INTERNATIONAL STANDARD

ISO/IEC 30113-1

First edition
2015-04-15

Information technology — User interface — Gesture-based interfaces across devices and methods —

Part 1: Framework

Technologies de l'information — Interface utilisateur — Interfaces fondés sur la gestuelle entre dispositifs et méthodes —

Partie 1: Cadre

Reference number
ISO/IEC 30113-1:2015(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

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 document 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 and IEC 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 meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT), see the following URL: Foreword — Supplementary information.

The committee responsible for this document is ISO/TC JTC 1, *Information technology*, Subcommittee SC 35, *User interfaces*.

ISO/IEC 30113 consists of the following parts, under the general title *Information technology — User interfaces — Gesture-based interfaces across devices and methods*:

— *Part 1: Framework*

— *Part 11: Single-point gestures for common system actions*

Introduction

Gestures are used for performing a variety of commands (such as scrolling a Web page up) as an alternative input method (to typing or using a mouse to select objects).

Given the limited number of basic gestures, the same gesture is often used for a variety of different commands in different situations. It is important that wherever possible, these different commands are similar to one another (i.e. by having a similar effect on different objects) so that users are not confused about what a gesture will do in a given situation.

Standardized gesture descriptions and commands minimize user confusion when interacting with various software systems and applications on various ICT devices. This International Standard is aimed at designers and developers of software applications.

This International Standard is intended to help users to more easily navigate and control application software on various ICT devices by standardizing gestures and gesture commands.

This part of ISO/IEC 30113 defines a framework of gesture-based interfaces to support interoperability among gesture-based interfaces with various input devices and methods.

Subclause A.1 gives informative description about the structure of ISO/IEC 30113 in detail.

Information technology — User interface — Gesture-based interfaces across devices and methods —

Part 1: Framework

1 Scope

This part of ISO/IEC 30113 defines a framework and guidelines for gesture-based interfaces across devices and methods in supporting interoperability.

NOTE Some of these devices include mice, touch screens, touch pads, 3D mice, joysticks, game controllers, wired gloves, depth-aware cameras, stereo cameras, Web cameras.

This part of ISO/IEC 30113 does not define or require specific technology for recognizing gesture of users. It focuses on the description of a gesture and its functions for utilizing ICT systems.

NOTE Operation of a physical keyboard is not addressed in this part of ISO/IEC 30113.

2 Conformance

A gesture-based interface is conformant to this part of ISO/IEC 30113 if it meets all requirements of Clause 5.

3 Terms and definitions

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

3.1
gesture
movement or posture of the whole body or parts of the body

3.2
gesture-based interface
gesture interface
user interface that provides information and controls for a user to accomplish specific tasks with the interactive system by his/her gestures

[SOURCE: ISO 9241-171: 3.29]

3.3
gesture command
instruction to the system resulting from a gesture input by the user, e.g. select, move, delete

[SOURCE: ISO/IEC 14754:1999, 4.5]

3.4
gesture software
software for implementing gesture-based interface functionality including gesture recognition, command processing, and feedback generation

Note 1 to entry: Gesture recognition software is usually contained within the operating system and specific device drivers. Information on gestures that are recognized is made available to the operating system and/or the application software, so that the intended command(s) are performed in response to the gesture.

4 Overview of gesture-based interface

4.1 General

Users can use gestures to interact with interface objects. Interface objects have representational properties (e.g. how are they rendered to user) and operational properties (e.g. what do they do) that can be effected by gestures.

Human-machine interaction involves a loop of execution and evaluation. A machine offers feed forward and a user manipulates interface objects (execution). The machine displays feedbacks and new feed forward (evaluation) and the user adjusts manipulation, and so on. The user produces gestures and the machine understands them based on the properties of the gestures that it recognizes.

For a successful interaction, the machine needs an input device in order to collect gesture properties. Those properties will be analysed by gesture software to compare those properties to pre-defined gesture command properties, and then decide to operate associated functions.

Figure 1 illustrates a model of human-machine interaction based on a gesture-based interface. It presents a schematic diagram of relationships among the user, gesture command, input device and machine (ICT system) when the user utilizes a gesture-based interface during human-machine interaction. The gesture-based interface includes hardware (physical) and software (logical) components. The input device is the hardware which recognizes the gesture and sends its associated input signal to the ICT system. The gesture software finds a command which is pre-defined and mapped to the input signal. The application software generates its feedback to the user using the output device.

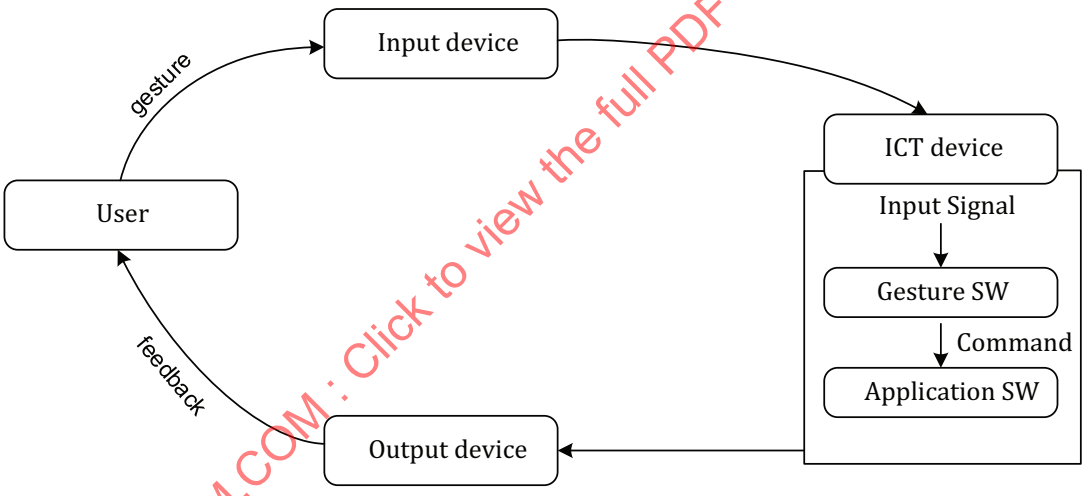

Figure 1 — Loop of human-machine interaction with a gesture-based interface

4.2 User's actions for gesture input

A user generates actions for gesture inputs which are two-dimensional motions relative to its supporting surface, two-dimensional or three-dimensional finger/hand/body postures/motions in a space, postures/motions of fingers on a surface and so on. A gesture can also be generated by a tool, as an extension of the body (such as: a wand, a pen, a mouse, a remote control or a glove).

Some gestures are controlled by a discrete body part such as one finger, several fingers, hand movement or fingers associated to hand movement. Facial expression, eye gaze and eyelid blinking can also provide a user's action for gesture input. Other gestures might be generated with a whole body or a coordination of several body parts coordination. They could involve arms, hand and fingers, and their coordination.

Physiological constraints which apply to gesture generation are important to take into account before defining gestures. For example, some gestures are difficult to be produced with a mouse in the hand on a 2D surface, however, easy to be produced with a finger on a 2D surface.

All gestures involve a clear and identifiable start, one or more action(s) and a clear and identifiable end (as further discussed in A.3.4.2). Before performing a gesture, the user can initiate a gesture recognition (where required to do so) by doing some action such as holding down a specific button on a device. Gesture recognition might be automatically supported by the system without the need of any action beyond the start of the gesture. The user generates a specific gesture (such as drawing 'L by moving a mouse') by motions between a start and end state. The user ends the gesture input by arriving at some state that is recognized by the system as indicating the end of the gesture. This end state might be included within the gesture or might be presented with another input modality (such as a voice command).

The gesture generated by the user is then interpreted as a command in by the operating system or a specific software application when the ICT system recognizing the gesture correctly.

4.3 Gesture input device

A gesture input device receives the interactions provided by a user and generates input signals to be interpreted by the gesture software. Example of useful gesture input devices include mice, touch screens, touch pads, 3D mice, joysticks, game controllers, wired gloves, depth-aware cameras, stereo cameras, Web cameras and so on.

4.4 ICT system

Gesture software analyses the signals received from gesture input devices. The functions of the gesture software include gesture recognition, command assignment and gesture feedback.

The gesture software recognizes pre-defined gestures from actions exercised by a user with a gesture input device. Then the gesture software sends the associated gesture command to application software.

While the user generates a gesture, the gesture software might invoke a feedback signal via the ICT system to the user. The feedback helps the user to notice whether the gesture command is properly activated or not. The feedback might be rendered using sound, visual display and/or tactile display.

4.5 Cultural Adaptability

Since gestures are one of input mechanisms such as a keyboard and a voice command, they are subject to internationalization/localization. Some gestures might be culturally dependent.

EXAMPLE Bulgarians nod to say "no" and shake their head for "yes", while Americans nod to say "yes" and shake their head for "no".

4.6 Accessibility

Due to the complexity, some gestures might not be properly and/or completely exercised by users with disabilities and/or elderly users. When gestures are defined for an ICT system, consideration of accessibility for all users (including the disabled and/or the elderly) is important.

5 Requirements and recommendations

5.1 Activating/finishing a gesture

A gesture-based interface shall provide one (or more) method(s) for activating and finishing a gesture.

EXAMPLE A mouse with two buttons is used as a gesture input device and holding down the secondary button of the mouse activates a gesture. By releasing the button, the user's action for gesture input is finished.

These methods may be managed by the user or automatically managed by the system.

5.2 Performing a gesture

A gesture-based interface shall provide one (or more) method(s) for making a gesture.

NOTE Making a gesture is valid only when the system is actively receiving gesture input. In some systems, the receipt of gesture input by a system can be activated and deactivated by the user.

EXAMPLE In a specific mouse gesture, the method for gesture formation is to move the mouse horizontally or vertically within one stroke.

5.3 Feedback for confirming a gesture

A gesture-based interface should provide one (or more) feedback signal(s) to notify the user of the current state of performing the gesture.

NOTE 1 Feedback can represent several states such as interface object selection, interface object activation, interface object manipulation, gesture command initialisation state, gesture command performing state, gesture command ending state and feedback about function execution.

Feedback should be expressed through one or more of the visual, tactile or audible modalities.

NOTE 2 When focus indicates an object is selected, the gesture command will apply to that specific object.

EXAMPLE 1 A visual trail line showing the movement of the pointer (mouse pointer) is displayed on a screen when a gesture is performed.

EXAMPLE 2 An ICT system makes a sound as a signal announcing that the gesture command is recognized.

EXAMPLE 3 Changes to an object's state are displayed after they are made by a gesture.

5.4 Feed forward

A gesture-based interface should provide clear feed forward signals to notify the user what kind of gestures are done and when they are done.

NOTE As gestures are dynamic, dynamic feed forward is more effective.

EXAMPLE A visual clue helps the user to identify that a certain interface object can respond to some gestural shortcuts.

5.5 Cancelling a gesture

The gesture-based interface should provide at least one cancelation method that can be used during the input of a gesture.

EXAMPLE If gestural input exceeds a specified time limit, gesture command is cancelled.

5.6 Criteria of gesture size

To minimize misunderstanding of gesture input, the gesture-based interface should provide criteria for ignoring very small or large movements when a gesture is formed.

EXAMPLE Within a particular system, the minimum movement of a pointer for gesture input is greater than 30 pixels along horizontal, vertical or diagonal direction.

5.7 Controlling the criteria

To accommodate individual capabilities, the gesture-based interface should allow a user to modify the criteria for ignoring very small and very large movements when a gesture is formed.

5.8 Changing correspondence of a gesture to a gesture command

A gesture-based interface should provide a method for users to easily change the correspondence of a gesture to a gesture command.

NOTE This allows users to create simplified gesture sets for their individual use.

5.9 Descriptions of individual gestures within the part

Descriptions of individual gestures within the ISO/IEC 30113 series should utilize the format given in A.3.

Annex A
(informative)

Outline for describing the ISO/IEC 30113 series

This annex provides a framework for parts of the ISO/IEC 30113 series and a structure for describing individual gestures within each of the parts.

A.1 Framework for organizing gesture standards

It is important that the categories used to organize individual gestures within the ISO/IEC 30113 series are relatively stable and are used to aid both standards developers and users of standards in finding the appropriate parts in this series.

The gesture standards in the ISO/IEC 30113 series will be organized based on their application and the complexity of the human interaction components involved in making the gesture in order to provide clear guidance to standards developers and users of standards where a particular gesture belongs.

Application areas are organized from the most general to the most specific in terms of applicability. Five application areas identified for separate sets of parts of this standard:

a) gestures that are used for common system functions will be standardized within the 1x series of parts (e.g. part 11, part 12);

b) gestures that are used for device control will be standardized within the 2x series of parts;

c) gestures that are used for document and image viewing will be standardized within the 3x series of parts;

d) gestures that are used for music and video playing will be standardized within the 4x series of parts;

e) gestures that are used for Web navigation will be standardized within the 5x series of parts.

Further series of parts might be added to this framework in the future, if needed.

Within each series, the parts are further organized in terms of complexity of the human interaction components from simplest to most complex. This complexity is based on the number of points of interest involved in producing/recognizing the gesture, the part of the body in producing/recognizing the gesture, and/or technologies involved in producing/recognizing the gesture.

The concept of a point of interest is essential to the specification, production and recognition of a gesture. The points of interest involved in a gesture are the minimum number of independent points that are needed to fully recognize a gesture and to differentiate it from other gestures, when the movement of these points is considered.

It is important for users which part of their body is used to specify and produce a gesture. The concept of the body part for the gesture would be critical in making the gesture-based interactions more natural and usable. At the same time, it might be considered that some users might not be able to use their body parts for generating the gesture.

The technologies for producing/recognizing gestures are also essential in determining their complexity. The gestures can be generated using a device such as a mouse, a pen, a wand, a touch pad, and so on. Sometimes the gestures are produced in the air using body parts. They are dynamic and/or static. Though the basic concept of the gestures is same, some gestures are affected by the technologies.

While many gestures of interest are distinguished using only one point of interest, there is a need to consider more complex gestures that involve more points of interest:

a) gestures that involve a single point of interest (e.g. gestures involving a pen, stylus, mouse, finger, or other pointing device) will be standardized within the x1 series of parts (e.g. part 11, part 21, part 31, part 41, part 51);

b) gestures that involve multiple points of interest (e.g. gestures involving the set of fingers, whole hands, body postures, facial expressions, or other combinations of multiple points of interest) will be standardized in the x2 series.

NOTE 1 The complexity of a gesture is also of particular significance in terms of how accessible its use is.

NOTE 2 Gestures can involve the positioning and/or movement of these points of interest in two-dimensional or three-dimensional space. Given that most gestures involve two-dimensional space, those gestures that involve three-dimensional space will be included in the same part of the ISO/IEC 30113 series with two-dimensional gestures based on their application area and number of points of interest.

NOTE 3 Individual gestures will be located in the lowest numbered part to which they might apply.

This structure of the proposed parts of ISO/IEC 30113 that will focus on standardizing specific gestures is illustrated in Table A.1. It is recognized that the completion of these anticipated 10 parts will take considerable time, and it is the intent of ISO/IEC JTC 1/SC 35 to develop more general parts before more specific parts as much as is practical.

Table A.1 — Proposed organization of the parts of ISO/IEC 30013 standardizing specific gestures

	x1: Single point	x2: Multi-point
1n Common system actions	11: Single-point gestures for common system action	12: Multi-point gestures for common system action
2n Device control	21: Single-point gestures for device control	22: Multi-point gestures for device control
3n Document and image viewing	31: Single-point gestures for document and image viewing	32: Multi-point gestures for document and image viewing
4n Music and Video playing	41: Single-point gestures for music and video playing	42: Multi-point gestures for music and video playing
5n Web navigation	51: Single-point gestures for Web navigation	52: Multi-point gestures for Web navigation

A.2 Component of individual gesture standards

Each part of ISO/IEC 30113 that standardizes individual gestures will include:

a) a discussion of the application area that is covered by the gestures in the part;

b) a discussion of the complexity that are involved in the gestures in the part. This discussion might include examples of typical body parts/devices that are used to produce the gesture and examples of typical technologies that could recognize the gesture;

c) individual gestures will be identified in terms of the movements and positioning of their points of interest based on their complexity. It is recognized that the same movements and points of interest when applied to different objects will have slightly different outcomes, but it is preferable to consider these as variations on a general gesture rather than as separate gestures.

A.3 Descriptions of individual gestures within the part

A.3.1 Properties of gestures

Gestures are one of a number of alternate methods for users to request that systems perform actions. Other methods include the use of command lines and mouse actions. It is important that these different methods (and their descriptions) be as compatible as possible to support interoperability between different methods of requesting the same action.

Gestures within the ISO/IEC 30113 series will be described according to ISO/IEC/TR 11580 and will take account of ISO/IEC DTS 11581-41.2.

According to ISO/IEC/TR 11580, actions have the following properties:

— internal properties which provide a basis for defining the gesture and distinguishing between different types and instances of gestures;

— information properties which provide textual descriptions of the gesture;

— representation properties provide information on the physical points of interest and their actions that are involved in a gesture;

— operational properties describe the intended functionalities of the gesture and variations of it depending on any objects that it involves.

A.3.2 Internal properties of gestures

A.3.2.1 Unique (internal) identifier

Each gesture will be assigned a unique internal identifier. This internal identifier might be used within the set of standards and/or within application programming interfaces to identify the particular gesture.

The format of these identifiers will be: GN-nnnn, where “G” indicates that it is a gesture, N is the part number of this standard series, and nnnn identifies the particular gesture within this set of gestures.

A.3.2.2 Internal states

Each gesture will have a number of states that include activation and completion of the gestures along with any intermediate states that are required in making/recognizing the gesture (gesture formation). The sequence of states might be discussed within the text description (see A.3.4.2). However, given the dynamic nature of gestures, there is no need for a system to present current information on the state of a gesture either as an internal property or as a separate information property.

A.3.3 Information properties of gestures

A.3.3.1 Entire gesture

Information properties of gestures deal with the entire gesture.

A.3.3.2 Text name of the gesture

Each gesture will be assigned a unique meaningful text name. This text name can be used by people to refer to the particular gesture.

A.3.3.3 Text description of the gesture

Each gesture will be assigned a text description that provides information on the purpose and/or use of the gesture. This text description could be used as the basis of help information provided by the system to users.

A.3.3.4 Graphic representation of the gesture

Each gesture will be assigned a graphic that illustrates the complete gesture. Further detailed illustrations of individual states of the gesture can be provided in the discussion of their representation properties of their states (see A.3.4.2).

EXAMPLE 1 The graphic representation of a gesture involving a single point of interest only the moving left of could be:

EXAMPLE 2 The graphic representation of a gesture involving a single point of interest first moving left and then moving right could be (even though this representation does not show perfectly horizontal movements):

NOTE The exact movements will be described as individual states of the gesture (see A.3.4.2).

A.3.4 Representation properties of gestures

A.3.4.1 Components of a gesture

Representation properties of gestures deal with parts of the gesture.

All gestures involve the transition from an initial state to a final state. Gestures can also include a number of intermediate states. Each gesture will be described in terms of the states of one or more points of interest with relation to one or more objects of interest.

All gestures will have some object(s) of interest in the user interface with which they interact, whether the interaction involves pointing at that object or not.

— Many gestures involve pointing at an object directly.

EXAMPLE 1 Some objects include: a control, a page of an e-Book, a Web page.

— Many gestures have similar actions on a variety of different objects.

— Some gestures, including those that do not point at an object, present information to some input location (object) within the system.

EXAMPLE 2 Sign language involves gestures that are intended as system inputs.

NOTE It is not the intent of ISO/IEC 30113 to standardize gestures used in sign languages. It is recognized that there are many different standard sign languages as there are many different natural languages and that such matters are beyond the scope of this series of standards. However, this does not preclude the adoption of particular gestures that are widely used in one or more sign languages to represent concepts (as opposed to signing used for text entry).

A.3.4.2 States of a gesture

A.3.4.2.1 Number of states involved in the gesture

It is important to know the number of states that are involved making/recognizing a gesture.

A.3.4.2.2 Initial state of a gesture

The start of the gesture usually involves some motion(s) being recognized and might involve some object(s). Thus the initial state of a gesture involves some change that is recognized by the system as initiating the gesture.

EXAMPLE 1 A pointer that is positioned over an object starts to move, indicating the start of a gesture.

EXAMPLE 2 Two pointers start to move together over some object.

EXAMPLE 3 A user starts gesturing with multiple points of interest when the system is waiting for text input.

A.3.4.2.3 Intermediate states of a gesture

Important intermediate states (positions/changes of direction/changes of speed for one or more points of interest) between the start and the end of the gesture involve changes that are significant in identifying the gesture or differentiating it from other gestures.

There are 0, 1, or more intermediate states possible between the start state and the end state.

A.3.4.2.4 Final state of a gesture

The final state of a gesture typically involves the lack of a further movement for some minimum period of time. In place of a movement will be some recognition of one (or more) condition(s) that signal the completion of the gesture.

EXAMPLE 1 A pointer that is positioned over an object stops moving for at least 1 s, indicating the end of the gesture.

EXAMPLE 2 A pointer that was in contact with a surface is lifted off that surface.

It is important that the end point of a gesture be clearly distinctive from any intermediate state.

A.3.4.2.5 Termination of a gesture without completion

It is not necessary to explicitly discuss how a gesture might be terminated prior to its completion. Failure to achieve the final state of a gesture (including adding any recognizable additional states) automatically creates something other than the individual gesture being described. This might result in the recognition of some other gesture (with some of the same states) or in the recognition that no recognizable gesture has occurred. While this is important, it is beyond the ability to be fully discussed within the description of an individual gesture.

A.3.4.2.6 Identifying states of a gesture

Individual states of a gesture can be identified in terms of:

- The **order** in which that state occurs within the gesture.
  - This is a number starting with 1 for the initial state and incrementing by 1 for each additional state.
- The **starting position** of the point(s) of interest.
  - This is typically represented in terms of the object the gesture interacts with, some other object in the user interface, and/or other points of interest.
- Any movement(s) of the point(s) of interest or completion condition(s).
  - There is typically only one movement per point of interest per state.
  - This movement can involve:
    - a direction of movement;
    - (optionally) a distance of the movement;
    - (optionally) a speed or time limit of the movement.
  - It is important to describe the permitted variations in the movement (or completion condition) over which the movement and positioning of the points of interest (or other completion conditions) are recognized as belonging to the same gesture.

