INTERNATIONAL STANDARD

ISO 22311

First edition
2012-11-15

Societal security — Video-surveillance — Export interoperability

Sécurité sociétale — Videosurveillance — Interopérabilité de l'export

Reference number
ISO 22311:2012(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) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

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.

ISO 22311 was prepared by Technical Committee ISO/TC 223, *Societal security*.

Introduction

Video-surveillance is a crucial asset in intelligence collection, crime prevention, crisis management, forensic applications etc. The minimum requirement in societal security is for the authorities to be able to rapidly use the data collected by different CCTV systems from given locations.

This International Standard provides an export interoperability profile which constitutes the exchange format and minimum technical requirements that ensure that the digital video-surveillance contents exported are compatible with the replay systems, establish an appropriate level of quality and contain all the context information (metadata) necessary for their processing.

It is crucial for societal security that present and future video-surveillance systems implement this interface to allow efficient forensic processing of the material produced, often in massive quantities.

This International Standard also contains provisions to ensure that privacy measures can be implemented to protect the rights of the individuals.

This International Standard does not impose implementation methods or technological solutions. It relies heavily on individual technical standards separately developed and concentrates on minimum necessary profiles or subsets thereof to achieve its societal security objectives.

This International Standard is a blend of profiles of standards and practices, which combined, will achieve a minimum level of interoperability.

This implementation has only been possible because of standards produced by the following bodies:

— ISO/IEC JTC 1/SC 29/WG 11, Coding of moving pictures and audio (MPEG);

— ISO/IEC JTC 1/SC 29/WG 1, Coding of still pictures (JPEG);

— IEC/TC 79, Alarm systems and electronic security (including its European equivalent CENELEC/TC 79, Alarm systems and electronic security);

— ITU, International Telecommunication Union;

— IETF, Internet Engineering Task Force;

— SMPTE, Society of Motion Picture and Television Engineers;

— NATO, Standardization Agency.

The normative Annex A contains a metadata dictionary.

The importance of having images stored and presented to the user in such a way that their use is facilitated is presented in the informative Annex B.

Societal security — Video-surveillance — Export interoperability

1 Scope

This International Standard is mainly for societal security purposes and specifies a common output file format that can be extracted from the video-surveillance contents collection systems (stand alone machines or large scale systems) by an exchangeable data storage media or through a network to allow end-users to access digital video-surveillance contents and perform their necessary processing. The means of exchange are not part of this International Standard.

This common output file format relies on a combination of several technical standards that individually are not restrictive enough to provide the requested interoperability. These standards are formally referenced to avoid duplications or divergence. When appropriate to improve the interoperability, subsets or a limited number only of these standards are called.

Since video-surveillance recording often includes taking records of citizens, requirements relating to privacy, use of the records and their disposal are also considered.

Based on the above mentioned technical standards, the following format components are covered:

- Video;
- Audio;
- Metadata:
  - Descriptive (location, camera identifier, etc.)
  - Dynamic (date, time, pan, tilt, zoom, identification results, etc.)
- Encapsulation/packaging for the output file;
- Data/access security and integrity;
- Provisions for privacy;
- Informative data regarding the presentation to users.

2 Normative references

The following referenced documents are indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 10918-1:1994, *Information technology — Digital compression and coding of continuous-tone still images: Requirements and guidelines — Part 1*

ISO/IEC 10918-5, *Information technology — Digital compression and coding of continuous-tone still images: JPEG File Interchange Format (JFIF) — Part 5*[1)]

ISO/IEC 14496-2:2004, *Information technology — Coding of audio-visual objects — Part 2: Visual*

ISO/IEC 14496-3:2009, *Information technology — Coding of audio-visual objects — Part 3: Audio*

1) To be published.

ISO/IEC 14496-10:2012, *Information technology — Coding of audio-visual objects — Part 10: Advanced Video Coding*

ISO/IEC 14496-12:2012, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format*

ISO/IEC 14496-14:2003, *Information technology — Coding of audio-visual objects — Part 14: MP4 file format*

ISO/IEC 14496-15:2010, *Information technology — Coding of audio-visual objects — Part 15: Advanced Video Coding (AVC) file format*

ISO/IEC 15444-1:2004, *Information technology — JPEG 2000 image coding system: Core coding system — Part 1*

ISO/IEC 23000-10, *Information technology — Multimedia application format (MPEG-A) — Part 10: Surveillance application format*[2)]

IEC 62676-1-1, *Video surveillance systems for use in security applications — Part 1-1: Video system requirements*[3)]

IEC 62676-2-3, *Video surveillance systems for use in security applications — Part 2-3: Video transmission protocols — IP interoperability implementation based on web services*[4)]

ITU-T/Rec G.711, *Pulse code modulation (PCM) of voice frequencies*

SMPTE RP210.11-2008, *Metadata Dictionary Contents*

SMPTE 335M-2001, *Metadata Dictionary Structure*

SMPTE 336M-2007, *Data Encoding Protocol Using Key-Length Value*

3 Terms and definitions

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

3.1
video-surveillance
surveillance by video means

3.2
forensic
related to or used in courts of law

NOTE This applies to video-surveillance used to produce legal evidence.

3.3
metadata
information to describe audiovisual content and data essence in a format defined by ISO or any other authority

EXAMPLE Time and date, text strings, location identifying data, audio and any other associated, linked or processed information.

3.4
static metadata
data associated with a digital image aside from the pixel values that does not change over time (or at least does not change over the addressed sequence)

2) To be published.

3) To be published.

4) To be published.

3.5
dynamic metadata
data associated with a digital image aside from the pixel values, which can change for each frame of a video sequence

3.6
CCTV system
surveillance system comprised of cameras, recorders, interconnections and displays that are used to monitor activities in a store, a company or more generally a specific infrastructure and/or a public place

3.7
logical structure
arrangement of data to optimize their access or processing by given user (human or machine)

3.8
geo-location
specific location defined by one of several means to represent latitude, longitude, elevation above sea level, and coordinate system

NOTE Geo-location generally means the meaningful specification of the position of a point or object on the earth. The term itself does not carry a prescription of the coordinate system to be used. Additional attributes associated with a geo-location are not a part of a geo-location specification.

3.9
scene location
collection of geo-locations that defines the perimeter of the viewable scene of a camera

NOTE The coordinate system is the same for each geo-location in the collection. There is at least one geo-location in the scene location. The geo-locations are ordered in either clockwise or counter clockwise order. Single geo-location scenes interpret the geo-location as the centre of the scene.

4 General

4.1 Concept

4.1.1 Video-surveillance systems generic architecture

A CCTV system usually consists of hardware, software and human elements.

A CCTV system for security applications presented as functional blocks, which portray the various parts and functions of the system, as well as the interactions with the human stakeholders is outlined in Figure 1.

This International Standard specifies the packaging and format of the data exchanged between the video and system management functional blocks of a CCTV system and societal security end-users external to the system as covered by the dotted zone of Figure 1.

Informative data regarding the presentation to users are given in Annex B.

The interactions between the components of the different functional modules of a CCTV system are specified by the IEC 62676 series of normative documents. This International Standard is an implementation of the provisions of 6.1.3 in IEC 62676-1-1 (to be published), which calls for publication of all necessary information allowing intended usage of the data produced. Accordingly this International Standard introduces the necessary extra requirements to the IEC 62676 series, without being in contradiction with them.

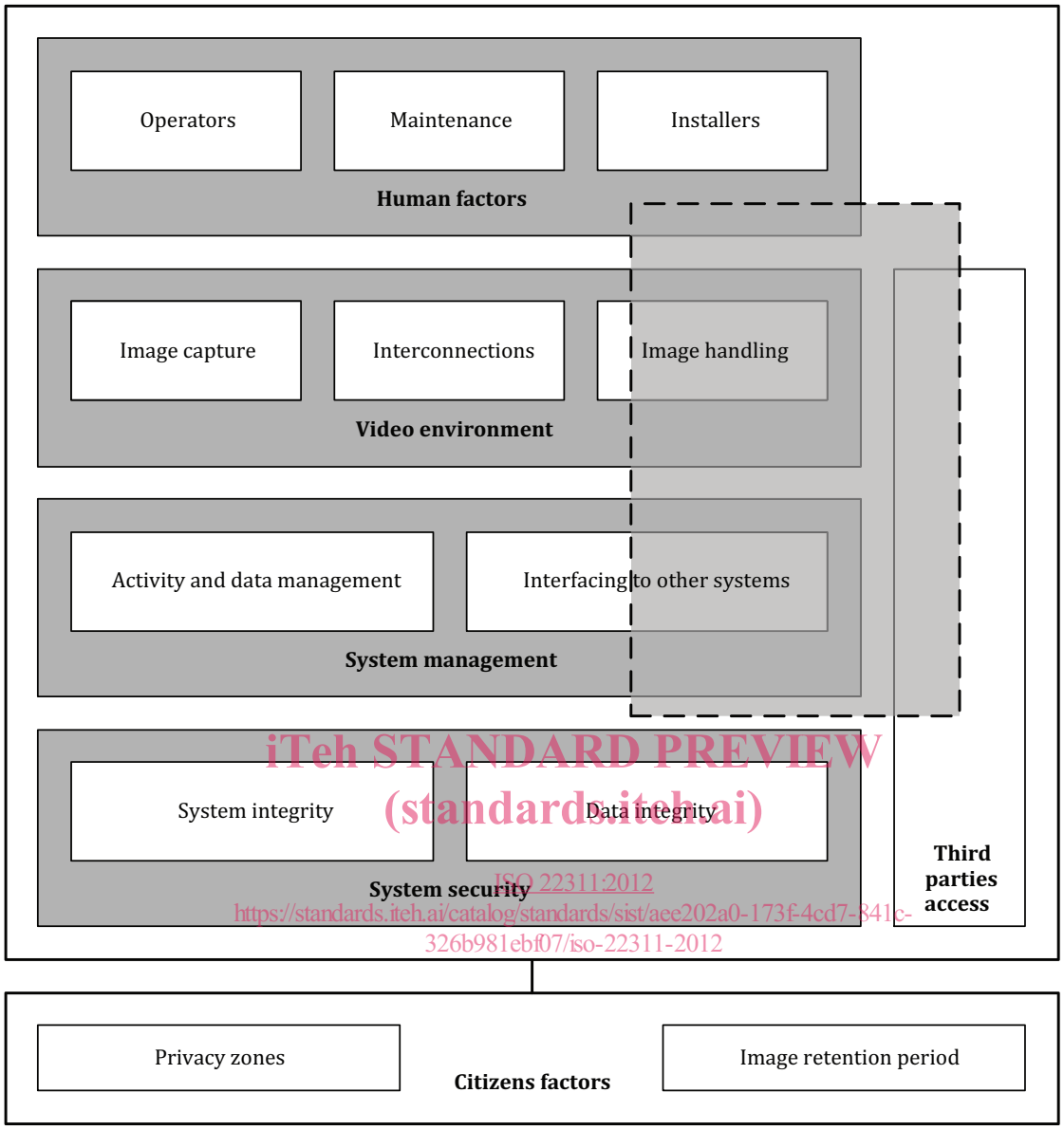

Figure 1 — Functional blocks of a CCTV system for security applications

4.1.2 Minimum requirements for interoperability

Societal security supposes interoperability of digital video-surveillance systems; this International Standard defines the minimum requirements applicable to the file formats used to export the collected contents (video, audio and associated metadata) to achieve this interoperability[5)].

All collected information shall be referenced to Coordinated Universal Time (UTC).

This format shall allow the file export of time slices of data coming from a selection of sources and preserve the time correlation between the contents, whatever export process (removable media or data transmission) is used.

The format shall be such that compatible, comparable processing of files exported by different systems (covering a same scene) with a common time base will be possible.

5) Pending legal authorization, this International Standard may allow streaming contents formatted as above, from a compliant CCTV system towards an external law enforcement facility, as its request or pushed by the CCTV system to allow hand-over in crisis situations.

Provisions will allow for the implementation of the applicable security, integrity and privacy protection measures.

As detailed hereafter, this export file format relies on several requirements for the elementary lower level formats and related protocols, respectively:

— Video;

— Audio;

— Metadata:

  — Static (localization, camera identifier, etc.)

  — Dynamic (date, time, pan, tilt, zoom, identification results, etc.)

— Container structure to integrate the above;

— Data security and integrity;

— Provisions for privacy.

The implementation of this International Standard shall be such that widely available Operating System (OS) independent tools will allow for minimal processing of received standard files by societal security organizations, such as forensic investigators, ensuring as a minimum the following and any combination thereof:

— Videos and metadata display;

— Direct access to the metadata without display of the videos;

— Selection of content time slots;

— Access to the sources defined by name or scene-location.

4.2 Relation with other standards

The different formats and prescriptions mentioned above are to a large extent subsets of existing standards; as a result, this International Standard is a blend of profiles of standards and practices which combined will achieve the expected minimum level of interoperability.

5 Requirements

5.1 General

The minimum requirement for a video-surveillance system to comply with this International Standard is to produce export files based on the following requirements.

These export files are essentially self-contained Audio-Video Packages which contain all information necessary to use the data, including location and time of all the scenes, without requiring access to the source system.

NOTE The requirements detailed hereafter are compatible with the implementation of the IEC 62676 series that make use of the data and of the stream payloads defined therein.

5.2 Requirements for the Audio-Video Package

5.2.1 General

The "Audio-Video Package" shall consist of a structure container, made of an indexing descriptor document (XML) and a collection of video and audio data, as well as metadata files. All of these files shall be in one hierarchical folder (or file system as may be the appropriate terminology for the operating system under which they are created).

A folder shall contain any number (and without any limitation) of the individual files described below and will cover an unlimited number of elementary contiguous time slots.

For easy random access to information, data shall be split into Directory Time Slots (DTS) and File Time Slots (FTS), where time is given in Coordinated Universal Time (UTC), as for GPS. Duration of a DTS is not critical; for operator convenience (typically, fast search of events by investigators), this duration shall be constant for a system and a new DTS shall be initiated at least once per day, starting with one hour.

As a minimum, any exported package shall contain once the descriptor of each of the Directory Time Slots covering (partially or in full) the corresponding time slot. This descriptor shall contain a time-stamped table of the changes that may have occurred in the configuration since the beginning of the DTS as well as the list of the contained FTS with their start time.

The duration of a File Time Slot (FTS) shall be comprised between 1 and 600 s. Within a time slot indexes shall allow to point accurately to any specific frame and time.

NOTE 1 It is allowed to zip or tar the package (again according to the operating environment of the source) and to transport it by any convenient means including placement on a portable storage media or file transfer protocol. These processes are not a part of this International Standard. Only the format of the package is covered in this International Standard.

An Audio-Video Package shall accordingly be organized in a structure as figured below with, in one top folder, a collection of time slices, arranged in DTS and FTS, of the relevant audio, video and other dynamic data contents and for each of them a description giving all available information relative to the corresponding time slice. Such a generic package organization is detailed in Figure 2.

For practical reasons, data naturally belong to groups, typically corresponding to a physical infrastructure, to data collected by a same Network Video Recorder (NVR) or to an administrative district; relationship with such groups shall be part of the description.

As a minimum interoperability condition, the naming rule shall be: DTS_info.xml.

