TECHNICAL SPECIFICATION

ISO/IEC TS 20071-15

First edition
2017-12

Information technology — User interface component accessibility —

Part 15: Guidance on scanning visual information for presentation as text in various modalities

Technologies de l'information — Accessibilité du composant interface utilisateur —

Partie 15: Recommandations relatives à la numérisation des informations visuelles en vue d'une présentation sous forme de texte selon différentes modalités

Reference number
ISO/IEC TS 20071-15:2017(E)

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

This document was prepared by ISO/IEC JTC 1, *Information technology*, SC 35, *User interfaces*.

A list of all parts in the ISO/IEC 20071 series can be found on the ISO website.

Introduction

There are a wide variety of visual objects in a user's environment that provide information that could be scanned and processed to output text-based information about or related to the object. Providing text-based information can provide accessibility to diverse users in various contexts of use, including:

— persons who cannot see the information (due to vision or environmental limitations);

— persons who cannot approach the information closely enough to see it (due to physical or environmental limitations);

— people who cannot understand the information (due to cognitive or linguistic limitations);

— where the information is provided in a format that human cannot directly understand (e.g. barcodes, QR codes);

— where information content, beyond what is in the scanned image, can be obtained through additional processing.

This document provides guidance on various aspects of the user interface of applications that scan visual information. This includes selecting the types of information that are of interest to the user, helping the user locate visual objects of interest, helping the user position the object or device used for scanning, scanning the information, processing the scanned image, and outputting textual information in various modalities.

The guidance contained in this document can be applied to a wide range of devices, applications, and contexts of use.

Information technology — User interface component accessibility —

Part 15: Guidance on scanning visual information for presentation as text in various modalities

1 Scope

This document provides guidance on various aspects of the user interface of applications that scan visual information that are used directly by humans, including:

— initiating the scanning application;

— setting user's preferences and configuring the scanning application;

— identifying the types of information currently of interest to the user;

— locating visual objects of interest to the user;

— creating a static image via scanning the visual object;

— identifying the information content provided by the visual object;

— processing scanned information and outputting the results to the user.

This document provides increased accessibility by addressing the user accessibility needs of diverse users in diverse contexts.

This document contains guidance that can be applied to a variety of devices, including:

— specialized devices that are dedicated to scanning and processing visual information;

— mobile devices (such as smartphones and tablets);

— general purpose computers with camera capabilities;

— office machines with scanning functions.

This document contains guidance that can be applied to various types of software, including:

— stand-alone scanning applications;

— applications including scanning functionalities;

— (scanning) applications that interoperate with other applications.

This document contains guidance that can be used for outputting scanned information in various modalities, including:

— audio outputs;

— visual outputs;

— tactile outputs;

— storing information for future use within the application performing the scanning;

— electronic outputs (to other applications, systems, or devices including those directly connected and those connected via the Internet).

This document does not:

— apply to fully automated scanning that is not under direct human control;

— apply to applications that scan visual information for editing an image or just displaying it;

— provide guidance on the design of specific hardware devices involved in scanning;

— provide guidance on the specific objects that can be recognized or the specific software used to recognize these objects;

— provide guidance on the internal functioning of software that recognizes specific types of objects.

This document is intended for use by developers of applications that include user controlled scanning functionalities. It does not expect that an application includes all of these functionalities. It can be used for those functionalities that an application does provide.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— IEC Electropedia: available at http://www.electropedia.org/

— ISO Online browsing platform: available at https://www.iso.org/obp

3.1
scanning
use of a camera or similar function to obtain image data

3.2
recognition
<scanning> processing of *scanned image* (3.6) to locate and identify visual *objects of interest* (3.4)

3.3
visual information
recognizable information presented by a visual object

3.4
object of interest
type of object that has been chosen by a user

3.5
information of interest
type of information that has been chosen to be processed by a *scanning* (3.1) system

3.6
scanned image
data resulting from *scanning* (3.1)

3.7
scanning application
software that scans visual objects, recognizes and processes the *objects of interest* (3.4), and outputs the information about the objects

3.8
cropping
<scanning> cutting off part(s) of a *scanned image* (3.6) so that its dimensions, and possibly its information content, are changed

[SOURCE: ISO/IEC 5127:2017, 3.5.9.1.08, modified — "a document or graphic" has been replaced by "a scanned image".]

4 Conformance

It is recognized that all of the functionalities recommended in this document might not be implemented.

If an organization claims conformity of a system with this document, then it shall document:

a) the decisions of which recommendations from this document are and are not applicable to the system;

b) how it addresses each of the applicable recommendations in this document or the justifications for any course of action that deviates from any of the applicable recommendations.

In making such a claim, an organization should be specific about the basis on which it is made and should have evidence to support the claim. The claim can be based on self-assessment or by an assessment carried out by another party.

Annex B provides assistance in documenting conformance to this document.

5 Overview of scanning visual information for presentation as text in various modalities

5.1 General

Humans recognize a multitude of objects in the real world and in scanned image, despite the fact that their vision of the objects can vary somewhat from different viewpoints, in different sizes and scales or even when they are translated or rotated.

As a type of computer vision technology, object recognition is the task of finding and identifying the visual objects of interest. Utilizing this technology allows for a wide range of applications that can provide information associated with the object. It can identify, locate, and retrieve information of the physical features of the object, and search for detailed information relevant to the scanned image.

Utilizing this technology can help users with diverse capabilities in diverse contexts to identify and understand objects as well as environments. Users, who will be aided by this technology include:

— users who cannot see the visual object being scanned;

— users who cannot directly understand the visual object being scanned;

— users who cannot sufficiently approach the visual object being scanned.

Contextual examples of use that can be provided by this technology include:

— where the information is not provided in a format that a human can directly understand (e.g. barcodes, QR codes);

— where the information content beyond what is in the scanned image can be obtained through additional processing.

In order to be of use to the widest range of users, the results will be provided by textual or other informative forms of presentation.

NOTE Annex A contains example use cases for scanning visual information for presentation as text.

5.2 Contexts for scanning

A scanning application captures (or scans) visual objects, identifies the visual objects of interest in the scanned images, finds out the visual information of interest, and provides the scanning results to the user.

In everyday life, people can encounter a variety of applications that use scanning functionalities.

EXAMPLE 1 When the customer places the printed barcode of a product under the barcode reader installed on the wall of the store, the customer can get the product information recognized by the reader through the display attached to the reader (see A.1).

EXAMPLE 2 By connecting the document scanner and the refreshable braille display to a PC and activating the optical character recognition (OCR) application, the user can read the content of the document in braille (see A.3).

EXAMPLE 3 By installing and activating a colour-aware application on the mobile device, a person who cannot distinguish colour can recognize the colour of the selected garment without the help of others (see A.4).

EXAMPLE 4 By installing and activating an application on the mobile device that identifies the number of visual objects specified by the user, it is easy to determine how many objects exist in the direction the mobile device camera points (see A.5).

EXAMPLE 5 Someone who cannot see waits for a bus putting on a wearable scanning system that recognizes shapes and text. This person can identify the bus he/she wants to take through the information the system reads by voice (see A.6).

EXAMPLE 6 When a traveller captures a landmark image using a mobile device, on which the scanning application analyses the image and automatically opens the web browser to search information about the landmark, the traveller enjoys the pleasures of travelling more than he/she would without the mobile device and the scanning application (see A.7).

EXAMPLE 7 When a person using the address book management software purchases a business card scanner and scans business cards, he or she then conveniently stores the name and phone number without typing them (see A.8).

5.3 Framework for scanning

The procedure of scanning, which inputs visual objects by scanning, identifies visual objects of interest and generates visual information of interest, shown in Figure 1.

Scanning can be done on a wide range of visual objects including (but not limited to) visual text, graphic codes and/or other identifiable visual objects. Visual text includes text on a printed material, a signboard, a road sign or a business card, etc. Graphic code includes (but not limited to) barcodes, icons, emoticons, traffic signs, etc. Other identifiable visual objects include daily goods, transportation, landmark, etc.

The scanning process consists of three steps.

First is the scan input step, which scans the visual object to obtain the scanned image. In this step, the user positions the visual object of interest within the camera angle, adjusts the size of the object. It might be necessary to adjust the exposure or focus the object to get a good quality image. The scanned image will be used in subsequent steps or can be stored for future use.

Second is the processing step, which extracts the visual information of interest of the scanned image. This step is the core of the scanning application.

Pre-processing is the initiating phase of the processing step, which identifies the visual objects of interest in the scanned image according to the purpose of the scanning application. Subsequently, main-processing extracts the visual information from the scanned (visual) objects.

Post-processing is optionally performed by the user's behest to carry out additional tasks (e.g. translation, database search).

Typically, the processing step proceeds without the user's intervention, once initiated.

Third is the output step, which provides the user with textual information about the visual information of interest extracted in the processing step. Textual information is presented in suitable modalities for user characteristics.

NOTE Some modalities include visual, audio, and tactile presentation.

In some cases, the textual information of the output step is used as an argument when the scanning application invokes other applications without providing the scanning results directly to the user. In this case, the user is provided with the result of the callee instead of the caller.

These three steps can be repeated by request of the users or until the scanned result presents the desired outcome.

Setup defines the user preference settings, required during the scanning process. Specifying the type of visual object of interest is a part of the user preference setting. Defining the output modality is also considered as the user preference setting.

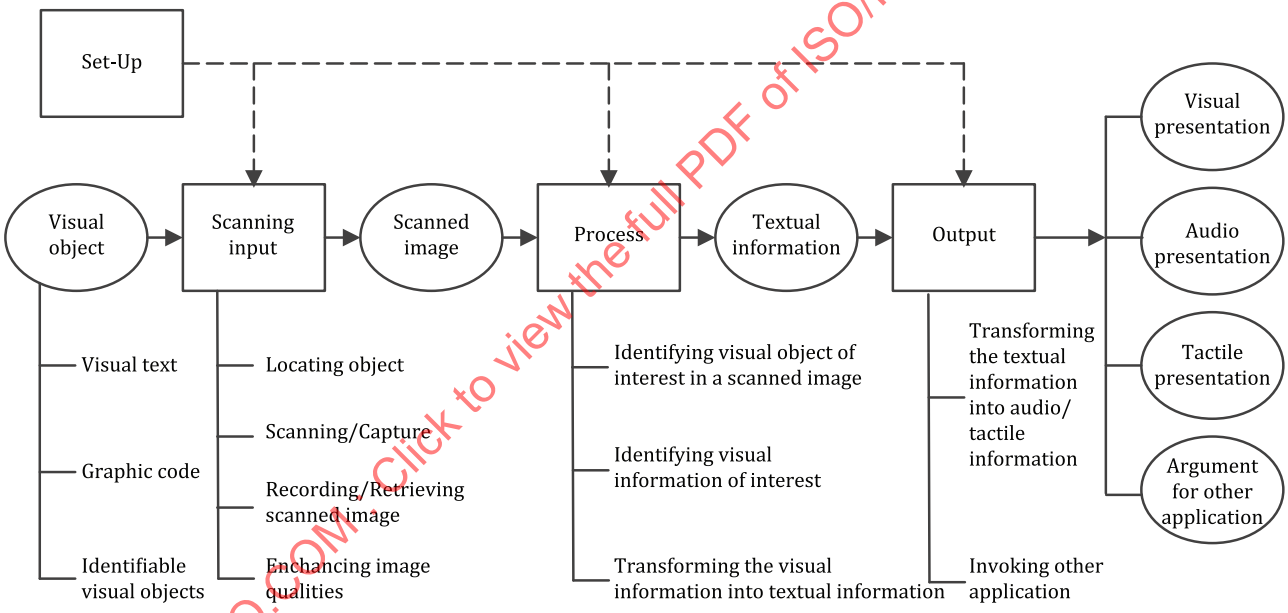

Figure 1 — Procedure of scanning visual information for presentation as text in various modalities

5.4 Types of devices

Typically, a scanning application consists of three components for (scanning) input, processing (scanned input), and outputting (results). Depending on the type of components, a scanning application can be implemented as an all-in-one device or as individual devices that are connected with or without wire connection. Scanning applications can be implemented on:

— specialized scanning devices that are dedicated to scanning and processing visual information;

— mobile devices (such as smartphones and tablets);

— general purpose computers with camera capabilities;

— office machines with scanning function.

NOTE 1 This distinction is not intended to limit the range of usage. There might be other type of devices or system that could be configured.

Specialized devices are those that are dedicated for one or more types of visual information in which the hardware and software are integrated within. Scanning applications can also be implemented on a range of mobile devices or general purpose computers containing cameras with processing capabilities as well as output capabilities. A wider variety of processing can be obtained by connecting to other applications that might be on the device or remotely connected.

The scanning input component consists of the scanning input device, such as a camera or a scanner that is used to obtain scanned images of visual objects.

NOTE 2 Scanners tend to provide their own light for viewing an object and thus are typically used at short distances from the object being scanned.

NOTE 3 Cameras can be used at longer distances from objects and can make use of existing lighting or might provide additional lighting, when needed.

NOTE 4 Regardless of whether a scanner or a camera was used, a scanned image that has been obtained will be referred to as “scanned image” within this document.

NOTE 5 The scanning input component can include a step for increasing the quality of the scanned image. This step can be performed automatically or manually in accordance to the user's instruction.

Scanning components can be fixed (in their location), hand-held or wearable.

EXAMPLE 1 Wall-mounted fingerprint scanners and cameras on laptop computers are some examples of scanning input devices that are typically fixed in a single location where it is easier to take the object to the scanner than the scanner to the object.

EXAMPLE 2 Barcode scanners, OCR pen scanners, smartphone cameras, tablet cameras are some examples of hand-held scanning input devices.

EXAMPLE 3 Cameras mounted on glasses, smart watch cameras, wearable ring cameras are some examples of wearable scanning input devices. The processing component is used to store the scanned information, to recognize objects of interest from the scanned image, and to process recognizable information within the scanned image to generate appropriate information. Data that are used for analysing a scanned image could be stored either in the scanning device itself or in storage, which is connected to the processing component via the Internet or cloud. Likewise, the processing could take place in the scanning device, via remote connection or by a combination of locations.

The processing component includes the scanning application and the hardware on which the application is running. The scanning application of the processing component identifies all the visual objects of interest from the scanned image, received from the scanning input component and then the component recognizes the visual information from these objects.

The processing component is also responsible for providing the user with the ability to configure the scanning application and set user preferences (including identifying the types of information currently of interest to the user).

Various output components (that are either part of the scanning device or that could be connected to the scanning device) can present the information generated by the scanning application.

When an assistive technology (AT) device is installed in the scanning system, the output component provides the information to the user through the AT device.

5.5 Type of software

In everyday life, people often encounter software that uses scanning functionality. Some software itself constitutes a scanning application. Some applications are designed to include scanning functionalities or to interact with an independent scanning application.

Scanning applications can be implemented as:

— stand-alone scanning applications;

— applications including scanning functionalities;

— applications that interoperate with other applications.

NOTE This distinction is not intended to limit the range of usage. There can be other type of software that could be configured.

Stand-alone scanning applications are software that was originally designed to scan and process visual information. In this kind of application, the processing results are mostly delivered directly to the user.

EXAMPLE 1 Scanning applications that identify products and their attributes from codes (e.g. barcodes or QR codes), reading text or colour of objects are examples of stand-alone scanning applications.

Applications can include scanning functions as a means of inputting information that will be further used by the application.

EXAMPLE 2 Mobile applications to handle business cards recognize the characters and automatically extract personal information, such as name, phone number, address, etc. This information can be used to add new contacts to an existing address book or call the person. In this application, the scanned image of the business card can be read from a file or can be scanned in real-time.

Applications that interoperate with other applications utilize other applications in order to process the recognized object. In some cases, several applications can be interoperated in sequence.

EXAMPLE 3 A scanning application recognizes an object (e.g. a barcode or QR code), sends the name of the object to a web browser, and receives search results from the web browser.

EXAMPLE 4 A scanning application recognizes English text, sends it to a translation software, receives Korean text, sends it to the web browser and receives search results in Korean.

5.6 Scanning modes

There are two main modes of scanning: user-initiated and automatic scanning.

NOTE 1 This distinction is not intended to limit the use of this document. There might be other modes or combinations of modes that could be used.

In user-initiated scanning, the scanning application captures the scanned image when the user initiates the scan. In automatic scanning, the scanning application recognizes the object of interest and automatically captures the scanning image.

The user directs the device to an object and moves the device or object slowly to recognize the object. The automatic scanning mode continues this process until the system automatically recognizes and captures the object, while in the user-initiated scanning mode, it is until the user has captured the object.

EXAMPLE 1 A barcode reader application uses automatic scanning.

EXAMPLE 2 A scanning application, which uses automatic scanning, goes through steps for detecting visual information from a scanned image.

EXAMPLE 3 A smartphone application, which recognizes text, scans a document when users push a button on the touch screen.

NOTE 2 In a scanning application, different scanning modes can be applied for each type of object.

6 Guidance on scanning visual information for presentation as text in various modalities

6.1 General guidance

6.1.1 Provide scanning results according to the purpose

A scanning application shall identify the visual objects of interest to the user.

If the application did not identify any objects, it shall inform the user that objects have not been identified.

NOTE Scanning applications can be used for a variety of purposes. If the purpose of the application differs from the purpose of use, the user cannot obtain the desired result.

EXAMPLE A QR-code scanner cannot recognize the data matrix code.

6.1.2 Provide scanning results equivalent to the object being scanned

A scanning application should provide scanning results that are equivalent to the objects of interest being scanned.

EXAMPLE 1 The barcode scanning result is equivalent to the text contained in the barcode.

EXAMPLE 2 The result of colour recognition is not always the same depending on the surrounding brightness.

NOTE The scanning result is also used as a cue for users to determine whether the scanning is successful or not (see 6.1.9).

EXAMPLE 3 An application reads “nothing found” when it did not identify the visual object of interest.

6.1.3 Provide focus adjustment

A scanning system shall provide a way to focus the visual objects of interest.

NOTE 1 There are many ways to focus on visual objects of interest, including:

— designing the system handheld type so that the user can move to obtain an accurate focus image;

— providing guidance so that the user can move the object to focus;

— using an autofocus function.

NOTE 2 With the autofocus function, it does not require the user to adjust the focus.

6.1.4 Provide exposure adjustment

A scanning system should provide a way to adjust the exposure of the visual objects of interest.

NOTE 1 Automatic exposure control is convenient because it does not need to adjust exposure of visual objects.

EXAMPLE 1 An application scans after adjusting the exposure through a trial scan.

EXAMPLE 2 In a scanning system with manual iris control, the user captures an image of the object while slowly adjusting the aperture.

NOTE 2 When the exposure is low, using an additional light source (e.g. flash) will help to obtain better exposed images.

EXAMPLE 3 If the exposure is insufficient, the scanning application informs the user “low light” or “use flash”.

6.1.5 Provide position adjustment

A scanning system should provide an accessible means for capturing or scanning the image when the visual object of interest is located within the camera angle.

NOTE 1 Providing visual guidance is helpful to a user to locate visual objects of interest in the target area, usually the centre of the image display.

NOTE 2 Visual guidance is often provided as a target area for locating and sizing object images (see 6.1.6).

EXAMPLE 1 A scanning application provides a cross-hair in the centre of the image display.

NOTE 3 Audio or vibration guidance that helps the user to move the object into the target area or recognizable distance can be provided for proper positioning.

EXAMPLE 2 A barcode application provides voice guidance, such as “Barcode is in top left of image frame. Please move your camera to top left for successful scanning.”

NOTE 4 Scanning applications, which indicate the presence of an object of interest, might not need to provide positioning guidance.

EXAMPLE 3 Alarm sounds, such as “ding-dong”, when the object of interest is recognized.

6.1.6 Providing distance adjustment

A scanning system should provide an accessible means to capture the visual object of interest whose resolution is sufficiently enough to recognize the information presented by the object.

NOTE 1 Capturing visual objects of interest within large images can have the effect of eliminating unnecessary backgrounds.

NOTE 2 Capture a sufficiently large visual object of interest can be made possible by:

— designing the handheld system so that the user can easily move closer to the objects of interest;

— providing a pedestal for placing the objects so that the system can capture an appropriate image of the object;

— using zoom lens or software-assisted zoom-in functionality;

— using a high-resolution camera or scanner (and crop unnecessary objects from the image).

NOTE 3 Guidance, which estimates the size of the object appearing in the image, is helpful to capture a larger object image.

EXAMPLE A semi-transparent rectangular guide is engraved on the display of the system to position the visual object of interest.

6.1.7 Provide orientation adjustment

A scanning system should provide an accessible means to adjust the orientation of the scanned object in the image.

NOTE 1 When the orientation of the scanned image is not appropriate, the system can prompt the user to change the orientation of the object.

NOTE 2 Some systems can automatically adjust object orientation.

6.1.8 Provide accessible initiation of scanning

A scanning system shall provide an accessible means to initiate the scan.

NOTE 1 There are several means of scanning initiations, including:

— activation of controls (e.g. pressing a button, clicking an icon);

— voice command;

— gesture input, etc.

NOTE 2 Some scanning applications, which provide an automatic scanning mode, start scanning when the user activates the application or the scanning system using a control, such as a power switch.

EXAMPLE A barcode reader automatically starts scanning as soon as the power switch is turned on.

6.1.9 Provide notification during scanning progress

A scanning application shall inform the user of the scanning progress (e.g. the start and end of scanning input, processing and output).

NOTE 1 If the scanning application provides various output modalities (e.g. visual, audio and sound, tactile), the application would be more accessible to the users.

EXAMPLE 1 A red triangle (►) animation on the display indicates that the scanning is in progress.

EXAMPLE 2 The scanning reads the result as “success” or “failure”.

EXAMPLE 3 If the scanning is successful, a distinct sound-based alarm (e.g. “ding-dong”) notifies the user. If in failure, a beep will sound.

NOTE 2 When the application successfully completes the scanning, it is not necessary to inform the user that the scanning was successful other than providing the scanning results (see 6.4.2).

If the scanning fails, the scanning application shall provide the cause and/or resolution of the failure so that the user can solve the problem.

NOTE 3 There are various causes of the failure, including:

— inappropriate exposure (see 6.1.4);

— out of focus (see 6.1.3);

— insufficient resolution (see 6.2.4);

— improper orientation (see 6.1.7);

— inappropriate size of visual objects of interest (see 6.1.6);

— inappropriate positioning of visual objects of interest (see 6.1.5).

6.1.10 Provide the scanning results in a textual representation

A scanning system shall provide the scanning results to the user in a textual representation.

NOTE 1 The user expects the scanning result to be equivalent to the visual information of interest (see 6.1.2).

EXAMPLE 1 A scanned colour is presented as “red” or “pink”.

NOTE 2 Scanned text can be represented as a text string, including the possible recognition of any recognizable formatting elements (e.g. title, paragraph, indent, bold, italics).

EXAMPLE 2 Output documents formatted according to the scanned text.

6.2 User preference settings

6.2.1 Enabling user preference settings

When scanning applications enable the user to set personal preferences, these settings should be configurable and accessible.

NOTE ISO 9241-171:2008, 8.2 provides detailed guidelines on user preference settings for applications.

6.2.2 Selecting scanning modes

Where multiple scanning modes are available, the scanning application shall allow the user to select the scanning mode currently used.

NOTE Some scanning systems can only be used in automatic scanning mode at all times.

6.2.3 Selecting purpose of use

If a scanning application can be used for various purposes, it shall allow the user to select the purpose of use.

EXAMPLE A user configures an application, which can recognize a QR code and a data matrix code, to recognize QR code.

NOTE Scanning applications designed solely for one purpose (i.e. dedicated systems) usually have no option for users to select the purpose of use.

6.2.4 Selecting image resolution

Where multiple scanning resolutions are available, the scanning application shall allow the user to adjust the scanning resolution currently used.

NOTE 1 A higher resolution image might take longer to get the scanning results, although accurate results can be obtained. If the resolution of the image is low, the processing time might be shortened, but accurate scanning results might not be obtained.

NOTE 2 See 6.3.2 for guidance on capturing an appropriate scanned image.

6.2.5 Selecting image enhancement

Where image enhancement is available, the scanning application shall allow the user to select the use of image enhancement.

NOTE 1 There are various image enhancement methods, including:

— colour correction;

— contrast enhancement;

— image cropping.

NOTE 2 See 6.3.3 for guidance on image enhancement.

6.2.6 Controlling presentation of visual guidance

If a scanning application provides a visual guidance programmatically, it shall allow the user to turn off/on the guidance.

NOTE Providing the visual guidance is helpful to the user to locate visual object of interest in the target area (see 6.1.5) or to capture a larger object image (see 6.1.6).

6.2.7 Selecting post-processing components

Where post-processing components are available, the scanning application shall allow the user to select the post-processing components to be used.

NOTE 1 There are various possible post-processing functions, including:

— translation components;

— search the web;

— keyword search.

NOTE 2 See 6.4.4 for guidance on post-processing.

6.2.8 Selecting modalities for textual representation

Where multiple output modalities are available, the scanning application shall allow the user to select the output modalities to be used.

EXAMPLE Turn visual display on/off, turn TTS on/off, turn vibration on/off, etc.

NOTE 1 Users can select multiple modalities at the same time (e.g. visual and audio).

NOTE 2 See 6.5.2 for guidance on providing applicable multimodalities.

6.2.9 Selecting other application(s) to be invoked

Where other applications are available to be used with the scanning application, the scanning application shall allow the user to select the other applications to be used.

EXAMPLE A web browser is selected to open a website whenever any URI is identified in the scanning.

NOTE See 6.5.3 for guidance on invoking other applications.

6.2.10 Storing and retrieving user preferences

A scanning application should enable users to store and to retrieve user preference settings.

NOTE 1 Storing and retrieving user preference settings reduces the overhead and effort required to create preference settings that are repeated in every device.

NOTE 2 User preference setting can be transferred to other devices with the help of platform software.

EXAMPLE A user loads a preference settings file from a universal serial bus (USB) drive onto a new device.

NOTE 3 In the case of public system, it is convenient to allow each user to set the user preference.

6.3 Scanning input

6.3.1 General

Visual objects are primarily of interest because they contain visual information that can be processed after being scanned.

To generate accurate visual information, it is important for the application to adjust the optics and associated electronics to obtain a good quality image containing the visual objects of interest in the scanning input stage.

6.3.2 Scanning/capturing appropriate images

The scanning input component should obtain an appropriate scanned image.

NOTE 1 An appropriate scanned image satisfies the following conditions:

— appropriate exposure (see 6.1.4);

— in focus (see 6.1.3);

— sufficient resolution (see 6.2.4);

— proper orientation (see 6.1.7);

— appropriate size of visual objects of interest (see 6.1.6);

— appropriate positioning of visual objects of interest (see 6.1.5).

NOTE 2 Providing the reason to the user, when the scan/capture component cannot obtain an appropriate scanned image, allows the user to easily solve the problem (see 6.1.9).

EXAMPLE Some reasons include: “Underexposure”, “Out of focus”, “Insufficient resolution”, “Improper orientation”, “Subject is too small”.

6.3.3 Enhancing image quality

Image enhancement shall not interfere, erase or inflate the visual objects (or information) of interest.

NOTE 1 Removal of a visual object, which is irrelevant to the visual information of interest, does not affect the image enhancement result.

EXAMPLE Some examples that do not modify, erase, or inflate the visual information of interest include:

— border removal of document image to prepare OCR scanning;

— enhancement of images;

— editing images to erase unnecessary areas;

— cropping of unnecessary areas;

— combining into one image by dividing a document image consisting of several columns, in order to prepare OCR scanning.

A scanning application that provides multiple image enhancement component(s) should allow the users to select the components to use.

NOTE 2 Image enhancement components being selected can be kept as setting options (see 6.2.5).

6.3.4 Recording scanned images

Scanned images should be recorded or saved at any point in the scanning and processing of the visual information.

NOTE 1 A scanned image can either be recorded automatically or be saved upon user request.

NOTE 2 Some applications do not leave image files when scanning is finished.

EXAMPLE Some OCR applications convert scanned text images to text, read the text, and then erase the images.

6.3.5 Naming scanned images

If a scanning application records or saves the scanned image, the images shall be recorded with a unique, identifiable name.

NOTE 1 Providing a meaningful name for scanned images is efficient for searching or retrieving the images.

NOTE 2 Some scanning applications generate the name of the scanned images using the date and time of the acquired image (e.g. “DSC20161225”).

NOTE 3 Some scanning applications request the user to enter the name of the scanned image. In this case, it is important that the application makes sure that the name entered is not identical to the name of any image that already exists.

The name of the recorded image should be modifiable by the user.

NOTE 4 Allowing the user to modify the name of a recorded image makes it easy to manage scanned images.

NOTE 5 By using the visual information that the scanned image contains as a name, users can get more information about the image without opening the particular image file.

NOTE 6 Users need an input device (i.e. keyboard) to modify the name of the recorded image.

6.3.6 Retrieving scanned images for processing

If a scanning application explicitly records scanned images, it shall allow users to identify images that are retrieved and fed to the processing component.

6.4 Processing

6.4.1 General

Scanned image consists of various visual objects that are either of primary interest to the user or unnecessary objects. The processing is used to recognize objects of interest from the scanned image and to identify recognizable information within the scanned image to generate appropriate information of interest.

User needs are very broad. For example, some users want to know if a user-defined visual object of interest exists. On the other hand, some users want to know all the visual information contained in the scanned image. Thus, the function of the processing component varies greatly depending on the scanning applications.

6.4.2 Presenting results of processing

The processing component shall be able to identify (all) the visual information required by the users.

NOTE 1 The output of processing components can be in any format (e.g. text string, binary data, numbers, or a combination thereof).

The processing component shall notify the users when there is no visual information of interest (see 6.1.9).

NOTE 2 If the processing component cannot find the visual information of interest, it can repeat the scanning input and the processing procedures subsequently.

6.4.3 Pre-processing: selecting visual objects of interest

Pre-processing shall be able to identify (all) visual objects of interest required by the users.

NOTE 1 Zero, one or more visual objects of interest might be identified.

EXAMPLE 1 The QR code application identifies only one visual object (QR code).

EXAMPLE 2 An OCR application identifies all the characters that make up visual text as visual objects.

The pre-processing component shall notify the users when there is no visual object of interest (see 6.1.9).

NOTE 2 Providing the reason to the user, when the pre-processing component cannot identify any visual objects of interest, assists the user to take appropriate actions for finding objects.

EXAMPLE 3 If pre-processing component cannot identify visual object of interest because of out of focus, it informs the user "out of focus".

6.4.4 Post-processing: respecting the context

Post-processing shall not interfere, erase or inflate the context of the visual information of interest.

EXAMPLE Some examples that the post-processing results are equivalent to the visual information before the processing include:

— translations equivalent to the original (i.e. Korean to English);

— providing original and search results;

— documents formatted with the same text.

A scanning application that provides multiple post-processing components shall allow the users to select the component(s) to use.

NOTE Applicable post-processing components being selected can be kept as setting options (see 6.2.7).

6.5 Output

6.5.1 General

The scanning results of the output component are delivered to users or the output component might invoke other applications, and then it transfers the scanning results to the invoked applications.

6.5.2 Presenting accessible scanning results

A scanning application shall deliver the scanning result to the user in an accessible manner.

EXAMPLE Some examples of accessible text representation are shown below, but are not limited to:

— visual representation (e.g. textual presentation on a display);

— audio representation (e.g. text reading using TTS);

— tactile representation (e.g. textual presentation in the refreshable braille display);

— any combinations of above representations.

A scanning application that provides various accessible representations shall allow the users to select the user preference(s) to apply.

NOTE Applicable modalities of representations being selected can be kept as setting options (see 6.2.8).

6.5.3 Invoking other applications: passing the scanning results

If a scanning application invokes other application(s), it should pass the scanning results as the input of the other applications.

EXAMPLE A QR-code scanner identifies URI that is passed to the web browser being invoked.

If a scanning application provides multiple applications that can be invoked after the process, it shall allow the users to select the application(s) to be invoked.

NOTE Applicable applications to be invoked being selected can be kept as setting items (see 6.2.9).

Types of argument data should be prepared according to the types of arguments of the invoked applications.

6.5.4 Privacy protection and security

A scanning application should provide appropriate levels of privacy protection and security. The user should be able to control the levels of privacy and security.

EXAMPLE If a scanning application reads the scanning results aloud, people nearby can hear it.

Annex A
(informative)

Scanning visual information for presentation as text in various modalities — Use cases

A.1 Recognizing barcodes

A.1.1 Description

Scott who is visually impaired is lactose-intolerant and only consumes milk without lactose.

One day, he went to the grocery store to buy milk. He went to the dairy section to find milk products that he always buys. Since Scott cannot read the printed text on the milk cartons, he relies on a barcode scanning application on his smartphone to identify the products. Grabbing his smartphone in his right hand and a milk carton in his left hand, he brought his smartphone camera closer to the surface of the carton.

As soon as the application recognized the barcode, the smartphone alarmed “ding-dong” and audibly informed him of the product. After repeating this process several times, Scott was able to pick out the product he wanted.

A.1.2 Similar use cases

Similar use cases include:

— scanning applications that recognize barcodes, such as QR code, data matrix code, and UPC of the printed materials, of the surface of the products;

— scanning applications that identify the graphic signs, such as traffic signs of the road, toilet signs on the wall;

— scanning applications that recognize and read the text content of the 2D barcode provided for those who cannot read the printed material.

A.2 Recognizing short text

A.2.1 Description

Fred who is visually impaired has recently installed the optical character recognition (OCR) application on his smartphone. He often uses this application to distinguish banknotes when paying small amounts in cash. First, Fred selects and activates the scanning application on his smartphone, and he places the smartphone camera near the corner of the banknote that he has picked. As soon as the smartphone recognizes the number of the banknote, it reads the value of the banknote.

A.2.2 Similar use cases

Similar use cases include:

— scanning applications that recognize short text of the printed materials, such as daily goods, bills, flyers, greeting cards, etc.;

— scanning applications that scan text of an object (e.g. signboard, signpost, road sign, licence plate, etc.), recognize text on the object, and translate recognized text into other languages.

A.3 Recognizing text on prints

A.3.1 Description

Erin who is 9 years old has dyslexia. Recently, Erin's parents bought a PC scanner and an OCR software to help Erin's reading activities. Nowadays, when Erin wants to read a book, she turns on the PC, activates the OCR application, and places the book on the scanner attached to the PC. When she presses the “Scan” button of the scanner, her PC starts scanning the page of the book, recognizes the text content of the page, which then reads out the content.

Erin’s mother, Rachel, who is visually impaired, often uses this application as well. Today, she had to read the flyer that Erin brought from her school. First, Rachel connected the refreshable braille display to the PC and pressed the “Scan” button of the scanner. Soon the scanner started scanning the flyer and the OCR program output the content of the flyer to the braille device so that Rachel could find out what the flyer was all about.

A.3.2 Similar use case

A similar use case involves OCR applications that use scanners to obtain text content from a variety of text-rich printed publications, such as books, documents, newspapers, magazines, etc.

A.4 Identifying object colour

A.4.1 Description

Bryan who is visually impaired is going to wear a red T-shirt for a party at the office. He activates the colour scanning application in his smartphone and takes out his T-shirts in the closet one by one, bringing the smartphone camera closer to the T-shirt. The application then notifies Brian of the T-shirt colour with a beep sound. Using this application, Brian can find the red T-shirt to wear today.

Bryan often uses this application at work. He keeps documents of the same type in a colour-coded folder. For example, documents related to procurement are stored in a blue folder, while documents related to general affairs are stored in a red folder.

This morning, Brian had to look for a document on procurement of PCs that his boss requested in the blue folder. He activated the colour application in his smartphone and placed the smartphone camera to all the folders one by one.

Soon he was able to find the blue folder with the aid of the mobile colour application. Using the OCR scanner, he could find the document his boss has requested.

A.4.2 Similar use case

A similar use case involves colour scanning applications that notify the colour of the objects such as paper sheets, cloths, etc.

A.5 Identifying a specific object

A.5.1 Description

Jane who has a low-vision often uses a mobile application that recognizes objects based on the shape of things. The object-recognition application has dual modes. The first one is to count the number of the objects she wants to know. The second one is to find out where the desired object is located.

One day, Jane invited five friends home. She asked her son, Eric, an elementary pupil, to set the table as she talked. Eric put five sets of napkins, forks, and spoons on the table.

After finishing the snack preparation, Jane set the object-recognition application on her smartphone to recognize "fork". As she pointed the smartphone camera towards the table, the application audibly mentioned, "Found five forks" along with a "ding-dong" sound, thus notifying her of the placement of utensils.

Jane also uses the object-recognition application when visiting a place for the first time. After activation of the application, she sets the application mode and the target object as "detection mode" and "doorway", respectively.

As Jane pointed her smartphone camera hand forward and slowly turned her arm to left, when the application recognized the doorway, the smartphone vibrated and she could find the way to the doorway.

A.5.2 Similar use case

A similar use case involves scanning applications that notify the user of a landmark (e.g. crosswalk, overpass, traffic light), facilities (e.g. vending machines, ATM, kiosks), etc.

A.6 Recognizing a shape and number

A.6.1 Description

Tim who is visually impaired is often helped by a wearable system that recognizes the shapes of objects along with written texts.

One day, he had to take the No. 50 bus to go downtown. After activating the bus application recognizing the shape and number of the bus, he was able to identify his bus, while waiting at the bus stop. Whenever he heard a bus come in, he turned his head toward the bus in which his wearable scanning system recognized the bus numbers with a beep sound. As soon as he heard "Bus, route 50", he raised his hand to hail the bus.

A.6.2 Similar use cases

Similar use cases include:

— scanning applications that recognize objects of multiple attributes such as route buses, sign boards, traffic lights, etc.;

— automatic payroll system of a parking lot wherein the system scans and recognizes the licence plate of the vehicle that comes into the parking lot.

A.7 Recognizing landmarks

A.7.1 Description

32-year-old Susie likes to travel. When she plans to travel abroad, she used to rely on travel books to find scenic landmarks to visit. Recently, she installed a mobile tour-guide application for her destination. When Susie arrived at her destination, she activated the mobile tour-guide application and set the "search nearby" mode on.

After then, she held the smartphone in hand with the camera facing forward and turned her body to the right slowly. When the application recognized a landmark, it showed detailed information about the site.

A.7.2 Similar use case

A similar use case involves scanning applications that use scanned images and GPS to provide information about buildings or attractions around user's location.