号。

前缀和后缀，包括头衔、专业和学术资格、荣誉、奖励和世袭地位不应包括在视读区内。但是，如果签发国或签发机构认为这样的前缀或后缀是姓名法定的一部分，那么前缀或后缀可以出现在视读区。数字字符不应写在视读区的姓名域中，但是如果使用数字字符在签发国是合法的命名惯例，这些字符应该用罗马数字表示。任何前缀、后缀或罗马数字应登记到次要标识符域内。

在视读区可以使用本国字符。如果本国字符不是基于拉丁语，那么应将其抄写或转写成拉丁字符。

3.5 签发国或签发机构的表示方式

当签发国或签发机构和/或签发部门或机关所在地的名称用非拉丁字符的本国语文表示时，国名或其他地点名称应以本国语文的形式出现，同时还应：

- 转写成拉丁字符；或
- 翻译成使该名称可能更普遍为国际社会所熟知的一种或多种语文（至少其中一种必须是英文、法文或西班牙文）。

不同语文的名称应该用一个斜杠字符 (/) 分开，后面至少跟一个空格。

当签发国或签发机构或签发部门或机关所在地的名称用拉丁字母的一种语文表示，但是如果将名称翻译成另外一种语文或多种语文（特别是英文、法文或西班牙文）更为国际社会所熟悉时，用本国语文写入的名称应配有一种或多种译文。不同语文的名称应该用一个斜杠字符 (/) 分开，后面至少跟一个空格。

3.6 国籍的表示方式

对于持有人国籍域为视读区强制性元素的证件，签发国应自行决定使用三字代码（见第5节）或全称表示持有人国籍。

当国籍采取全称形式且签发国的本国语文或签发机构的工作语文为非拉丁字符的语文时，国籍应以本国语文的形式出现，同时还应：

- 转写成拉丁字符；或
- 翻译成使该国籍可能更普遍为国际社会所熟知的一种或多种语文（至少其中一种必须是英文、法文或西班牙文）。

不同语文的国籍应该用一个斜杠字符 (/) 分开，后面至少跟一个空格。

当签发国的本国语文或签发机构的工作语文使用拉丁字母，但是如果将国籍翻译成另外一种语文或多种语文（特别是英文、法文或西班牙文）更为国际社会所熟悉时，用本国语文写入的国籍应配有一种或多种译文。不同语文的国籍应该用一个斜杠字符 (/) 分开，后面至少跟一个空格。

3.7 出生地的表示方式

可自由选择是否纳入出生地。如果纳入出生地，可用镇、市、城郊和/或国家加以表示。

如果纳入镇、市或城郊的名称且签发国的本国语文并不使用拉丁字符，则镇、市或城郊的名称应以本国语文的形式出现，同时还应：

- 转写成拉丁字符；或
- 翻译成使该出生地可能更普遍为国际社会所熟知的一种或多种语文（至少其中一种必须是英文、法文或西班牙文）。

用不同语文表示的镇、市或城郊的名称应该用一个斜杠字符 (/) 分开，后面至少跟一个空格。

当签发国的本国语文或签发机构的工作语文使用拉丁字母，但是如果将镇、市或城郊的名称翻译成另外一种语文或多种语文（特别是英文、法文或西班牙文）更为国际社会所熟悉时，用本国语文写入的镇、市或城郊的名称应配有一种或多种译文。用不同语文表示的镇、市或城郊的名称应该用一个斜杠字符 (/) 分开，后面至少跟一个空格。

如果纳入国家名称，则应该用国际民航组织的三字代码（见第 5 节）表示，出生国没有代码的情况除外，在这种情况下，应该用全称写入出生国的名称，并且还应遵守上文所规定的针对镇、市和城郊的翻译和转写要求。

注：在选择纳入或省略出生地时，旅行证件的签发国或签发机构应该考虑到当前与该国或该领土相关的任何政治敏感因素，以及其是否是被其他国家的签证发放机关所认可的国家或领土。

3.8 日期的表示方式

机读旅行证件视读区中的日期应根据公历按照下列方式登记：

日

日应该用两位数表示，即 1 到 9 日的前面应加零。可以在该数字后面的月份之前留一个空格，也可以不留空格，直接跟上月份。

月

月份可以用签发国的本国语文或签发机构的工作语文完整打印出来，或使用缩写，最多使用四个字符位置。

如果签发国的本国语文或签发机构的工作语文不是英文、法文或西班牙文，则月份后面应跟一个斜杠字符 (/) 和用下表中所列的三种语文中的一种标明月份或月份的缩写（最多占用四个字符位置）。

如果签发国的本国语文或签发机构的工作语文是英文、法文或西班牙文，签发国或签发机构还可以在斜杠字符 (/) 后面加上其他两种语文中的一种（如表 1 所示）。

根据签发国或签发机构自行作出的选择，月份也可以用数字形式打印出来，特别是如果这样做可能便于那些不使用公历的国家使用机读旅行证件。这时，日期将写成 DDnMMnYY 或者 DDnMMnYYYY，其中 n 是一个空格或一个圆点。

表 1 月份的英文、法文和西班牙文缩写

月份	英文	法文	西班牙文
一月	JAN	JAN	ENE
二月	FEB	FÉV	FEB
三月	MAR	MARS	MAR
四月	APR	AVR	ABR
五月	MAY	MAI	MAYO
六月	JUN	JUIN	JUN
七月	JUL	JUIL	JUL
八月	AUG	AOÛT	AGO
九月	SEP	SEPT	SEPT
十月	OCT	OCT	OCT
十一月	NOV	NOV	NOV
十二月	DEC	DÉC	DIC

年

年份用最后两位数或四位数表示，前面留有一个空格，或者也可以直接跟在月份后面，不留空格。两种格式都是可以接受的。

当月份用数字形式表示时，签发国或签发机构可以使用两位或四位数表示年份，并用一个空格或圆点将月份和年份分开。

注：鼓励各国对于所有日期格式使用四位数代表年份。

示例：

用意大利文签发且将月份译成法文的机读旅行证件数据页上的 1942 年 7 月 12 日可表示为：

12nLUGn/JUILn42

其中，n = 一个空格，即 12 LUG /JUIL 42

或

12 July 42（仅使用英文）

或

12JUIL42（使用法文缩写）

或

12JUL 42（使用英文或西班牙文缩写）

或

12 07 42 或 12.07.42（使用数字格式）

或

12 07 1942 或 12.07.1942（使用数字格式，其中年份为四位数）。

未知的出生日期在出生日期一无所知的情况下，该数据元素应该用签发国或签发机构所使用的出生日期格式来表示，用 Xs 表示日期的未知元素（数字和/或字母）。

举例：

XXnXXnXX
XXnXXnXXXX
XXnXXXnXX

其中 n = 一个空格或（如果使用数字格式）一个圆点。

如果出生日期中只有一部分不清楚，则仅该部分（日、月、年）应按照签发国或签发机构所使用的日期格式用 Xs 来表示。

3.9 显示的持有人识别特征

Doc 9303 号文件明确了各类机读旅行证件必须在视读区中显示的强制性的和选择性的持有人识别特征，如人脸图像、签名或常用标记和/或单枚指纹，以及这些识别特征的位置、尺寸和缩放比例。

3.9.1 显示的人脸图像

为确保与脸部识别系统的兼容性，在拍摄照片用作机读旅行证件持有人的人脸图像时必须遵守以下指导原则和图示。

所显示的人脸图像应展现机读旅行证件合法持有人的真实面貌，不得通过数字手段修改或美化图像，从而使人物面貌发生任何变化。照片应在机读旅行证件签发日期之前六个月内所拍摄的。

3.9.1.1 姿态

照片应为头部和肩部的特写，拍照对象应正面直视镜头，两眼应清晰可见，保持正常的表情，嘴唇闭合。



姿态应该是两眼中心之间的想象平行线与长方形照片的上缘和下缘平行。

人脸图像从头顶（不考虑任何头发）到下颚、从鼻子到耳朵都应在焦点上。



如果需要显示一只耳朵的更多细节（有时称为“半剖面”），脸部应与想象的平面形成适当的夹角，以便显示该只耳朵的细节，同时保持与显露的这只耳朵相对的另半边脸的面部细节完全可见。

脸部的两个边缘必须清晰可见。拍照对象不得像拍肖像照一样侧过头。

脸部应精确聚焦，并且没有墨迹或折痕等瑕疵。

眼睛必须睁开，不得有头发遮住眼睛。

如果拍照对象佩戴眼镜，照片必须清楚显现眼睛，眼镜上不得出现反光。眼镜不得装佩有色镜片。可能的话，避免佩戴粗框眼镜，并确保镜框不遮盖眼睛的任何部分。只有常带眼镜，照片上才应出现眼镜。



除签发国特别批准的情况，不得佩带头部覆盖物。这类情况可能出于宗教、医疗或文化上的原因。人脸必须从发顶线到下巴及两耳前部均清晰可见。



不允许出现遮住脸部的遮盖物、头发、头饰或人脸装饰物。



签发国应自行决定人脸装饰物（例如鼻环、点缀物）可以多大程度上出现在标准照上。只有永久性佩带的人脸装饰物才应出现在照片上。

婴儿的人脸图像应该遵循与成人一样的规范。最好为婴儿在直立的状态下拍照，但是躺在白色或浅素色毛毯上为其拍摄人脸图像也是可以接受的。或者可把婴儿放在儿童椅上，但其头的后面应该有白色或浅素色的背景。婴儿的眼睛应睁开，不得看到协助孩子的手。



照片中不得出现其他人或物体。



3.9.1.2 布光、曝光和色彩平衡

在采集人脸图像时，照明应充足、均匀，以确保在脸部或背景上无阴影或反光。



拍照对象的双眼一定不能显示红眼。



照片必须有适当的亮度和对比度。





所显示的标准照应为黑白灰度[或黑白]等级或呈现持有人脸部的真实色彩。如果照片是彩色的，在布光和拍照时必须进行色彩平衡，使皮肤呈现真实色调。

应使用均匀淡色的背景，以便与人脸和头发形成对比。对于彩色标准照，建议采用淡蓝色、米色、淡褐色、暗灰色或白色作为背景。

3.9.1.3 标准照的放置位置

人脸图像应居中放在 V 区内，头顶（不考虑任何头发）应尽量靠近机读旅行证件的上部边缘。人脸图像中的头顶到下颚部分应占 V 区的最长尺寸的 70%到 80%，并保持持有人脸部头顶到下颚和耳朵到耳朵之间的长宽比。这个 70%到 80%的要求可能意味着需对照片进行裁切，以致看不到全部头发。

3.9.1.4 数字打印复制

无论是以纸质格式还是数字格式提供的标准照，都必须以数字形式打印在机读旅行证件上。签发国或签发机构应采取必要措施，确保所显示的标准照是防伪造和防替换的。

数字复制质量 数字复制应能再现证件合法持有人的准确、可以识别的图像。所采集的原始标准照的质量应至少相当于可以接受的最低质量的照片（分辨率相当于每毫米 6-8 线对）。为在数字复制中达到这种类似的图像质量，必须认真注意图像的采集、处理、数字化、压缩和打印以及制作标准照的工艺，包括机读旅行证件的最后加工。

边界 不得使用边界或边框为数字印刷的复制件添加轮廓。

与背景安全处理共存 数字打印的复制件应与 V 区内的背景安全处理共存，即背景安全印刷不应影响所显示的标准照的正常观看，反之亦然。

与机读旅行证件最后加工处理共存 显示的标准照应与最后加工处理共存，即最后加工处理不应影响所显示的标准照的正常观看，反之亦然。

3.9.1.5 向签发机关提交标准照

如果向签发机关提供印刷版的标准照，则照片，无论其是通过传统的照相技术还是通过数字手段制作的，应该打印在质量优良或质量达到照片水平的纸张上，并且尺寸应为所规定的最大尺寸。

如果向签发机关提供数字版的标准照，则必须遵守签发机关规定的要求。

所提交的标准照的尺寸应该为 45.0 毫米×35.0 毫米（1.77 英寸×1.38 英寸）。这一尺寸可提供足够的分辨率，便于将照片缩放至用在机读旅行证件上所需的大小，同时具有进行人脸识别的足够分辨率。

3.9.1.6 遵循国际标准

照片应遵循[ISO/IEC 19794-5]中所列出的相关规定。

3.9.2 显示的签名或常用标记

签名或常用标记显示在 IV 区，其可接受性由签发国或签发机构自行决定。所显示的签名或常用标记应是在机读旅行证件上创建的原件、原件数字打印的复制件，或者在 Doc 9303 号文件第 4 部分至第 7 部分所确定的、与制作不同类型的机读旅行证件具体相关的规范允许的情况下，可将签名或常用标记放在可牢固贴在机读旅行证件上的基材上。签发国或签发机构应采取必要的措施，确保所显示的签名或常用标记是防伪造和防替换的。所显示的签名或常用标记应满足下列要求。

定位 所显示的签名或常用标记在显示时应使 A-尺寸与机读旅行证件的参考（较长）边缘相平行，如图 2 所示。

尺寸 所显示的签名或常用标记的尺寸应使人的眼睛可以辨认（即尺寸缩小比例不超过百分之五十），而且保持原始签名或常用标记的长宽比（A-尺寸比 B-尺寸）。

使用数字打印时复制件的缩放比例 如果所显示的签名或常用标记被放大或缩小，应保持原始签名或常用标记的长宽比（A-尺寸比 B-尺寸）。

使用数字打印时复制件的裁切 签发国或签发机构应采取措施避免裁切或将裁切减少到最低程度。

颜色 所显示的签名或常用标记的颜色应与背景形成明显的对比。

边界 不允许使用边界或边框为所显示的签名或常用标记添加轮廓。

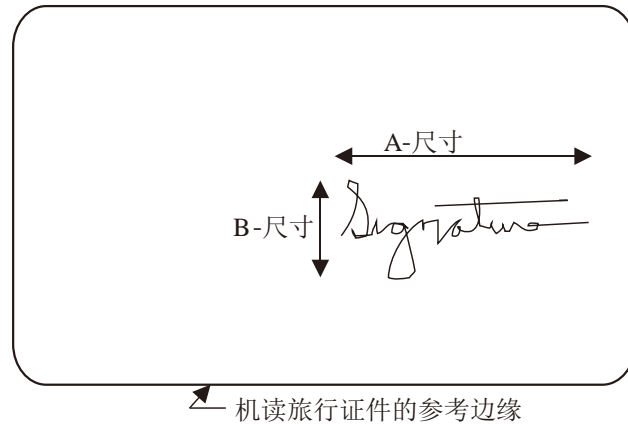


图2 所显示的签名或常用标记的定位

3.9.3 单枚指纹的显示

如果签发国或签发机构要求显示单枚指纹，它应是持有人在机读旅行证件基材上创建的原件，或者更可能是原件的数字打印复制件。签发国或签发机构应采取必要的措施，确保单枚指纹是防伪造和防替换的。单枚指纹应满足下列要求。

定位 所显示的单枚指纹的 A-尺寸（宽）应与机读旅行证件的参考边缘相平行，如图 3 所示。手指的上部应是单枚指纹距离机读旅行证件参考边缘最远的那个部分。（见 Doc 9303 号文件第 6 部分的图 10 和图 12。）

尺寸 所显示的单枚指纹应是原始指纹的 1 比 1 的复制件（A-尺寸对 B-尺寸）。

使用数字打印时复制件的缩放比例 不允许缩放单枚指纹的尺寸。

使用数字打印时复制件的裁切 签发国或签发机构应采取措，避免裁切或将裁切减少到最低程度。

颜色 所显示的单枚指纹的颜色应与背景形成明显的对比。

边界 不允许使用边界或边框为所显示的单枚指纹添加轮廓。

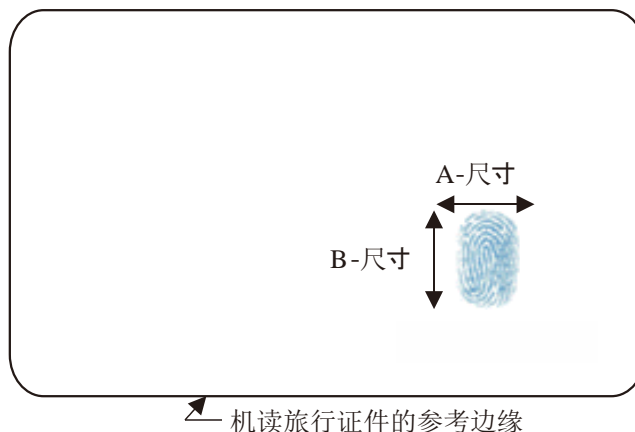


图 3 所显示的单枚指纹的定位

4. 机读区(MRZ)

4.1 机读区的用途

根据 Doc 9303 号文件制作的机读旅行证件包括一个机读区，以便于查验旅行证件，并减少旅行过程中行政程序所占用的时间。另外，机读区还为视读区的信息提供验证，并可用来为数据库查询提供搜索字符。它也可用来采集到港和离港的登记数据，或简单地显示数据库中的现有记录。

机读区以针对各类机读旅行证件的标准格式提供一套基本的数据元素，所有的接受国，不管其国家的文字或习惯如何，都可以使用这一格式。

机读区的数据经格式化后，全世界具有标准能力的机器都可以阅读。必须强调，机读区是为依照机读旅行证件的国际标准在国际上使用的数据而预留的。机读区的数据表示法不同于视读区的数据表示法。

4.2 机读区的特性

机读区的数据必须是视读和机读的。数据的表现方式必须符合一个共同的标准，这样所有按照 Doc 9303 号文件配置的机读设备都能识别每个字符，并采用标准协议（例如美国标准信息交换代码）进行通信。这种协议要与接受国规定的技术基础设施和处理要求相兼容。

为满足这些要求，规定了 OCR-B 字体作为在机读区存储数据的表现方式。在此界定的机读区被认可为进行全球信息交换必不可少的机读技术，因此在所有类型的机读旅行证件中都必须采用。

4.3 机读区的限制条件

机读区仅可使用的字符是各国都可以使用的一套公用字符（图 4）。国家特有字符一般只出现在应用这些字符的国家计算机处理系统中，在全球范围内并不使用这些字符。因此，它们不得出现在机读区内。

在机读区内，不允许使用变音符。尽管变音符对于区分姓名是有用的，但在机读区内使用会干扰机读设备，从而降低数据库搜索准确度，减慢旅客的通关速度。

在机读区内，数据可用字符位置的数量有限，并因机读旅行证件的类型而异。插入到机读区的数据元素的长度，必须符合 Doc9303 号文件第 4 部分至第 7 部分的数据元素目录中所规定各个域的尺寸。

在某些情况下，机读区内的名字所采取的表现形式不一定与视读区的相同。在视读区内，可以使用非拉丁字符和国家特有字符来更准确地代表签发国或签发机构的文字性数据。此类字符不得用在机读区内。

4.4 打印规格

机读数据应使用 OCR-B 字体、尺寸 1 和固定笔画宽度的字符打印，固定的宽度间距为 2.54 毫米（0.1 英寸），即水平打印密度为每 25.4 毫米（1.0 英寸）10 个字符。打印字符应限制在图 4 规定的字符范围内。



图 4 用于机读旅行证件的选自[ISO 1073-2]的 OCR-B 字符子集

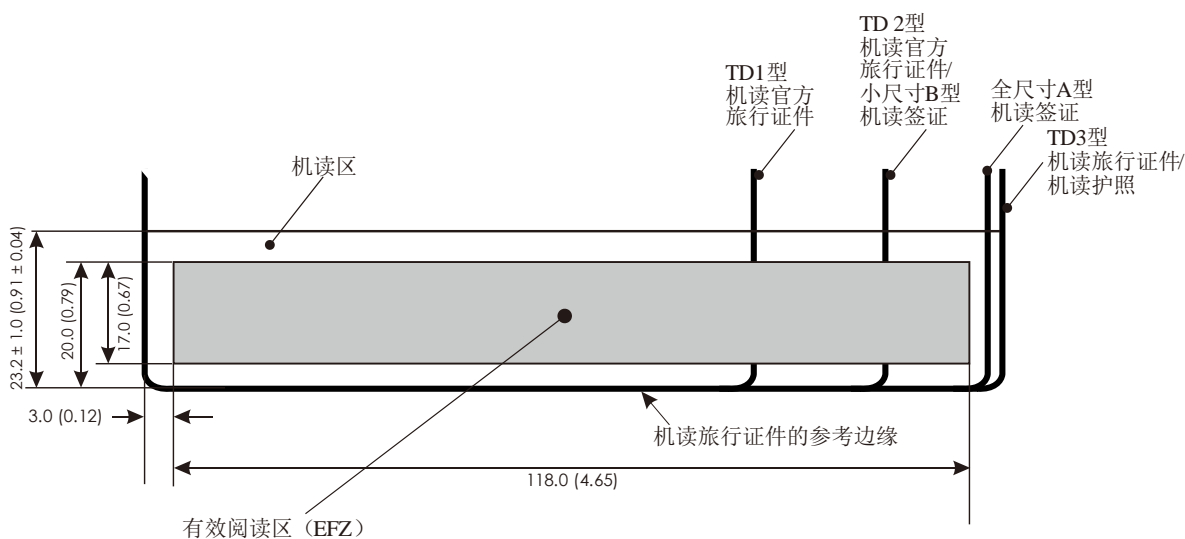
注：仅作为说明之用 — 所显示字符要比实际尺寸大。

4.5 机读要求和有效阅读区

有效阅读区 按照最大尺寸的机读旅行证件规定了一个固定尺寸的阅读区（大小为 17.0 毫米×118.0 毫米（0.67 英寸×4.65 英寸）的有效阅读区（ERZ）），以便用一个机读者就可以阅读所有尺寸的机读旅行证件。有效阅读区的位置如图 5 所示。提供有效阅读区不是为了增加与制作不同类型机读旅行证件相关的第 4、5、6 和 7 部分中规定的打印位置的公差。有效阅读区是为了兼顾到由于人工放置机读签证（MRV）和打开机读护照内页阅读时出现的册页鼓风效应而造成的不同情况。有效阅读区还兼顾到阅读有两行或三行机读数据的机读官方旅行证件的情况。

为了应对使用诸如复印机等设备对旅行证件安全所带来的威胁，在机读区允许设置安全特征。根据[ISO 1831]的规定，这种安全特征不得影响在 B900 范围内准确地阅读 OCR 字符。虽然根据 4.2 中的规定，OCR 字符必须明显可见，以确保所有机读旅行证件，包括那些在机读区带有安全特征的证件，都能被成功地阅读，但在机读区内的 OCR 字符应至少能在光谱的近红外区部分（即[ISO 1831]中规定的 B900 谱带）被机器阅读。

注：所示的有效阅读区（ERZ）的尺寸是以可用于所有机读旅行证件的标准化有效阅读区为根据的，以便能够使用单一机器阅读器。



尺寸单位为毫米（括号内的尺寸单位为英寸）

未按比例

图 5 机读旅行证件有效阅读区示意图

4.6 持有人姓名的书写惯例

为实现全球互操作性，机读区内的主要和次要标识符应使用图 4 所示的大写的 OCR-B 字符打印，不加变音符，并且应符合可用的字符位置的数量。因此，机读区内姓名的表达方式不同于视读区内姓名的表达方式。签发国或签发机构应仅使用允许的 OCR-B 字符对本国字符进行转写和/或按照 Doc 9303 号文件针对特定尺寸旅行证件的第 4 部分至第 7 部分中所规定的要求进行截取。第 6 节载有拉丁语、西里尔语和阿拉伯语系最常用字符的转写表。

使用拉丁字符转写（如果适用的话）的主要标识符应按照 Doc 9303 号文件针对特定尺寸旅行证件的第 4 部分至第 7 部分中所规定的要求写入机读区。主要标识符后面应跟着两个填充符（<<）。使用拉丁字符转写（如果适用的话）的次要标识符应从紧跟着这两个填充符的字符位置开始写入。

如果主要或次要标识符有一个以上的姓名成分，成分之间应该用一个填充符（<）分开。

填充符（<）应紧跟着最后的次要标识符插入（或如果姓名只有主要标识符的话，紧跟着主要标识符插入），一直插到机读行的最后一个字符位置。

姓名域的字符位置数量有限，并因机读旅行证件的类型而异。如果采用上述程序写入相关机读行的主要和次要标识符超过这个可用的字符位置，应按照 Doc 9303 号文件针对特定尺寸旅行证件的第 4 部分至第 7 部分中所规定的程序进行截取。在其他所有情况下，则不得截取姓名。

姓名截取的例子见 Doc 9303 号文件针对特定尺寸旅行证件的第 4 部分至第 7 部分。

前缀和后缀，包括头衔、专业和学术资格、荣誉、奖励和世袭地位（如博士、爵士、小、老、二世和三世），不得包括在机读区内，除非签发国认为它们是姓名的法定的一部分。在这些情况下，前缀或后缀应作为次要标识符的成分出现。

数字字符不得在机读区的姓名域中使用。

在机读区内不允许有标点符号字符。当它们作为姓名的一部分出现时，应对它们做如下处理：

撇号：

应省掉。在机读区，由撇号分开的姓名成分应结合到一起，不得插入填充符替代它的位置。

例子 视读区： D'ARTAGNAN
 机读区： DARTAGNAN

连字符：

如果在两个名字成分之间有连字符，在机读区应该用一个填充符（<）代替它。（即用连字符连接的名字应表示为不同的成分）。

例子 视读区： MARIE-ELISE
 机读区： MARIE<ELISE

逗号：

如果在视读区使用逗号将主要标识符和次要标识符分开，应在机读区省掉逗号。在机读区，主要标识符和次要标识符之间应使用两个填充符（<<）分开。

例子 视读区： ERIKSSON, ANNA MARIA
 机读区： ERIKSSON<<ANNA<MARIA

另外，如果在视读区使用逗号将两个姓名成分分开，该逗号在机读区应显示为一个填充符（<）。

例子 视读区： ANNA, MARIA
 机读区： ANNA<MARIA

其他标点符号字符：

所有其他标点符号字符均应从机读区略去（即在机读区它们的位置不得插入填充符）。

4.7 签发国或签发机构和持有人国籍的表示方式

在机读区，应使用第 5 节所列的三字代码填写签发国或签发机构和国籍域。

4.8 日期的表示方式

在机读旅行证件机读区内的日期应以六位数表示，其中包括年份的最后两位数（YY），后面紧跟着月份的两位数（MM）和日的两位数（DD）。结构如下：YYMMDD。

根据这个格式，1942 年 7 月 12 日将被表示为：420712。

如果出生日期的全部或部分不清楚，相关字符位置应该用填充符（<）填充。

4.9 机读区内的校验数位

校验数位由从一系列其他数字计算得出的一个数位组成。机读区内的校验数位是根据机读区内特定的数字数据元素计算而来的。通过校验数位，阅读器可以验证机读区内的数据是否得到正确的解读。

在机读旅行证件中，采用了一种特殊的校验数位计算方法。校验数位应在模数 10 的基础上进行计算，不断地重复 731 731……加权，具体方法如下：

步骤 1 从左到右，用相应顺序位置上的加权数乘相关数字数据元素的每一位数。

步骤 2 将每次乘法运算的乘积相加。

步骤 3 将得出的和除以 10（模数）。

步骤 4 余数即为校验数位。

对于数字没有占满所有可用字符位置的数据元素，应使用符号<填充空白位置，并且应确定它为零值，以便计算校验数位。

在对含有字母字符的数据元素进行校验数位计算时，字符 A 至 Z 的赋值应依次为 10 到 35，具体如下：

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35

校验数位计算中所包含的数据元素以及每类证件中的检验数位位置载于 Doc 9303 号文件针对特定尺寸旅行证件的第 4 部分至第 7 部分。校验数位计算的示例见提供了详实信息的本部分附录 A。

4.10 机读区的特征

除在此另有说明外，机读旅行证件应在以下事项方面符合[ISO 1831]：

- 所采用的基材的光学特性；
- 构成 OCR 字符的图像图案的光学和尺寸特性；和
- 与 OCR 字符在基材上的位置相关的基本要求。

机读数据应在固定长度域中从左到右分两行排列（上和下），但 TD1 型旅行证件除外，该类证件分三行（上、中、下）排列。应按照 Doc 9303 号文件针对特定尺寸旅行证件的第 4 部分至第 7 部分中的数据结构表所规定的顺序排列数据，并且数据在证件上的位置应遵循这些部分所规定的要求。在每个域中登记数据时应从左边字符位置开始。

如果登记的数据没有占满相关域的所有字符位置，应使用符号<填充未占位置。

4.11 机读区的质量规范

除在此另有规定外，打印质量一般应符合[ISO 1831] Range X。除另有说明外，下文规定的全部质量规范都应符合本部分第 2 节的要求，并且应适用于后期加工后的机读旅行证件以及已经放进护照或其他旅行证件的签证。

基材质量 [ISO 1831]，4.3 一直到 4.3.2，应只用作参考。

基材的不透明性在后期加工前后使用和测定的基材（对于签证，应为在将其放置进护照或其他旅行证件之前）应至少在[ISO 1831]的 4.4.1 和 4.4.3 所规定的中等不透明性的定义范围内。

基材光泽 光泽度没有规定。

荧光性 在紫外线光的照射下，基材在可见光谱中的反射系数不得显示出眼睛看得到的荧光，但那种出于安全原因而设定的可预测的荧光除外。

替代基材 不管是哪种基材，都应遵循前文所述的质量规范。

光谱带 OCR 打印应清晰可见，而且应使用黑色（[ISO 1831]中规定的 B425 到 B680）。OCR 打印还应在

[ISO 1831]规定的 B900 谱带下吸光（即近红外）。如果需透过可能加在证件表面的保护性材料对字符进行机读，这一特征必须顺利通过测试。

打印对比度（PCS） 后期加工后，根据[ISO 1831]规定测定的最小打印对比度（PCS/min）在 B900 光谱带下应为 $PCS/min \geq 0.6$ 。

字符笔画宽度 后期加工后的笔画宽度应符合[ISO 1831]的 5.3.1 中关于 Range X 的规定。

对比度变化率（CVR） 后期加工后，对比度变化率应如[ISO 1831]中所示的 Range X 数值，即 $CVR < 1.50$ 。

污点和外部标记 [ISO 1831]的 5.4.4.6 和 5.4.5.12 应适用于阅读表面（另见[ISO 1831]附录 B 的 B.6 和附录 C 的 C.5.10）。

失效 在[ISO 1831]的 5.4.5.9 中定义的“d”值在阅读表面上应等于 0.4。

行分离 见 Doc 9303 号文件针对特定尺寸旅行证件的第 4 部分至第 7 部分。

行间隔 见 Doc 9303 号文件针对特定尺寸旅行证件的第 4 部分至第 7 部分。

机读区行的歪斜 机读区行的实际歪斜率和机读区字符的歪斜率从参考边缘测量不得超过 3 度，同时机读区的歪斜或字符偏差不得导致机读区的行或其中任何部分超出 Doc 9303 号文件针对特定尺寸旅行证件的第 4 部分至第 7 部分所规定的打印区。

5. 代表国籍、出生地、签发国/机构所在地和其他情况的代码

A 部分 — 字母代码

以下为[ISO 3166-1]中规定的并定期更新的实体的二字和三字代码，还包括用星号标出的一些特定国家和组织。代码的当前版本可从[ISO 3166-1]维护机构 — [ISO 3166/MA]获得，该机构是国际标准化组织在国家代码方面的联系单位。

实体（简称）	二字代码	三字代码	实体（简称）	二字代码	三字代码
阿富汗	AF	AFG	安道尔	AD	AND
奥兰群岛	AX	ALA	安哥拉	AO	AGO
阿尔巴尼亚	AL	ALB	安圭拉	AI	AIA
阿尔及利亚	DZ	DZA	南极洲	AQ	ATA
美属萨摩亚	AS	ASM	安提瓜和巴布达	AG	ATG

阿根廷	AR	ARG
亚美尼亚	AM	ARM
阿鲁巴	AW	ABW
澳大利亚	AU	AUS
奥地利	AT	AUT
阿塞拜疆	AZ	AZE
巴哈马	BS	BHS
巴林	BH	BHR
孟加拉国	BD	BGD
巴巴多斯	BB	BRB
白俄罗斯	BY	BLR
比利时	BE	BEL
伯利兹	BZ	BLZ
贝宁	BJ	BEN
百慕大	BM	BMU
不丹	BT	BTN
多民族玻利维亚国	BO	BOL
博内尔岛、圣尤斯特歇斯岛 和萨巴岛	BQ	BES
波斯尼亚和黑塞哥维那	BA	BIH
博茨瓦纳	BW	BWA
布韦岛	BV	BVT
巴西	BR	BRA

英属印度洋领地	IO	IOT
文莱达鲁萨兰国	BN	BRN
保加利亚	BG	BGR
布基纳法索	BF	BFA
布隆迪	BI	BDI
佛得角	CV	CPV
柬埔寨	KH	KHM
喀麦隆	CM	CMR
加拿大	CA	CAN
开曼群岛	KY	CYM
中非共和国	CF	CAF
乍得	TD	TCD
智利	CL	CHL
中国	CN	CHN
圣诞岛	CX	CXR
科科斯（基灵）群岛	CC	CCK
哥伦比亚	CO	COL
科摩罗	KM	COM
刚果	CG	COG
刚果民主共和国	CD	COD
库克群岛	CK	COK
哥斯达黎加	CR	CRI

科特迪瓦	CI	CIV
克罗地亚	HR	HRV
古巴	CU	CUB
库拉索	CW	CUW
塞浦路斯	CY	CYP
捷克共和国	CZ	CZE
丹麦	DK	DNK
吉布提	DJ	DJI
多米尼加	DM	DMA
多米尼加共和国	DO	DOM
厄瓜多尔	EC	ECU
埃及	EG	EGY
萨尔瓦多	SV	SLV
赤道几内亚	GQ	GNQ
厄立特里亚	ER	ERI
爱沙尼亚	EE	EST
埃塞俄比亚	ET	ETH
福克兰群岛（马尔维纳斯）	FK	FLK ¹
法罗群岛	FO	FRO
斐济	FJ	FJI

芬兰	FI	FIN
法国	FR	FRA
法属圭亚那	GF	GUF
法属玻利尼西亚	PF	PYF
法属南领地	TF	ATF
加蓬	GA	GAB
冈比亚	GM	GMB
格鲁吉亚	GE	GEO
德国	DE	D
加纳	GH	GHA
直布罗陀	GI	GIB
希腊	GR	GRC
格陵兰	GL	GRL
格林纳达	GD	GRD
瓜德罗普	GP	GLP
关岛	GU	GUM
危地马拉	GT	GTM
根西岛	GG	GGY
几内亚	GN	GIN
几内亚—比绍	GW	GNB
圭亚那	GY	GUY
海地	HT	HTI

1. 就福克兰群岛（马尔维纳斯）主权问题，阿根廷和大不列颠及北爱尔兰联合王国政府之间存有争议。

赫德岛和麦克唐纳群岛	HM	HMD
教廷（梵蒂冈城国）	VA	VAT
洪都拉斯	HN	HND
中国香港特别行政区	HK	HKG
匈牙利	HU	HUN
冰岛	IS	ISL
印度	IN	IND
印度尼西亚	ID	IDN
伊朗（伊斯兰共和国）	IR	IRN
伊拉克	IQ	IRQ
爱尔兰	IE	IRL
马恩岛	IM	IMN
以色列	IL	ISR
意大利	IT	ITA
牙买加	JM	JAM
日本	JP	JPN
泽西	JE	JEY
约旦	JO	JOR
哈萨克斯坦	KZ	KAZ
肯尼亚	KE	KEN
基里巴斯	KI	KIR

朝鲜民主主义人民共和国	KP	PRK
大韩民国	KR	KOR
科威特	KW	KWT
吉尔吉斯斯坦	KG	KGZ
老挝人民民主共和国	LA	LAO
拉脱维亚	LV	LVA
黎巴嫩	LB	LBN
莱索托	LS	LSO
利比里亚	LR	LBR
利比亚	LY	LBY
列支敦士登	LI	LIE
立陶宛	LT	LTU
卢森堡	LU	LUX
中国澳门特别行政区	MO	MAC
前南斯拉夫马其顿共和国	MK	MKD
马达加斯加	MG	MDG
马拉维	MW	MWI
马来西亚	MY	MYS
马尔代夫	MV	MDV
马里	ML	MLI
马耳他	MT	MLT

马绍尔群岛	MH	MHL
马提尼克	MQ	MTQ
毛里塔尼亚	MR	MRT
毛里求斯	MU	MUS
马约特	YT	MYT
墨西哥	MX	MEX
密克罗尼西亚（联邦）	FM	FSM
摩尔多瓦共和国	MD	MDA
摩纳哥	MC	MCO
蒙古	MN	MNG
黑山	ME	MNE
蒙特塞拉特	MS	MSR
摩洛哥	MA	MAR
莫桑比克	MZ	MOZ
缅甸	MM	MMR
纳米比亚	NA	NAM
瑙鲁	NR	NRU
尼泊尔	NP	NPL
荷兰	NL	NLD
荷属安的列斯	AN	ANT
中立区	NT	NTZ
新喀里多尼亚	NC	NCL

新西兰	NZ	NZL
尼加拉瓜	NI	NIC
尼日尔	NE	NER
尼日利亚	NG	NGA
纽埃	NU	NIU
诺福克岛	NF	NFK
北马里亚纳群岛	MP	MNP
挪威	NO	NOR
阿曼	OM	OMN
巴基斯坦	PK	PAK
帕劳	PW	PLW
巴勒斯坦国	PS	PSE
巴拿马	PA	PAN
巴布亚新几内亚	PG	PNG
巴拉圭	PY	PRY
秘鲁	PE	PER
菲律宾	PH	PHL
皮特凯恩	PN	PCN
波兰	PL	POL
葡萄牙	PT	PRT
波多黎各	PR	PRI
卡塔尔	QA	QAT

留尼汪岛	RE	REU
罗马尼亚	RO	ROU
俄罗斯联邦	RU	RUS
卢旺达	RW	RWA
圣巴泰勒米岛	BL	BLM
圣赫勒拿岛、阿森松岛和特 里斯坦达库尼亚群岛	SH	SHN
圣基茨和尼维斯	KN	KNA
圣卢西亚	LC	LCA
圣马丁岛（法属）	MF	MAF
圣皮埃尔和密克隆	PM	SPM
圣文森特和格林纳丁斯	VC	VCT
萨摩亚	WS	WSM
圣马力诺	SM	SMR
圣多美和普林西比	ST	STP
沙特阿拉伯	SA	SAU
塞内加尔	SN	SEN
塞尔维亚	RS	SRB
塞舌尔	SC	SYC
塞拉利昂	SL	SLE
新加坡	SG	SGP
圣马丁岛（荷属）	SX	SXM
斯洛伐克	SK	SVK

斯洛文尼亚	SI	SVN
所罗门群岛	SB	SLB
索马里	SO	SOM
南非	ZA	ZAF
南乔治亚和南桑威奇群岛	GS	SGS
南苏丹	SS	SSD
西班牙	ES	ESP
斯里兰卡	LK	LKA
苏丹	SD	SDN
苏里南	SR	SUR
斯瓦尔巴和扬马延群岛	SJ	SJM
斯威士兰	SZ	SWZ
瑞典	SE	SWE
瑞士	CH	CHE
阿拉伯叙利亚共和国	SY	SYR
中国台湾省	TW	TWN
塔吉克斯坦	TJ	TJK
坦桑尼亚联合共和国	TZ	TZA
泰国	TH	THA
东蒂汶	TL	TLS
多哥	TG	TGO
托克劳群岛	TK	TKL

汤加	TO	TON
特立尼达和多巴哥	TT	TTO
突尼斯	TN	TUN
土耳其	TR	TUR
土库曼斯坦	TM	TKM
特克斯和凯科斯群岛	TC	TCA
图瓦卢	TV	TUV
乌干达	UG	UGA
乌克兰	UA	UKR
阿拉伯联合酋长国	AE	ARE
联合王国	GB	
— 英国公民		GBR
— 英国海外领土公民		GBD*
— 英国国民（海外）		GBN*
— 英国海外公民		GBO*
— 英国被保护人		GBP*

— 英国臣民		GBS*
美国	US	USA
美国本土外小岛屿	UM	UMI
乌拉圭	UY	URY
乌兹别克斯坦	UZ	UZB
瓦努阿图	VU	VUT
委内瑞拉玻利瓦尔共和国	VE	VEN
越南	VN	VNM
维尔京群岛（英属）	VG	VGB
维尔京群岛（美属）	VI	VIR
瓦利斯和富图纳群岛	WF	WLF
西撒哈拉	EH	ESH
也门	YE	YEM
赞比亚	ZM	ZMB
津巴布韦	ZW	ZWE

B 部分 — ISO 3166/MA 预留的其他代码

欧盟（EU）		EUE
--------	--	-----

C 部分 — 联合国旅行证件上使用的代码

联合国组织或该组织的一名官员	ZZ	UNO
联合国专门机构或该机构的一名官员	ZZ	UNA
持有联合国驻科索沃临时管理特派团（UNMIK）签发的旅行证件的科索沃居民		UNK

D 部分 — 其他签发机关的代码

非洲开发银行（ADB）		XBA
非洲进出口银行（AFREXIM）		XIM
加勒比共同体或其一名使者（CARICOM）		XCC
东部和南部非洲共同市场（COMESA）		XCO
欧洲理事会		XCE
西非国家经济共同体（ECOWAS）		XEC
国际刑事警察组织（INTERPOL）		XPO
马耳他最高军教团或其一名使者		XOM
南部非洲发展共同体		XDC

E 部分 — 没有确定国籍的人的代码

1954 年《无国籍人地位公约》第 1 条定义的无国籍人		XXA
经 1967 年议定书修订的 1951 年《难民地位公约》第 1 条定义的难民		XXB
不在上述 XXB 代码定义之内的难民		XXC
未确定国籍的人，签发国认为不管该人的地位是什么，没有必要为其确定上述任何 XXA、XXB 或 XXC 代码。该类可以包括既不是无国籍人，也不是难民，而是在签发国合法居住的不明国籍的人。		XXX

6. 建议各国使用的转写方法

以下表格载有拉丁语、西里尔语和阿拉伯语系最常用的国家字符。

A. 基于拉丁语的多国字符的转写

统一码	国家字符	说明	建议的转写
00C0	À	A 加沉音符号	A
00C1	Á	A 加尖重音符号	A
00C2	Â	A 加抑扬音调符号	A
00C3	Ã	A 加鼻音化符号	A
00C4	Ä	A 加分音符号	AE 或 A
00C5	Å	A 上加小圆圈	AA 或 A
00C6	Æ	AE 连字	AE
00C7	Ç	C 加变音符号	C
00C8	È	E 加沉音符号	E
00C9	É	E 加尖重音符号	E
00CA	Ê	E 加抑扬音调符号	E
00CB	Ë	E 加分音符号	E
00CC	Ì	I 加沉音符号	I
00CD	Í	I 加尖重音符号	I
00CE	Î	I 加抑扬音调符号	I
00CF	Ï	I 加分音符号	I
00D0	Ð	Eth 齿间擦音	D
00D1	Ñ	N 加鼻音化符号	N 或 NXX
00D2	Ò	O 加顿音符号	O
00D3	Ó	O 加尖重音符号	O
00D4	Ô	O 加抑扬音调符号	O
00D5	Õ	O 加鼻音化符号	O
00D6	Ö	O 加分音符号	OE 或 O
00D8	Ø	O 加斜线符号	OE
00D9	Ù	U 加沉音符号	U
00DA	Ú	U 加尖重音符号	U
00DB	Û	U 加抑扬音调符号	U
00DC	Ü	U 加分音符号	UE 或 UXX 或 U
00DD	Ý	Y 加尖重音符号	Y
00DE	Þ	北欧字母(冰岛)	TH

统一码	国家字符	说明	建议的转写
00DF	B	双 s (德国)	SS
0100	Ā	A 加长音符号	A
0102	Ǻ	A 加短音符号	A
0104	Ą	A 加元音鼻音化符号	A
0106	Ć	C 加尖重音符号	C
0108	Ĉ	C 加抑扬音调符号	C
010A	Č	C 加顿音符号	C
010C	Ď	C 加倒折符号	C
010E	Đ	D 加倒折符号	D
0110	Ð	D 加横线符号	D
0112	Ē	E 加长音符号	E
0114	Ĕ	E 加短音符号	E
0116	Ė	E 加顿音符号	E
0118	Ę	E 加元音鼻音化符号	E
011A	Ě	E 加倒折符号	E
011C	Ĝ	G 加抑扬音调符号	G
011E	Ğ	G 加短音符号	G
0120	Ġ	G 加顿音符号	G
0122	Ģ	G 加变音符号	G
0124	Ĥ	H 加抑扬音调符号	H
0126	H	H 加横线符号	H
0128	Ī	I 加鼻音化符号	I
012A	Ĭ	I 加长音符号	I
012C	Ĳ	I 加短音符号	I
012E	Į	I 加元音鼻音化符号	I
0130	İ	I 加顿音符号	I
0131	I	I 无点 (土耳其)	I
0132	IJ	IJ 连字	IJ
0134	Ĵ	J 加抑扬音调符号	J
0136	Ķ	K 加变音符号	K
0139	Ĺ	L 加尖重音符号	L
013B	Ļ	L 加变音符号	L
013D	Ł	L 加倒折符号	L
013F	Ł	L 加间隔符号	L
0141	Ł	L 加斜线符号	L
0143	Ń	N 加尖重音符号	N
0145	Ņ	N 加变音符号	N

统一码	国家字符	说明	建议的转写
0147	Ñ	N 加倒折符号	N
014A	Đ	Eng	N
014C	Ō	O 加长音符号	O
014E	Ö	O 加短音符号	O
0150	Ő	O 加双尖重音符号	O
0152	Œ	OE 连字	OE
0154	Ŕ	R 加尖重音符号	R
0156	Ŗ	R 加变音符号	R
0158	Ř	R 加倒折符号	R
015A	Ś	S 加尖重音符号	S
015C	Ŝ	S 加抑扬音调符号	S
015E	Ș	S 加变音符号	S
0160	Š	S 加倒折符号	S
0162	Ţ	T 加变音符号	T
0164	Ť	T 加倒折符号	T
0166	Ƨ	T 加斜线符号	T
0168	Ũ	U 加鼻音化符号	U
016A	Ū	U 加长音符号	U
016C	Ů	U 加短音符号	U
016E	Ů̊	U 加小圆圈	U
0170	Ű	U 加双尖重音符号	U
0172	Ū̊	U 加元音鼻音化符号	U
0174	Ŵ	W 加抑扬音调符号	W
0176	Ŷ	Y 加抑扬音调符号	Y
0178	ÿ	Y 加分音符号	Y
0179	Ž	Z 加尖重音符号	Z
017B	Ẓ̌	Z 加顿音符号	Z
017D	Ž̌	Z 加倒折符号	Z

B. 西里尔字符的转写

统一码	国家字符	建议的转写
0401	Ё	E(例外: 白俄罗斯语= IO)
0402	Ђ	D
0404	Є	IE(例外: 乌克兰语首字, 则=YE)
0405	Ѕ	DZ
0406	І	I
0407	Ї	I(例外: 乌克兰语首字, 则=YI)
0408	Ј	J
0409	Љ	LJ
040A	Њ	NJ
040C	Ќ	K(例外: 在塞尔维亚语和前南斯拉夫马其顿共和国所讲语言= KJ)
040E	Ў	U
040F	Џ	DZ(例外: 在塞尔维亚语和前南斯拉夫马其顿共和国所讲语言= DJ)
0410	А	A
0411	Б	B
0412	В	V
0413	Г	G(例外: 白俄罗斯语、塞尔维亚语和乌克兰语=H)
0414	Д	D
0415	Е	E
0416	Ж	ZH(例外: 塞尔维亚语=Z)
0417	З	Z
0418	И	I(例外: 乌克兰语=Y)
0419	Ї	I(例外: 乌克兰语首字, 则=Y)
041A	К	K
041B	Л	L
041C	М	M
041D	Н	N
041E	О	O
041F	П	P
0420	Р	R
0421	С	S
0422	Т	T
0423	У	U
0424	Ф	F

统一码	国家字符	建议的转写
0425	X	KH(例外: 塞尔维亚语和前南斯拉夫马其顿共和国=H)
0426	Ц	TS(例外: 塞尔维亚语和前南斯拉夫马其顿共和国=C)
0427	Ч	CH(例外: 塞尔维亚语=C)
0428	Ш	SH(例外: 塞尔维亚语=S)
0429	Щ	SHCH(例外: 保加利亚语=SHT)
042A	Ъ	IE
042B	Ы	Y
042D	Э	E
042E	Ю	IU(例外: 乌克兰语首字, 则=YU)
042F	Я	IA(例外: 乌克兰语首字, 则=YA)
046A	Ѹ	U
0474	V	Y
0490	Г	G G(例外: 在前南斯拉夫马其顿共和国所讲语言=GJ)
0492	F	
04BA	h	C

C. 阿拉伯字母的转写

统一码	阿拉伯字母	名称	机读区
0621	ء	hamza	XE
0622	آ	alef 上加 madda	XAA
0623	أ	alef 上加 hamza	XAE
0624	ؤ	waw 上加 hamza	U
0625	إ	alef 下加 hamza	I
0626	ئ	yeh 上加 hamza	XI
0627	ا	alef	A
0628	ب	beh	B

统一码	阿拉伯字母	名称	机读区
0629	ة	tehmarbuta	XTA/XAH ¹
062A	ت	teh	T
062B	ث	theh	XTH
062C	ج	jeem	J
062D	ح	hah	XH
062E	خ	khah	XKH
062F	د	dal	D
0630	ذ	thal	XDH
0631	ر	reh	R
0632	ز	zain	Z
0633	س	seen	S
0634	ش	sheen	XSH
0635	ص	sad	XSS
0636	ض	dad	XDZ
0637	ط	tah	XTT
0638	ظ	zah	XZZ
0639	ع	ain	E
063A	غ	ghain	G
0640	-	tatwheel	(没有编码)
0641	ف	feh	F
0642	ق	qaf	Q
0643	ك	kaf	K
0644	ل	lam	L
0645	م	meem	M

1. 一般使用 XTA，当 teh marbuta 出现在名字成分的末尾时除外，这时应使用 XAH。

统一码	阿拉伯字母	名称	机读区
0646	ن	noon	N
0647	ه	heh	H
0648	و	waw	W
0649	ى	alefmaksura	XAY
064A	ي	yeh	Y
064B	◌َ	fathatan	(没有编码)
064C	◌ِ	dammatan	(没有编码)
064D	◌ِ	kasratan	(没有编码)
064E	◌َ	fatha	(没有编码)
064F	◌ُ	damma	(没有编码)
0650	◌ْ	kasra	(没有编码)
0651	◌◌◌	shadda	[双次] ²
0652	◌◌◌	sukun	(没有编码)
0670	◌◌◌	用作上标的 alef	(没有编码)
0671	أ	alefwasla	XXA
0679	ٹ	tteh	XXT
067C	ٲ	teh 加小圆圈	XRT
067E	پ	peh	P
0681	ح	hah 上加 hamza	XKE
0685	ح	hah 上加 3 个点	XXH
0686	چ	tcheh	XC
0688	ڌ	ddal	XXD
0689	د	dal 加小圆圈	XDR
0691	ڌ	rreh	XXR
0693	ر	reh 加小圆圈	XRR
0696	ر	reh 上下各加小圆圈	XRX

2. Shadda 表示双次的意义：重复一个拉丁字符或序列，如 عبااس 即为 EBBAS，فضضة 即为 FXDZXDZXA H。

统一码	阿拉伯字母	名称	机读区
0698	ژ	jeh	XJ
069A	ڙ	seen 上下各加小圆圈	XXS
069C	ڙڙ	seen 上下各加 3 个点	(没有编码)
06A2	ڙ	Feh 将点移至下方	(没有编码)
06A7	ڙ	qaf 上加 1 个点	(没有编码)
06A8	ڙ	qaf 上加 3 个点	(没有编码)
06A9	ڪ	keheh	XKK
06AB	ڪ	kaf 加小圆圈	XXK
06AD	ڱ	ng	XNG
06AF	گ	gaf	XGG
06BA	ڻ	noonghunna	XNN
06BC	ڻ	noon 加小圆圈	XXN
06BE	ھ	hehdoachashmee	XDO
06C0	ه	heh 上加 yeh	XYH
06C1	ه	heh goal	XXG
06C2	ه	heh goal 上加 hamza	XGE
06C3	ه	tehmarbuta goal	XTG
06CC	ی	farsiyeh	XYA
06CD	ی	yeh 加小尾巴	XXY
06D0	ی	yeh	Y
06D2	ے	yeh barree	XYB
06D3	ے	yeh barree 上加 hamza	XBE

7. 偏差

随着世界各国不断采用机读旅行证件，越发复杂的情况和偏差的增多使得有必要通过一种标准化机制报告一国偏离标准或通常做法的情况。偏差指机读旅行证件中包含未完全符合国际民航组织规范以及 ISO 和 RFC 规范性标准的元素。偏差一般出现在国家签名认证机构（CSCA）或证件签名者证书（DSCs）中。但是，有些国家还指出在其机读旅行证件内也存在与逻辑数据结构和机读区域相关的问题。本节的目的是详述签发国可借以公布其所存在偏差的机制。

虽然旅行证件可能会包含一些偏差，但在边境管理系统中它们还是可用的。而对于有效的证件，可以使用数年。因此，信赖方应确定自己的处理公布偏差的程序。

7.1 运作经验

很长一段时间以来，管理偏差的唯一做法是签发国通过外交手段提供一般性的建议。本节讲述了涉及大量机读旅行证件的各种偏差，可通过报告这些偏差协助边境管理部门确定旅行证件是有效的、伪造的或是被替换的。运作中出现的一些错误包括机读区、逻辑数据结构和公钥基础设施方面的偏差。

尽管机读区已经使用了很多年，但仍存在一些众所周知的错误，近期的例子包括：

- 机读区的出生日期与视读区页面上的出生日期对不上。
- 机读区的国籍错误地报告成出生国，而非国籍。

在大多数情况下，机读区未符合规范的旅行证件将会被签发国召回。由于证件的签发和随后的补发之间存在一定的时间间隔，旅行者可能不得不使用有偏差的机读旅行证件。在这段时间里，公布偏差可缓解可能对旅行者造成的问题。³

对于逻辑数据结构和公钥基础设施方面的偏差，有些可以长时间不被发现，这是由于许多国家还未按照 Doc 9303 号的规定实行被动和主动认证。但是，大力提倡签发国公布偏差，以协助国际社会处理在采用机读旅行证件时遇到的技术问题。⁴

7.2 偏差列表做法

本节所述的做法旨在为签发国提供一种用于公布和分发旅行证件偏差列表的标准化方法。这一做法所依据的是在编制国家签名认证机构证书总表过程中所建立的原则（见 Doc 9303 号文件第 12 部分），根据这些原则，有关每个国家不合规情况的经签署的偏差列表将通过国际民航组织的公钥目录予以提供，或者由签发国通过网站或简易目录访问协议服务器予以提供。公钥目录旨在为分发与偏差管理相关的信息提供支持。

-
3. 本节不涉及仅影响单个证件或少量电子机读旅行证件的不合规的情况。签发国可自行决定是否要召回并补发单个证件。
 4. 如果存在与公钥基础设施证书相关的安全问题，正确的做法是如 Doc 9303 号文件第 12 部分所述的那样将其撤销。本节不宜提供更多的指导。

偏差分为四个具体方面：

- 密钥和证书；
- 逻辑数据结构（LDS）；
- 机读区（MRZ）；
- 芯片。

对于每一个类别，偏差将仅从一个层级上描述，如：

类别：	LDS
错误	DG2

更多的信息将通过各国提供的运行参数和/或在报告框架内设置的一个使通知国可添加任何所需说明性文字的自由文本域来提供。通知国可在自由的文本域内添加可获取更多信息的链接。对于证书方面的错误，签发机构可选择发放新的证书，但这并不是强制要求。

是否将不合规情况通知信赖方由签发国自己决定。在决定是否编制偏差列表时，各国应该考虑到，随着旅行者自助边检越来越普及，如果不能将与不合规的旅行证件相关的信息通知出去，可能会给旅行者造成延误和不便，而这将对签发国和整个边检过程产生不良影响。

偏差列表是报告涉及数以千计而非几个或几百个旅行证件的偏差的一种方式。对于少量的不合规的旅行证件，各国宜直接处理。

7.3 方法

7.3.1 偏差元素

机读旅行证件的组成部分包括纸片和无线射频辨识芯片等元素，每个元素都通过安全特征得到某种保护，这些安全特征可被界定，因而可在旅行证件的有效期内通过查验系统加以测试。在物理旅行证件上采用的安全特征既有显性的，也有隐性的。本节仅考虑机读区、逻辑数据结构和公钥基础设施内的偏差元素。

机读区是机读旅行证件数据页上的一个尺寸固定的区域，包含强制性和选择性数据，这些数据使用便于通过光学字符识别方法进行机读的格式加以编排。Doc 9303 号文件提供了与机读区相关的规范，包括：

- 目的；
- 限制；
- 转写；和
- 机读区行的数据结构。

机读区与规范的符合性定期通过查验系统进行测试，做法是将机读区的数据与对应的视读区页上的数据进行对比以及重新计算机读区的校验数位。

机读旅行证件无线射频辨识芯片上存储的数据的真实性和完整性由被动认证技术加以保护。这一安全机制以数字签名和公钥基础设施（PKI）为基础。

Doc 9303 号文件第 10 部分对机读旅行证件的逻辑数据结构进行了定义。虽然没有特定的测试可用于确定是否符合规范，但是存储在逻辑数据结构内的数据有一部分是机读旅行证件机读区或视读区页上数据的子集。因此，适用于机读区和视读区页的测试同样可用于测试机读区和视读区的数字化数据。逻辑数据结构的真实性可通过由查验系统准确实施被动认证程序来确定，而主动认证则由芯片进行。简要描述如下：

被动认证（PA）以数字签名为依据，包含以下公钥基础设施成分：

1. **国家签名认证机构（CSCA）**：每个国家设立一个国家签名认证机构作为其在电子机读旅行证件方面的国家级信任单位。国家签名认证机构为一个或多个（国家）证件签名者颁发公钥证书。此外，国家签名认证机构还发布包含所有被撤销证书的证书撤销列表（CRLs）。⁵
2. **证件签名者（DS）**：证件签名者对存储在机读旅行证件上的数据进行数字签名；该签名被存储在每份证件的证件安全对象上。

主动认证（AA）：实施主动认证时，每个芯片都包含有自己的主动认证密钥对。私钥被存储在芯片的安全存储器中，而公钥被存储在逻辑数据结构的第 15 数据组中。

7.3.2 发布偏差列表

偏差列表不得由国家签名认证机构直接发布，而是应由国家签名认证机构授权一个偏差列表签名者（Doc 9303 号文件第 12 部分）负责汇编、签署和发布偏差列表。

发布偏差列表所应采取的程序应反映在负责发布的国家签名认证机构所公布的认证政策中。

7.3.3 接收偏差列表

各接收国制定其关于接收偏差列表的政策，并确定在查验证件过程中处理偏差的方式。一般而言，这类政策属于保密信息。

是否允许存在偏差的机读旅行证件继续使用将由接收国自行决定。

5. 由于证书撤销列表是一种安全报告机制，且会不断重新发布，所以没必要报告其中的缺陷，本节也不宜涉及这些内容。

7.3.4 偏差的类别

7.3.4.1 密钥和证书

证书和密钥方面的偏差仅限于下列情况：

问题	说明
证书	描述至域或域的扩展
密钥	描述至域或域的扩展
主动认证	仅描述至错误/问题

注：如果报告国决定签发新的证书，则不得将证书纳入偏差列表中，但可以通过自由文本域对其加以说明。

7.3.4.2 逻辑数据结构（LDS）

逻辑数据结构方面的偏差仅限于下列情况：

问题	说明
数据组通用信息	描述至编码错误
数据组	描述至数据组
安全对象数据	描述至具体问题（如证件签名者证书）

7.3.4.3 机读区（MRZ）

机读区方面的偏差仅限于下列情况：

问题	说明
与视读区的匹配	描述至域
校验数位	描述至相关校验数位
信息编码错误	描述至机读区域

7.3.5 偏差类型的定义

偏差类别和相应参数可能会不断扩充，Doc 9303 号文件将使其保持在最新状态。

每一偏差由一个偏差描述元素加以描述。偏差用一个客体标识符即偏差类型来标示，并可通过参数对其进行更加详细的说明。域描述可包含更多的信息，如偏差的性质为何无法通过主要偏差类型加以充分描述。

偏差类型	参数	描述
证书/密钥偏差		
标识符—偏差—证书或密钥 (id-Deviation-CertOrKey)	无	未包括在下列更为详细的偏差范围之内的与证书或密钥相关的一般性偏差。
标识符—偏差—证书或密钥—证件签名者的签名 (id-Deviation-CertOrKey-DSSignature)	无	证件签名者证书上的签名有误。
标识符—偏差—证书或密钥—证件签名者编码 (id-Deviation-CertOrKey-DSEncoding)	证书域 (CertField)	证件签名者证书包含错误编码。
标识符—偏差—证书或密钥—国家签名认证机构编码 (id-Deviation-CertOrKey-CSCAEncoding)	证书域 (CertField)	国家签名认证机构证书包含编码错误。
标识符—偏差—证书或密钥—主动认证密钥泄密 (id-Deviation-CertOrKey-AAKeyCompromised)	无	主动认证密钥可能泄密，不应信赖。
逻辑数据结构偏差		
标识符—偏差—逻辑数据结构 (id-Deviation-LDS)	无	未包括在下列更为详细的偏差范围之内的与逻辑数据结构相关的一般性偏差。
标识符—偏差—逻辑数据结构—数据组格式错误 (id-Deviation-LDS-DGMalformed)	数据组 (Datagroup)	某一数据组的标志长度值编码被破坏。
标识符—偏差—逻辑数据结构—数据组散列错误 (id-Deviation-LDS-DGHashWrong)	数据组 (Datagroup)	安全对象数据中某一数据组的散列值有误。
标识符—偏差—逻辑数据结构—安全对象数字签名错误 (id-Deviation-LDS-SODSignatureWrong)	无	安全对象数据中所包含的签名有误。
标识符—偏差—逻辑数据结构—数据组通用信息不一致 (id-Deviation-LDS-COMinconsistent)	无	数据组通用信息和安全对象数据不一致。

偏差类型	参数	描述
机读区偏差		
标识符—偏差—机读区 (id-Deviation-MRZ)	无	未包括在下列更为详细的偏差范围之内的与机读区相关的一般性偏差。
标识符—偏差—机读区—错误数据 (id-Deviation-MRZ-WrongData)	机读区域 (MRZField)	机读区的某个域包含错误数据（如与视读区不一致），但可用导出的基本访问控制密钥打开芯片。 若基本访问控制密钥无法使用，则应该在偏差列表中另外加入标识符—偏差—芯片。
标识符—偏差—机读区—错误校验数位 (id-Deviation-MRZ-WrongCheckDigit)	机读区域 (MRZField)	机读区某一域的校验数位计算错误。
芯片偏差		
标识符—偏差—芯片 (id-Deviation-Chip)	无	芯片无法使用，如基本访问控制密钥错误、天线坏了或其他物理缺陷。

7.3.6 识别存在偏差的证件

可通过几种不同的方法识别受到偏差影响的证件：

- 通过用于为这些证件签名的证件签名者证书；证件签名者可通过以下方式加以识别：
 - 签发人的唯一甄别名以及证书的序列号（签发人和序列号（issuerAndSerialNumber）），
 - 用作识别证件签名者唯一标志的主体密钥标识符（subjectKeyIdentifier），或
 - 证件签名者证书散列（证书散列（certificateHash））；所使用的散列函数与偏差列表签名中所用的函数是一样的。
- 通过一系列签发日期（签发日期开始时间、签发日期结束时间（startIssuingDate, endIssuingDate））
- 通过证件号列表（证件号列表（listOfDocNumbers））。

对于偏差列表的发布者和接收者来说，每一种方法都有其优缺点。这些优缺点包括：

- 通过证件签名者加以识别的做法使得查验系统只有在实施被动认证后才能识别偏差。另外，通过证件签名者加以识别的做法可能太过粗略，无法准确地仅识别出有缺陷的证件，即偏差影响的可能仅是某一特定证件签名者所签名的一部分证件。

- 签发日期并非机读区的一部分，并且电子逻辑数据结构中一般不会包含签发日期。因此，这种方法不适用于自动处理。此外，在有些签发国，签发日期可能并不是对护照进行个人化处理的实际日期，而是申请日期，因此可能不够准确，无法仅识别出受到影响的证件。
- 如果证件号没有按顺序发放，则很难整理出证件号列表。如果许多证件都受到同一缺陷的影响，证件号列表会迅速变大到难以汇编的程度。

建议提供尽可能多的关于受影响证件的识别信息。如果给出了多种识别方法，则必须同时满足相关条件才能将证件确认为受到影响。信赖国应自行决定使用缺陷列表条目中给出的哪些识别方法来识别受影响证件。

7.4 公布

偏差列表可以通过国际民航组织的公钥目录和/或由签发机关通过网站或简易目录访问协议服务器进行公布。缺陷列表的一级分发点是公钥目录。

	偏差列表
一级分发点	公钥目录
二级分发点	网站/简易目录访问协议服务器

7.4.1 由签发国公布

偏差列表可通过签发机关的网站或简易目录访问协议服务器进行公布。

7.4.2 在公钥目录上公布

公钥目录可作为偏差列表的中央存储库。

公布偏差列表的程序如下所述：

1. 作为《公钥目录接口规范》和《公钥目录程序手册》中规定的证书上传常用过程的一部分，将偏差列表发送至可写公钥目录。
2. 国际民航组织公钥目录办公室按照《公钥目录程序手册》的规定验证所上传的偏差列表的签名。
3. 将有效的偏差列表移至只读公钥目录。
4. 分发国将决定是否将偏差列表公开发布，或者仅提供给公钥目录的成员国内。

7.4.3 信赖方

为核实偏差列表，信赖方需通过带外传输收到签发国相应的国家签名认证机构证书。信赖方可自行决定如何处理在签发国偏差列表中有相应记录的机读旅行证件。

7.5 技术规范

7.5.1 偏差列表规范

偏差列表应根据[RFC 3852]的规定，作为一种签名数据进行编制。所有偏差列表在编制时都必须采取唯一编码规则格式，以保持列表中签名的完整性。

偏差所涉范围将由以下各项来制约：

- 日期范围（包括签发日期和失效日期）；
- 颁发者名称和序列号；
- 证件签名者证书的主体密钥标识符；
- 电子机读旅行证件编号列表。

将利用这些值的适当组合来准确地制约受影响的机读旅行证件的范围。在组合这些值时，将用“和”进行连接。不得用“或”作为各项值的连接符。

7.5.1.1 签名数据类别

应适用[RFC 3852]中的处理规则。

M 强制性 — 必须有这个域。

R 建议的 — 应该有这个域。

X 不得使用 — 不得填写这个域。

O 选择性 — 可以有这个域。

值		说明
签名数据 (SignedData)		
版本 (version)	m	值 = v3
摘要算法 (digestAlgorithms)	m	
封装数据内容信息 (encapContentInfo)	m	
电子内容类型 (eContentType)	m	标识符 — 缺陷列表 (id-DefectList)
电子内容 (eContent)	m	编码的内容缺陷列表 (DefectList)

值		说明
证书 (certificates)	m	各国必须纳入缺陷列表签名者证书，并且应该纳入国家签名认证机构证书，这些证书可用于验证签名者信息 (signerInfos) 域中的签名。
证书撤销列表 (crls)	x	
签名者信息 (signerInfos)	m	建议各国仅在该域中提供一个签名者信息。
签名者信息 (SignerInfo)	m	
版本 (version)	m	该域的值由安全标识符域决定。有关该域的规则，见[RFC 3852]第 5.3 节。
安全标识符 (sid)	m	
主体密钥标识符 (subjectKeyIdentifier)	r	建议各国在签发者和序列号 (issuerandSerialNumber) 的基础上对该域进行支持。
摘要算法 (digestAlgorithm)	m	用于根据封装数据内容信息 (encapContentInfo) 和签名属性 (signedAttrs) 计算散列值的算法的标识符。
签名属性 (signedAttrs)	m	制作国似宜在签名中纳入更多的属性，但除用于验证签名值外，接收国没必要对这些属性进行处理。签名属性 (signedAttrs) 必须包含签名时间 (参见 PKCS#9)。
签名算法 (signatureAlgorithm)	m	用于计算签名值和任何相关参数的算法的标识符。
签名 (signature)	m	签名生成过程的结果。
非签名属性 (unsignedAttrs)	x	

7.5.1.2 ASN.1 规范

DeviationList

{ iso-itu-t(2) international-organization(23) icao(136) mrt(1) security(1) deviationlist(7)}

DEFINITIONS IMPLICIT TAGS ::=

BEGIN

IMPORTS

— Imports from RFC 3280 [PROFILE], Appendix A.1

AlgorithmIdentifier

FROM PKIX1Explicit88

{ iso(1) identified-organization(3) dod(6)
internet(1) security(5) mechanisms(5) pkix(7)
mod(0) pkix1-explicit(18) }

— Imports from RFC 3852

SubjectKeyIdentifier, Digest, IssuerAndSerialNumber

FROM CryptographicMessageSyntax2004

{ iso(1) member-body(2) us(840) rsdsi(113549)
pkcs(1) pkcs-9(9) smime(16) modules(0)
cms-2004(24) };

DeviationListVersion ::= INTEGER {v0(0)}

DeviationList ::= SEQUENCE {

version DeviationListVersion,
digestAlgorithm AlgorithmIdentifier OPTIONAL,
deviations SET OF Deviation
}

Deviation ::= SEQUENCE{

documents DeviationDocuments,
descriptions SET OF DeviationDescription
}

DeviationDescription ::= SEQUENCE{

description PrintableString OPTIONAL,
deviationType OBJECT IDENTIFIER,
parameters [0] ANY DEFINED BY deviationType OPTIONAL,
nationalUse [1] ANY OPTIONAL

— nationalUse 域在国家内部使用，

— 不受国际民航组织规范的约束。

}

```

DeviationDocuments ::= SEQUENCE {
    documentType    [0] PrintableString (SIZE(2)) OPTIONAL,
        — per MRZ, e.g. ‘P’
    dscIdentifier   DocumentSignerIdentifier OPTIONAL,
    issuingDate    [4] IssuancePeriod OPTIONAL,
    documentNumbers [5] SET OF PrintableString OPTIONAL
}

```

```

DocumentSignerIdentifier ::= CHOICE{
    issuerAndSerialNumber [1] IssuerAndSerialNumber,
    subjectKeyIdentifier [2] SubjectKeyIdentifier,
    certificateDigest [3] Digest -- if used, digestAlgorithm must be present in DeviationList
}

```

```

IssuancePeriod ::= SEQUENCE {
    firstIssued   GeneralizedTime,
    lastIssued    GeneralizedTime
}

```

— CertField 用于说明证书的哪一部分受到编码错误的影响。主体部分通过对应的 CertificateBodyField 值加以确认，扩展通过用于识别拓展的对应的客体标识符加以确认。

```

CertField ::= CHOICE {
    body CertificateBodyField,
    extension OBJECT IDENTIFIER
}

```

```

CertificateBodyField ::= INTEGER {
    generic(0), version(1), serialNumber(2), signature(3), issuer(4),
    validity(5), subject(6), subjectPublicKeyInfo(7),
    issuerUniqueID(8), subjectUniqueID(9)
}

```

```

Datagroup ::= INTEGER
    { dg1(1), dg2(2), dg3(3), dg4(4), dg5(5), dg6(6),
      dg7(7), dg8(8), dg9(9), dg10(10), dg11(11),
      dg12(12), dg13(13), dg14(14), dg15(15), dg16(16),
      sod(20), com(21) }

```

```

MRZField ::= INTEGER
    { generic(0), documentCode(1), issuingState(2), personName(3),
      documentNumber(4), nationality(5), dateOfBirth(6),
      sex(7), dateOfExpiry(8), optionalData(9) }

```


— Base Object Identifiers

id-icao OBJECT IDENTIFIER ::= { 2 23 136 }

id-icao-mrtd OBJECT IDENTIFIER ::= { id-icao 1 }

id-icao-mrtd-security OBJECT IDENTIFIER ::= { id-icao-mrtd 1 }

id-icao-DeviationList OBJECT IDENTIFIER ::= { id-icao-mrtd-security 7 }

id-icao-DeviationListSigningKey OBJECT IDENTIFIER ::= { id-icao-mrtd-security 8 }

— Deviation Object Identifiers and Parameter Definitions

id-Deviation-CertOrKey OBJECT IDENTIFIER ::= { id-icao-DeviationList 1 }

id-Deviation-CertOrKey-DSSignature OBJECT IDENTIFIER ::= { id-Deviation-CertOrKey 1 }

id-Deviation-CertOrKey-DSEncoding OBJECT IDENTIFIER ::= { id-Deviation-CertOrKey 2 }

id-Deviation-CertOrKey-CSCAEncoding OBJECT IDENTIFIER ::= { id-Deviation-CertOrKey 3 }

id-Deviation-CertOrKey-AAKeyCompromised OBJECT IDENTIFIER ::= { id-Deviation-CertOrKey 4 }

id-Deviation-LDS OBJECT IDENTIFIER ::= { id-icao-DeviationList 2 }

id-Deviation-LDS-DGMalformed OBJECT IDENTIFIER ::= { id-Deviation-LDS 1 }

id-Deviation-LDS-DGHashWrong OBJECT IDENTIFIER ::= { id-Deviation-LDS 2 }

id-Deviation-LDS-SODSignatureWrong OBJECT IDENTIFIER ::= { id-Deviation-LDS 3 }

id-Deviation-LDS-COMInconsistent OBJECT IDENTIFIER ::= { id-Deviation-LDS 4 }

id-Deviation-MRZ OBJECT IDENTIFIER ::= { id-icao-DeviationList 3 }

id-Deviation-MRZ-WrongData OBJECT IDENTIFIER ::= { id-Deviation-MRZ 1 }

id-Deviation-MRZ-WrongCheckDigit OBJECT IDENTIFIER ::= { id-Deviation-MRZ 2 }

id-Deviation-Chip OBJECT IDENTIFIER ::= { id-icao-DeviationList 4 }

END

8. 参考文献(规范性)

- [ISO 1073-2] ISO 1073-2:1976, 光学识别字母数字字符集第2部分: 字符集 OCR-B — 印刷图像的形状和尺寸
- [ISO 1831] ISO 1831:1980, 光学字符识别印刷规范
- [ISO 3166-1] ISO 3166-1:2013, 国家及其地区的名称代码 — 第1部分: 国家代码
- [ISO 3166/MA] ISO 3166 维护机构 https://www.iso.org/iso/home/standards/country_codes.htm
- [ISO/IEC 7810] ISO/IEC 7810:2003, 识别卡 — 物理特性
- [ISO/IEC 19794-5] ISO/IEC 19794-5:2005, 信息技术 — 生物特征数据交换格式 — 第5部分: 人脸图像数据
- [RFC 3852] 加密报文语法 — 2004年7月
- [RFC 5280] D. Cooper, S. Santesson, S. Farrell, S. Boeyen, R. Housley, W. Polk, “互联网 X.509 公钥基础设施证书和证书撤销列表 (CRL) 简介”, 2008年5月。

第 3 部分附录 A 校验数位计算方法示例(资料性)

示例 1 校验数位应用在日期域上

以 1952 年 7 月 27 日为例，日期使用数字格式，其计算方法为：

	日期:	5	2	0	7	2	7		
	加权:	7	3	1	7	3	1		
步骤 1 (乘法运算)	乘积:	35	6	0	49	6	7		
步骤 2 (乘积之和)		$35 + 6 + 0 + 49 + 6 + 7 = 103$							
步骤 3 (除以模数)		$\frac{103}{10} = 10$, 余数为 3							

步骤 4 校验数位为余数 3。因此，该日期和其校验数位应写作 **5207273**。

示例 2 校验数位应用在证件号码域上

以号码 **AB2134** 为例，在一个 9—字符的固定长度域中（例如护照号）对其编码，其计算方法为：

样本数据元素:	A	B	2	1	3	4	<	<	<
赋予的数字值:	10	11	2	1	3	4	0	0	0
加权:	7	3	1	7	3	1	7	3	1
步骤 1 (乘法运算) 乘积:	70	33	2	7	9	4	0	0	0
步骤 2 (乘积之和)	$70 + 33 + 2 + 7 + 9 + 4 + 0 + 0 + 0 = 125$								
步骤 3 (除以模数)	$\frac{125}{10} = 12$, 余数为 5								

步骤 4 校验数位是余数 5。因此，该号码和其校验数位应写作 **AB2134<<<5**。

复合校验数位计算示例

所有机读旅行证件的复合校验数位计算方法都是一样的。但在进行计算时，不同类型证件的数字位置和数量则有所不同。为完整起见，在此各举一例。

样本数据元素:	7	1	2	2	<	<	<	<	<	<
赋予的数字值	7	1	2	2	0	0	0	0	0	0
加权:	1	7	3	1	7	3	1	7	3	1
步骤 1 (乘法运算) 乘积:	7	7	6	2	0					

样本数据元素:	<	<	<	<	<
赋予的数字值	0	0	0	0	0
加权:	7	3	1	7	3
步骤 1 (乘法运算) 乘积	0	0	0	0	0

步骤 2 (乘积之和)	91 + 6 + 3 + 7 + 12 + 5 + 56 + 27 + 0 + 49 +
步骤 2 (乘积之和)	0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 +
步骤 2 (乘积之和)	0 + 0 + 0 + 0 + 0 + 9 + 4 + 0 + 21 + 1 +
步骤 2 (乘积之和)	14 + 21 + 9 + 35 + 0 + 7 + 7 + 6 + 2 + 0 +
步骤 2 (乘积之和)	0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 +
步骤 2 (乘积之和)	=392
步骤 3 (除以模数)	$\frac{392}{10} = 39$, 余数为 2

步骤 4 校验数位是余数 2。因此，机读区中行数据和其复合校验数位可以写成如下形式：
3407127M9507122YT0<<<<<<<<<<<<<2。

示例 5 TD2 机读旅行证件复合校验数位的计算方法

以下列机读区的下行数据为例，对复合校验数位进行编码，其计算方法为：

下机读行（字符位置 1-35）：

HA672242<6YT05802254M9601086<<<<<<<<

样本数据元素:	H	A	6	7	2	2	4	2	<	6
赋予的数字值:	17	10	6	7	2	2	4	2	0	6
加权:	7	3	1	7	3	1	7	3	1	7
步骤 1 (乘法运算) 乘积:	119	30	6	49	6	2	28	6	0	42

样本数据元素:	5	8	0	2	2	5	4	9	6	0
赋予的数字值:	5	8	0	2	2	5	4	9	6	0
加权:	3	1	7	3	1	7	3	1	7	3
步骤 1 (乘法运算) 乘积:	15	8	0	6	2	35	12	9	42	0

样本数据元素:	1	0	8	6	<	<	<	<	<	<
赋予的数字值:	1	0	8	6	0	0	0	0	0	0
加权:	1	7	3	1	7	3	1	7	3	1
步骤 1 (乘法运算) 乘积:	1	0	24	6	0	0	0	0	0	0

样本数据元素:	<
赋予的数字值:	0
加权:	7
步骤 1 (乘法运算) 乘积:	0

步骤 2 (乘积之和) 119 + 30 + 6 + 49 + 6 + 2 + 28 + 6 + 0 + 42 +

步骤 2 (乘积之和) 15 + 8 + 0 + 6 + 2 + 35 + 12 + 9 + 42 + 0 +

步骤 2 (乘积之和) 1 + 0 + 24 + 6 + 0 + 0 + 0 + 0 + 0 + 0 +

步骤 2 (乘积之和) 0

步骤 2 (乘积之和) =448

步骤 3 (除以模数) $\frac{448}{10} = 44$, 余数为 8

步骤 4 校验数位为余数 8。因此，机读区下行数据和其复合校验数位可以写成如下形式：

HA672242<6YT05802254M9601086<<<<<<<8

第 3 部分附录 B 机读旅行证件中阿拉伯文字的转写(资料性)

B.1 阿拉伯文字

阿拉伯文字是阿拉伯文所用的文字，阿拉伯文是摩洛哥、阿曼等 24 个国家的官方语文。还有其他语文也使用阿拉伯文字，特别是伊朗的波斯文、阿富汗的普什图文和达里文、巴基斯坦的乌尔都文，以及库尔德文、亚述文、豪萨文和维吾尔文等许多其他语文。过去，它曾用于中亚的一些语文，如塔吉克文和乌兹别克文。

阿拉伯文字是草写体，一个字母的形状是多变的，这取决于它是独立的一个词（孤立的）、位于词首（开头）、词中（中间）还是词尾（末尾）。例如，ب (beh) 这一字母在位于 بكر (Bakr) 的词首时形状变为 ب — 注意阿拉伯文的阅读顺序为从右往左，因此，第一个字母在右侧。这里，我们不关注字母的不同形状（字形），只关注由独立的字母形状表示的基本字母代码。

阿拉伯文和使用阿拉伯文字的其他语文在书写时通常仅使用辅音。因此，محمد (默罕默德) 这个名字的书写形式仅包含四个辅音，大概等于拉丁语的“Mhmd”。翻译人员可自己决定是否加上元音，以实现语音上的对应。阿拉伯文还可“附加母音符”，如果需要通过添加母音符号 (“harakat”) 改变发音的话。但是，harakat 通常省略。

标准的阿拉伯文字包括 32 个辅音、18 个元音和双元音以及 3 个其他符号。此外，在用于阿拉伯文以外的语文时，阿拉伯文字中还有 100 多个本国特殊字符，但有些已经过时，不再使用。

B.2 机读旅行证件中的阿拉伯文字

B.2.1 视读区

视读区有一个针对姓名的强制性域（参见 Doc 9303 号文件第 4 部分至第 7 部分中与各种尺寸的机读旅行证件相关的规范）。Doc 9303 号文件第 3 部分的第 3.1 段写到：

“当强制性数据元素用不使用拉丁字母的本国语文表示时，还应提供转写。”

因此，当姓名用阿拉伯文字进行书写时，应增加一个拉丁字符表示法。Doc 9303 号文件将这种表示法称为“转写”，但这通常是一种对等译音，更确切地说应该将其称为“音译”。

例如：

用阿拉伯文字表示的姓名¹： **ابو بكر محمد بن زكريا الرازي**

音译为拉丁字符是：**Abū Bakr Mohammed ibn Zakarīa al-Rāzi**

首先应注意的是，Doc 9303 号文件第 3 部分的第 3.2 段允许签发国决定是否在视读区中使用变音符号（如 **al-Rāzi** 中的 **ā**）。

其次，应注意的是这种特定的拉丁字符音译方式只是许多可能的音译方式中的一种。“阿拉伯姓名变体数据库”网站²就**محمد**提供了以下 16 种不同表示方式：

- | | | |
|---------------|--------------|---------------|
| 1. Muhammad | 2. Moohammad | 3. Moohamad |
| 4. Mohammad | 5. Mohamad | 6. Muhamad |
| 7. Muhamad | 8. Mohamed | 9. Mohammed |
| 10. Mohemmed | 11. Mohemmed | 12. Muhemmed |
| 13. Muhamed | 14. Muhammed | 15. Moohammed |
| 16. Mouhammed | | |

在有些国家，一般会将末尾的“d”替换为“t”，这样**محمد**就有 32 种不同的表示方式。

所用的音译方案取决于使用阿拉伯文字的地方的语文和地方口音（波斯文、普什图文和乌尔都文等非阿拉伯文语文也使用阿拉伯文字）、说拉丁文的人所使用的语文，以及所用的音译方案。

B.2.2 机读区

Doc 9303 号文件本部分第 4 节对机读区进行了描述。

机读区以适用于各类机读旅行证件的标准格式提供了一套基本的数据元素，所有的接收国，不管其国家的文字或习惯如何，都可以使用这一格式。机读区的数据经格式化后，全世界具有标准能力的机器都可以阅读，因此，机读区的数据表示法不同于视读区的数据表示法。国家特有字符一般只出现在使用这些字符的国家的计算机处理系统中，不能全球通用。因此，这种字符不得出现在机读区。

在机读护照中，机读区的姓名域包含 39 个字符位置，且仅可使用 OCR-B 的 A-Z 子集和填充符。因此，机读区中不得使用阿拉伯字符，必须使用“对应的”OCR-B 字符来表示。

在机读区将以阿拉伯文字表示的姓名转化为拉丁字符，而所使用的字符被限制为 OCR-B 的 A-Z 和填充符，这种做法是有问题的。此外，允许使用基于语音的音译将会引致的不确定性意味着数据库搜索可能变得徒劳无益。

例如，在上文所举的例子中：

用阿拉伯文字书写的姓名：**ابو بكر محمد بن زكريا الرازي**

1. Abū Bakr al-Rāzi 是大约 1100 年前波斯的一位伟大的科学家和医生。在波斯文（Farsi）中，他的名字的最后两个波斯字母通常拼为“yeh”（ی），但为了避免引起混乱，我们使用的是“yeh”（ي）这一标准的阿拉伯文拼写方式。

2. 见 <http://www.kanji.org/cjk/arabic/araborth.htm>。

在机读区将其音译成拉丁字符的一种方式:

ABU<BAKR<MOHAMMED<IBN<ZAKARIA<AL<RAZI

但是, 仅“Mohammed”这个名字机读区就可能有至少 32 种不同的表示方法。“Zakaria”可以写作“Zakariya”; “ibn”可以写作“bin”; “al”可以写作“el”。仅这些不同的表达方式就可以产生 256 种备选结果。

与此形成对比, 使用 Buckwalter 表(见下文)对例如上述名字محمد中的四个阿拉伯字符进行音译的结果是“mHmd”。在这种情况下, 每个阿拉伯字符对应一个拉丁字符。不考虑语音上的对应。

用 Buckwalter 表对上述姓名进行完整转写的结果是:

Abw<bAkr<mHmd<bn<zkryAY<AlrAzY

遗憾的是, Buckwalter 表会用到小写形式(a-z)和特殊字符(‘,|,>,\$,<,},*,_~,~), 因此, 不适合用于机读区(见 <http://www.qamus.org/transliteration.htm>)。

B.3 对视读区的建议

B.3.1 视读区内字符的音译

如上文所述, Doc 9303 号文件第 3 部分第 3.1 段要求在使用非拉丁文字的本国文字时在视读区加入“转写”。与此项要求相关的 Doc 9303 号文件第 3 部分第 3.4 段特别说明了姓名方面的要求。

“转写”和“音译”这两个词容易引起混淆。“转写”是指严格地对非拉丁文字进行逐字翻译。“音译”是一种不那么严格的表示方法, 一般以语音为基础(在说出姓名时所“发出的音”)。当然, 有时候用一种语文发出的音在另一种语文中并没有与之相对等的发音, 并且要取决于目标语文的情况, 例如, “ch”、“sh”和“th”在英文、法文和德文中的读法是不一样的。可以比较一下数学家和诗人 عمر خيام 的姓名的英文音译“Omar Khayyam”和德文音译“Omar Chajjam”。

有许多种“音译”方案:

- 德国标准化协会: DIN31635 (1982 年)
- 德国东方学会 (1936 年)
- 国际标准化组织: ISO/R233 (1961 年)、ISO233 (1984 年) [3]、ISO233-2 (1993 年)
- 英国标准协会: BS4280 (1968 年)
- 联合国地名专家组 (UNGEGN): UN (1972 年) [4]
- Qalam (1985 年)
- 美国图书馆协会—国会图书馆: ALA-LC (1997 年) [1]
- 伊斯兰百科全书 (新版): EI (1960 年) [2]

一些国家在出生或公民登记簿中用阿拉伯文字和拉丁文字记录其公民的姓名, 而拉丁文字姓名是经核准的对阿拉伯文字姓名的音译。这些国家似宜继续在视读区纳入经核准的拉丁文字音译字符。

建议

如上文所述，Doc 9303 号文件第 3 部分的第 3.1 段和第 3.4 段要求在视读区提供对等的拉丁字符，因此，签发国可自行决定是采取基于语音的音译，还是沿用机读区的转写做法（如下所述）。

B.3.2 音译方案

一些音译方案见下文所示：

统一码	阿拉伯字母	名称 ³	DIN 31635	ISO 233	UN GEGN	ALA-LC	EI	
0621	ء	hamza	'	'	'	'	'	
0622	أ	alef 上加 madda	'ā	'ā	ā	ā	Ā	
0627	ا	alef	Ā	'				
0628	ب	beh	B	b	b	b	B	
0629	ة	teh marbuta	h,t	ṭ	h,t	h,t	a,at	
062A	ت	teh	T	t	t	t	<u>T</u>	
062B	ث	theh	<u>T</u>	ṭ	th	th	<u>Th</u>	
062C	ج	jeem	Ĝ	ġ	j	j	<u>Dj</u>	
062D	ح	hah	ḥ	ḥ	ḥ	ḥ	ḥ	
062E	خ	khah	ḫ	ḫ	kh	kh	<u>Kh</u>	
062F	د	dal	D	d	d	d	D	
0630	ذ	thal	<u>D</u>	ḏ	dh	dh	<u>Dh</u>	
0631	ر	reh	R	r	r	r	R	
0632	ز	zain	Z	z	z	z	Z	
0633	س	seen	S	s	s	s	S	
0634	ش	sheen	Š	š	sh	sh	Sh	
0635	ص	sad	Ṣ	ṣ	ṣ	ṣ	ṣ	
0636	ض	dad	Ḍ	ḏ	ḏ	ḏ	ḏ	
0637	ط	tah	ṭ	ṭ	ṭ	ṭ	ṭ	
0638	ظ	zah	Ẓ	ẓ	ẓ	ẓ	ẓ	
0639	ع	ain	'	'	'	'	'	
063A	غ	ghain	Ĝ	ġ	gh	gh	<u>Gh</u>	
0640	.	tatwheel	[图形填充符，无音译]					
0641	ف	feh	F	f	f	f	F	
0642	ق	qaf	Q	q	q	q	ḳ	
0643	ك	kaf	K	k	k	k	K	
0644	ل	lam	L	l	l	l	L	
0645	م	meem	M	m	m	m	M	
0646	ن	noon	N	n	n	n	N	
0647	ه	heh	H	h	h	h	H	
0648	و	waw	W	w	w	w	W	
0649	ي	alef maksura	Ā	ỳ	y	y	Ā	
064A	ي	yeh	Y	y	y	y	Y	

3. 统一码和 ISO/IEC 10646 中给出的名称。

转写到机读区是 CANXXON<<TERESA

虽然机读区的表示方式看上去不太美观（并且可能招致抱怨），但这种做法是为了便于机读，使得原来的姓名可以通过数据库搜索等手段来查找。因此，机读区把这一姓名定为 CAÑON，从而与 CANON 区分开来。

B.4.2 统一码的使用

在内部处理上，计算机使用编码方案来表示不同语文的字符。常用的一种编码方案是统一码，它与 ISO/IEC 10646 标准几乎相同（下文的表格中使用了统一码的字符索引）。

阿拉伯文字所有字符的表示方式都可以在统一码中找到。统一码字符索引通常表示为一个四位的十六进制数（十六进制以 16 作为基数，用数字 0-9 和字母 A-F 表示 16 个可能的数字）。所有阿拉伯字符都在第 06 行，因此，06 就是数字的前两位（即 06XX）。

例如：

ابو بكر محمد بن زكريا الرازي

可用统一码将其编码为：

ابو Alef (ا) - Beh (ب) - Waw (و) => 0627 + 0628 + 0648
 بكر Beh (ب) - Kaf (ك) - Reh (ر) => 0628 + 0643 + 0631
 محمد Meem (م) - Hah (ح) - Meem (م) - Dal (د) => 0645 + 062D + 0645 + 062F
 بن Beh (ب) - Noon (ن) => 0628 + 0646
 زكريا Zain (ز) - Kaf (ك) - Reh (ر) - Yeh (ي) - Alef (ا) => 0632 + 0643 + 0631 + 064A + 0627
 الرازي Alef (ا) - Lam (ل) - Reh (ر) - Alef (ا) - Zain (ز) - Yeh (ي) =>
 0627 + 0644 + 0631 + 0627 + 0632 + 064A

B.5 关于机读区的建议

B.5.1 影响机读区转写的因素

Doc 9303 号文件第 3 部分的第 4.1 段指出：“……机读区可为视读区的信息提供验证，并可用来为数据库查询提供搜索字符。”第 4.1 段还指出“机读区的数据经格式化后，全世界具有标准能力的机器都可以阅读”，以及“机读区的数据表示法不同于视读区的数据表示法。”但是，第 4.2 段指出，“机读区的数据必须是既可机读，又可视读的”。

这样做的目的是在机读区中将阿拉伯姓名转写为对应的拉丁字符，确保姓名只有一种可能的表示方式。这对于避免歧义以及使数据库和敏感人物列表搜索尽可能准确以便确保识别结果的可靠性是有必要的。同时，机读区必须以尽可能容易辨认的方式来表示视读区所显示的姓名，以便使机读区的数据是可视读的，方便用于进行先进的旅客处理和类似目的。

B.5.2 现有的转写方案

现有几种转写方案：标准阿拉伯文技术转写系统（SATTS）、Buckwalter 和 ASMO 449。这些方案见下文所示：

统一码	阿拉伯字符	名称	SATTS	Buckwalter	ASMO 449
0621	ء	hamza	E	'	A
0622	آ	alef 上加 madda	(无)		B
0623	أ	alef 上加 hamza	(无)	>	C
0624	ؤ	waw 上加 hamza	(无)	&	D
0625	إ	alef 下加 hamza	(无)	<	E
0626	ئ	yeh 上加 hamza	(无)	}	F
0627	ا	alef	A	A	G
0628	ب	beh	B	b	H
0629	ة	teh marbuta	?	p	I
062A	ت	teh	T	t	J
062B	ث	theh	C	v	K
062C	ج	jeem	J	j	L
062D	ح	hah	H	H	M
062E	خ	khah	O	x	N
062F	د	dal	D	d	O
0630	ذ	thal	Z	*	P
0631	ر	reh	R	r	Q
0632	ز	zain	:	z	R
0633	س	seen	S	s	S
0634	ش	sheen	:	\$	T
0635	ص	sad	X	S	U
0636	ض	dad	V	D	V
0637	ط	tah	U	T	W
0638	ظ	zah	Y	Z	X
0639	ع	ain	"	E	Y
063A	غ	ghain	G	g	Z
0640	ـ	tatwheel	(无)	_	0x60
0641	ف	feh	F	f	A
0642	ق	qaf	Q	q	B
0643	ك	kaf	K	k	C
0644	ل	lam	L	l	D
0645	م	meem	M	m	E
0646	ن	noon	N	n	F
0647	ه	heh	?	h	G
0648	و	waw	W	w	H
0649	ى	alef maksura	(无)	Y	I
064A	ي	yeh	I	y	J
064B	◌◌◌	fathatan	(无)	F	K
064C	◌◌◌◌	dammatan	(无)	N	L
064D	◌◌◌◌◌	kasratan	(无)	K	M
064E	◌◌◌◌◌◌	fatha	(无)	a	N
064F	◌◌◌◌◌◌◌	damma	(无)	u	O
0650	◌◌◌◌◌◌◌◌	kasra	(无)	i	P
0651	◌◌◌◌◌◌◌◌◌	shadda	(无)	~	Q
0652	◌◌◌◌◌◌◌◌◌◌	sukun	(无)	o	R
0670	◌◌◌◌◌◌◌◌◌◌◌	用作上标的 alef	(无)	`	(无)

仔细查看表格可以看出，这些方案所用的拉丁字符超出了 A-Z 的范围，因此，从根本上说不适合用于机读区中。

ASMO 449 方案在分配拉丁字符时比较随意，而 Buckwalter 方案将一些有对等发音的字符写的很相似。

标准阿拉伯文技术转写系统未对 heh (ه)和 teh marbuta (ة)或最后的 yeh (ي)和 alif maksura (آ)进行区分，并且这一系统无法对 alif madda (إ)进行转写。

B.5.3 其他需考虑的方面

在使用所建议的转写方案时，不能不考虑到机读旅行证件的运作环境。特别是，机读区的姓名应该在外形和格式上尽可能地接近其他来源所提供的姓名。航空公司所使用的并通过旅客信息预报（API）方案提交给移民当局的旅客姓名记录（PNR）就是一个例子。虽然机读区中的转写几乎总是无法与视读区中的音译（以及旅客姓名记录等基于语音的其他表示方法）完全一致，但此处所建议的方案试图使两个区中的姓名看起来是明显相似的。

为此，与基于拉丁语的多国字符转写表中的做法一样，使用了“X”这一字符作为“换码”字符，区别在于这里只用到一个“X”，并且用在其要修饰的字符之前，而非字符之后（如“XTH”对“NXX”）。在每个“X”之后跟上一个或两个字符来代表一个阿拉伯字母。这样使用“X”是可行的，因为现有的阿拉伯音译和转写方案中并不存在“X”。

[“X”在阿拉伯字符和拉丁字符转写中的不同用法不太可能造成混淆。为了可以顺利地进行逆向转写，必须明确源文字，且最好以签发国为基础。]

在一些转写条目中，在开头的“X”之后会加上另一个“X”：例如，alef 上加 madda 表示为“XAA”，alef wasla 表示为“XXA”。采用这一方法主要是为了避免引入会造成机读区不易肉眼阅读的其他字符。

这一做法的目的是让从现有系统中查看机读区原始数据的操作人员忽略任何“X”字符。所得到的姓名应该与其他来源所提供的姓名相类似。机读区原始数据还将缺少元音，而视读区中的音译和旅客姓名记录等其他来源一般会包含元音。但是，如果操作人员被告知元音未被包括在内，那么他们就应该认为机读区数据是基于语音的一种合理音译。

转写还将不包括“太阳字母”之前冠词的同化（sandhi），因为这本质上是一种语音现象，因此拼写可能与视读区的基于语音的音译对不上（例如，在视读区中，“AL-RAZI”可能写为“AR-RAZI”）。

“shadda”（用于表示将字母写两次的标志）意味着带有该标志的字符在机读区中将被重写一次（两次）。搜索算法应该考虑到“shadda”并不会总是存在。

B.5.4 标准阿拉伯文的转写建议方案

以 Buckwalter 转写表为基础，并考虑到各种音译方案（第 B.3.2 段）所列出的常见的基于语音的对等字符，形成了仅使用拉丁字符 A-Z 的一个转写建议方案。由于有用“X”表示不同字母的先例（第 B.5.3），这里将“X”这一字符用作一个“换码”字符，用于表示跟在“X”之后的一个或两个字符代表一个阿拉伯字母。

06A5	پ	feh 下加 3 个点	XF	
------	---	-------------	----	--

注 1: 没有编码

B.5.5 关于转写表的说明

B.5.5.1 Alef 上加 madda

ALA-LC 罗马化表[1]未对 alef 上加 madda (اِ) 进行表示。但是, Interpol [5]和 Dr Hoogland [6]都建议将其转写为 XAA。

B.5.5.2 Alef 上加 hamza

ALA-LC 罗马化表[1]未对 alef 上加 hamza (أ) 进行表示。但是, Interpol [5]建议将其转写为 XAE。

B.5.5.3 Waw 上加 hamza

ALA-LC 罗马化表[1]未对 waw 上加 hamza (وْ) 进行表示。由于 waw 上加 hamza 通常音译为“U”, 因此这里使用 U 进行表示。

B.5.5.4 Alef 下加 hamza

ALA-LC 罗马化表[1]未对 alef 下加 hamza (اِ) 进行表示。这里所使用的转写方式是 I, 因为该拉丁字母并未在别处使用, 而且 alef 下加 hamza 通常出现在 ابراهيم (Ibrahim) 等名字的开头, 这种情况中的 alef 下加 hamza 通常音译为“I”。

B.5.5.5 Yeh 上加 hamza

ALA-LC 罗马化表[1]未对 yeh 上加 hamza (يْ) 进行表示。这里所使用的转写方式是 XI, 因为 yeh 上加 hamza 用于 فانز (Faiz) 等名字中, 而这种情况中的 yeh 上加 hamza 通常音译为“I”。

B.5.5.6 Teh marbuta

Teh marbuta (ة) 在 ALA-LC 罗马化表[1]中表示为 H 或 T 或 TAN, 具体取决于相关语境。Dr Hoogland [6]建议将其表示为 XTA。这里, teh marbuta 有两种不同的转写方式: 一般情况下都使用 XTA, teh marbuta 出现在名字成分的末尾时除外, 这时使用的是 XAH。这是因为女性姓名通常使用 teh marbuta 来修饰男性姓名, 如 فاطمة (Fatimah)。搜索算法应该将这两种可能性纳入考虑。

B.5.5.7 Hah and heh

根据 Interpol [5]的建议，将 hah (ح) 和 heh (ه) 的转写方式进行了对调。Hah 现在表示为 XH，而 heh 表示为 H。

B.5.5.8 Tatwheel

Tatwheel (ٲ) 是一个图形符号，未进行转写。

B.5.5.9 Alef maksura

根据 Dr Hoogland [6]的建议，现在将 alef maksura (ع) 转写为 XAY。其他字符转写为 XY_，这便使得之前的 XY 变得不协调。

B.5.5.10 短元音 fatha、damma、kasra、fathatan、dammatan 和 kasratan

选择性的短元音 (haracat) 一般不用于姓名中，因此未对其进行转写。

B.5.5.11 Shadda

Shadda (ّ) 表示其下面的辅音发两次音，因此它的转写方式是将对应的字符写两次。搜索算法应该注意，shaddah 是选择性的，字符有时候会写两次，有时则不会。

注意 ﷲ (Allah) 这种特殊情况。

B.5.5.12 Sukun

Sukun (◌ْ) 表示未将元音包括在内，它是选择性的，因此未对其进行转写。

B.5.5.13 用作上标的 alef

未对用作上标的 (◌ِ) (“元音-剑号-alef”) 进行转写。

B.5.5.14 Alef wasla

根据 Interpol [5]的建议，现在将 alef wasla (أ) 转写为 XXA。其他字符转写为 XA_，这便使得之前的 XA 变得不协调。Dr Hoogland [6]也建议将其转写为 XXA。

B.5.6 其他语文的转写建议方案

讲波斯文的国家有伊朗(波斯文)、阿富汗(达里文)、塔吉克斯坦和乌兹比克斯坦。

讲普什图文的国家有阿富汗和巴基斯坦西部。

讲乌尔都文的国家有巴基斯坦和印度。

统一码	阿拉伯字母	语文	名称	机读区
0679	ٹ	乌尔都文	tteh	XXT
067E	پ	波斯文、乌尔都文	peh	P
067C	ت	普什图文	teh 加小圆圈	XRT
0681	خ	普什图文	hah 上加 hamza	XKE
0685	ځ	普什图文	hah 上加 3 个点	XXH
0686	چ	波斯文、乌尔都文	tcheh	XC
0688	ڈ	乌尔都文	ddal	XXD
0689	د	普什图文	dal 加小圆圈	XDR
0691	ڑ	乌尔都文	rreh	XXR
0693	ر	普什图文	reh 加小圆圈	XRR
0696	ړ	普什图文	reh 上下各加一个点	XRX
0698	ژ	波斯文、乌尔都文	jeh	XJ
069A	ښ	普什图文	seen 上下各加一个点	XXS
06A9	ک	波斯文、乌尔都文	keheh	XKK
06AB	ک	普什图文	kaf 加小圆圈	XXK
06AD	ښ		ng	XNG
06AF	گ	波斯文、乌尔都文	gaf	XGG
06BA	ں	乌尔都文	noon ghunna	XNN
06BC	ن	普什图文	noon 加小圆圈	XXN
06BE	ھ	乌尔都文	heh doachashmee	XDO
06C0	ه	乌尔都文	heh 上加 yeh	XYH
06C1	ہ	乌尔都文	heh goal	XXG
06C2	ه	乌尔都文	heh goal 上加 hamza	XGE
06C3	ة	乌尔都文	teh marbuta goal	XTG
06CC	ی	波斯文、乌尔都文	farsi yeh	XYA ⁵
06CD	ی	普什图文	yeh 加小尾巴	XXY
06D0	ی	普什图文	yeh	Y ⁶
06D2	ے	乌尔都文	yeh barree	XYB
06D3	ئ	乌尔都文	yeh barree 上加 hamza	XBE

B.5.7 标准阿拉伯文的转写示例

上文的例子

ابو بكر محمد بن زكريا الرازي

在机读区中可编码为:

ابو Alef (ا) - Beh (ب) - Waw (و) => ABW
 بكر Beh (ب) - Kaf (ك) - Reh (ر) => BKR
 محمد Meem (م) - Hah (ح) - Meem (م) - Dal (د) => MXHMD
 بن Beh (ب) - Noon (ن) => BN

5. “farsi yeh” (ی) 这一字母在功能上等同于标准的“yeh” (ي)，但在单独使用和用在末尾时在形状上等同于标准的“alef maksura” (ى)，因此，可以转写为“Y”或“XAY”。数据库匹配算法应该将这一点纳入考虑。

6. “Pashto yeh” (ي) 这一字母在功能上等同于标准的“yeh” (ي)。

زكريا Zain (ز) - Kaf (ك) - Reh (ر) - Yeh (ي) - Alef (ا) => ZKRYA
 الرازي Alef (ا) - Lam (ل) - Reh (ر) - Alef (ا) - Zain (ز) - Yeh (ي) => ALRAZY

即 ABW<BKR<MXHMD<BN<ZKRYA<ALRAZY

这种转写方式的优点是:

1. 用阿拉伯文字表示的姓名总是可以转写为同一种拉丁表示方法。这意味着更有可能实现数据库匹配;
2. 该过程是可以逆向操作的—可以复原到用阿拉伯文字表示的姓名。

复原用阿拉伯文字表示的姓名:

ABW A=Alef (ا) - B=Beh (ب) - W=Waw (و) => ابو
 BKR B=Beh (ب) - K=Kaf (ك) - R=Reh (ر) => بكر
 MXHMDM=Meem (م) - XH=Hah (ح) - M=Meem (م) - D=Dal (د) => محمد
 BN B=Beh (ب) - N=Noon (ن) => بن
 ZKRYA Z=Zain (ز) - K=Kaf (ك) - R=Reh (ر) - Y=Yeh (ي) - A=Alef (ا) => زكريا
 ALRAZY A=Alef (ا) - L=Lam (ل) - R=Reh (ر) - A=Alef (ا) - Z=Zain (ز) - Y=Yeh (ي) => الرازي

略去 *harakat* 和其他变音符号的理由是, 这些符号是选择性的, 基本上不会用到。因此, 这些符号的处理方式应该与欧洲国家字符上的变音符号 (如 é、è、ç) 相同, 这些变音符号是作为发音之用的。

同样地, *harakat* 是选择性的这一特点会对数据库的准确匹配性产生不利影响。

机读区	阿拉伯字母名称	阿拉伯字母	统一码
XAE	alef 上加 hamza	أ	0623
XAH	teh marbuta (另见 XTA)	ة	0629
XAY	alef maksura	ى	0649
XBE	yeh barree 上加 hamza	ے	06D3
XC	tcheh (波斯文、乌尔都文)	چ	0686
XDH	thal	ذ	0630
XDO	heh doachashmee	ھ	06BE
XDR	dal 加小圆圈 (普什图文)	ڊ	0689
XDZ	dad	ض	0636
XE	hamza	ء	0621
XF	feh 下加 3 个点	پ	06A5
XGG	gaf (波斯文、乌尔都文)	گ	06AF
XGE	heh goal 上加 hamza (乌尔都文)	ه	06C2
XH	hah	ح	062D
XI	yeh 上加 hamza	ی	0626
XJ	jeh (乌尔都文)	ژ	0698
XKE	hah 上加 hamza (普什图文)	خ	0681
XKH	khah	ځ	062E
XKK	keheh (波斯文、乌尔都文)	ک	06A9
XNN	noon ghunna (乌尔都文)	ن	06BA
XNG	ng	ڱ	06AD
XRR	reh 加小圆圈 (普什图文)	ر	0693
XRT	teh 加小圆圈	ڙ	067C
XRX	reh 上下各加 1 个点 (普什图文)	ږ	0696
XSH	sheen	ش	0634
XSS	sad	ص	0635
XTA	teh marbuta (see also XAH)	ة	0629
XTG	teh marbuta goal (乌尔都文)	ه	06C3
XTH	theh	ث	062B
XTT	tah	ط	0637
XXA	alef wasla	أ	0671
XXD	ddal (乌尔都文)	ڈ	0688
XXG	heh goal (乌尔都文)	ه	06C1
XXH	hah 上加 3 个点 (普什图文)	ځ	0685
XXK	kaf 加小圆圈 (普什图文)	ک	06AB
XXN	noon 加小圆圈 (普什图文)	ن	06BC
XXR	rreh (乌尔都文)	ڙ	0691
XXS	seen 上下各加 1 个点 (普什图文)	ښ	069A
XXT	tteh (乌尔都文)	ٹ	0679
XXY	yeh 加小尾巴 (普什图文)	ی	06CD
XYA	farsi yeh (波斯文、乌尔都文)	ی	06CC
XYB	yeh barree (乌尔都文)	ے	06D2
XYH	heh 上加 yeh (乌尔都文)	ه	06C0
XZZ	zah	ظ	0638

B.7 计算机程序

B.7.1 阿拉伯字符转写为机读区字符

这一用 Python 进行编写的程序是将阿拉伯字符（用统一码表示）转换为机读区格式的一个示例。

阿拉伯字符被包含在“Arabic souce.txt”这样一个文件中，而对应的机读区数据被写入“MRZ output.txt”这样的文件中。

```
*****
# # -*- coding: iso-8859-15 -*-

import unicodedata
import encodings.utf_8_sig
import codecs

# TRANSLITERATE
def Arabic_to_MRZ(unicode_string):
    transform = {0x20: '<', 0x21: 'XE', 0x22: 'XAA', 0x23: 'XAE', 0x24: 'U',
                 0x25: 'I', 0x26: 'XI', 0x27: 'A', 0x28: 'B', 0x29: 'XAH',
                 0x2A: 'T', 0x2B: 'XTH', 0x2C: 'J', 0x2D: 'XH', 0x2E: 'XKH',
                 0x2F: 'D', 0x30: 'XDH', 0x31: 'R', 0x32: 'Z', 0x33: 'S', 0x34: 'XSH',
                 0x35: 'XSS', 0x36: 'XDZ', 0x37: 'XTT', 0x38: 'XZZ', 0x39: 'E',
                 0x3A: 'G', 0x41: 'F', 0x42: 'Q', 0x43: 'K', 0x44: 'L',
                 0x45: 'M', 0x46: 'N', 0x47: 'H', 0x48: 'W', 0x49: 'XAY',
                 0x4A: 'Y', 0x71: 'XXA', 0x79: 'XXT', 0x7E: 'P', 0x7C: 'XRT',
                 0x81: 'XKE', 0x85: 'XXH', 0x86: 'XC', 0x88: 'XXD', 0x89: 'XDR',
                 0x91: 'XXR', 0x93: 'XRR', 0x96: 'XRX', 0x98: 'XJ', 0x9A: 'XXS',
                 0xA4: 'XV', 0xA5: 'XF', 0xA9: 'XKK', 0xAB: 'XXK', 0xAD: 'XNG',
                 0xAF: 'XGG', 0xBA: 'XNN', 0xBC: 'XXN', 0xBE: 'XDO', 0xC0: 'XYH',
                 0xC1: 'XXG', 0xC2: 'XGE', 0xC3: 'XTG',
                 0xCC: 'XYA', 0xCD: 'XXY', 0xD0: 'Y', 0xD2: 'XYB', 0xD3: 'XBE'}
    name_in = unicode_string
    name_out = ""
    for c in name_in:
# check for shadda (double)
        if ord(c) == 0x51:
            name_out = name_out + char
        else:
            if ord(c) in transform:
                char = transform[ord(c)]
                name_out = name_out + char
    print name_out
    return name_out

#
# MAIN - Arabic to MRZ
#

# open input and output files
```

```

fin = encodings.utf_8_sig.codecs.open('Arabic source.txt', 'r') #b, 'utf-8-sig', 'ignore', 1)
fout = open('MRZ output.txt', 'w')

# loop through the input file

try:
    for arabic_name in fin:
        MRZ_name = Arabic_to_MRZ(arabic_name)
        fout.write(MRZ_name)
        fout.write('\n')
finally:
    fin.close()
fout.flush()
fout.close()

*****

```

B.7.2 机读区字符转写为阿拉伯字符

这一用 Python 进行编写的程序是将机读区字符转换为阿拉伯字符（用统一码表示）的一个示例。

机读区字符被包含在“MRZ souce.txt”这样一个文件中，而对应的阿拉伯数据被写入“Arabic output.txt”这样的文件中。

```

*****

# # -*- coding: iso-8859-15 -*-

import unicodedata
import encodings.utf_8_sig
import codecs

# TRANSLITERATE
def MRZ_to_Arabic(ascii_string):
    transform = { '<': 0x20, 'XE': 0x21, 'XAA':0x22, 'XAE': 0x23, 'U': 0x24,
                  'I': 0x25, 'XI': 0x26, 'A': 0x27, 'B': 0x28, 'XAH': 0x29,
                  'T': 0x2A, 'XTH': 0x2B, 'J': 0x2C, 'XH': 0x2D, 'XKH': 0x2E,
                  'D': 0x2F, 'XDH': 0x30, 'R': 0x31, 'Z': 0x32, 'S': 0x33, 'XSH': 0x34,
                  'XSS': 0x35, 'XDZ': 0x36, 'XTT': 0x37, 'XZZ': 0x38, 'E': 0x39,
                  'G': 0x3A, 'F': 0x41, 'Q': 0x42, 'K': 0x43, 'L': 0x44, 'M': 0x45,
                  'N': 0x46, 'H': 0x47, 'W': 0x48, 'XAY': 0x49, 'Y': 0x4A, 'XXA': 0x71,
                  'XXT': 0x79, 'P': 0x7E, 'XRT': 0x7C, 'XKE': 0x81, 'XXH': 0x85,
                  'XC': 0x86, 'XXD': 0x88, 'XDR': 0x89, 'XXR': 0x91, 'XRR': 0x93,
                  'XRX': 0x96, 'XJ': 0x98, 'XXS': 0x9A, 'XV': 0xA4, 'XF': 0xA5,
                  'XKK': 0xA9, 'XXK': 0xAB, 'XNG': 0xAD, 'XGG': 0xAF,
                  'XNN': 0xBA, 'XXN': 0xBC, 'XDO': 0xBE, 'XYH': 0xC0,
                  'XXG': 0xC1, 'XGE': 0xC2, 'XTA': 0x29, 'XTG': 0xC3, 'XYA': 0xCC,
                  'XXY': 0xCD, 'I': 0xD0, 'XYB': 0xD2, 'XBE': 0xD3}

    name_in = ascii_string
    name_out = ""

```

```
# if this character is not X, does it appear by itself in the table?
search_string = ""
last_string = ""
iloop = 0
while iloop < len(name_in):
    search_string = search_string + name_in[iloop]
    if search_string in transform:
        if search_string <> last_string:
            name_out = name_out + chr((transform[search_string]))
            #insert shadda if double found
        else:
            name_out = name_out + chr(0x51)
        if search_string <> '<':
            name_out = name_out + chr(0x06)
        else:
            name_out = name_out + chr(0x00)
        #remember last string
        if search_string <> '<':
            last_string = search_string
        else:
            last_string = ""
        #clear the search string once found
        search_string = ""
    iloop = iloop + 1
print name_out
return name_out

#
# MAIN - MRZ to Arabic
#

# open input and output files

fin = open('MRZ source.txt', 'r')
fout = open('Arabic output.txt', 'wb') #b', 'utf-8-sig', 'strict', 1)
fout.write(encodings.utf_8_sig.codecs.BOM)

# loop through the input file

try:
    for MRZ_name in fin:
        Arabic_name = MRZ_to_Arabic(MRZ_name)
        Arabic_name = Arabic_name + chr(0x0D) + chr(0x00) + chr(0x0A) + chr(0x00)
        fout.write(Arabic_name)
finally:
    fin.close()
fout.flush()
fout.close()
```

```
*****
```

B.8 参考文献(资料性)

- [1] ALA-LC 罗马化表：非罗马文字转写方案，Randal K. Berry (编辑)，美国国会图书馆，1997 年。
- [2] 《伊斯兰百科全书》，新版，莱顿，1960 年。
- [3] ISO 233:1984：《文件 — 将阿拉伯字符转写为拉丁字符》，国际标准化组织，1984 年 12 月 15 日。
- [4] 《联合国地名罗马化系统：使用现状报告》，由联合国地名专家组罗马化系统工作组编制，第 2.1 版，2002 年 6 月。
- [5] 创新公共服务小组对发布日期为 2008 年 2 月 15 日的文件 — 《机读旅行证件中阿拉伯文字的转写 — 技术报告 — 第 2.3 版》的意见，国际刑警组织，里昂，2008 年 3 月 17 日。
- [6] 私人信件，Jan Hoogland 博士，奈梅亨大学阿拉伯语系，荷兰，2008 年 3 月 23 日。
- [7] 对《机读旅行证件中阿拉伯文字的转写 — 技术报告 AMA 13052008》的意见，Mr. Abdalla M. Askar，阿联酋身份管理局。

