INTERNATIONAL STANDARD

ISO/IEC 30137-4

First edition
2021-06

Information technology — Use of biometrics in video surveillance systems —

Part 4: Ground truth and video annotation procedure

Reference number
ISO/IEC 30137-4:2021(E)

COPYRIGHT PROTECTED DOCUMENT

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: 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.

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 or www.iec.ch/members_experts/refdocs).

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) or the IEC list of patent declarations received (see patents.iec.ch).

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

For an explanation of 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 www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 37, *Biometrics*.

A list of all parts in the ISO/IEC 30137 series can be found on the ISO and IEC websites.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

Considerable improvements in the performance of automated face recognition (AFR) have resulted in applications such as automated border controls, where facial images encoded in ePassports are compared with the face presented by a traveller at a control point. The success of these first generation AFR systems has encouraged suppliers to consider other applications, where the subject is not necessarily aware of the use of biometric comparison and where the environment for collection of images can be far from optimal. The inferior performance in such less-controlled identification applications can necessitate a greater involvement by trained personnel.

The ISO/IEC 30137 series provides guidance on the use of biometric technologies (primarily automated face recognition) in video surveillance systems (VSS) for several scenarios, including real-time operation against watchlists and post-event analysis of video data. The ISO/IEC 30137 series includes guidance on the selection and placement of cameras through to system specification, testing and maintenance. The ISO/IEC 30137 series uses the term VSS to replace the older but commonly used term, closed circuit television (CCTV).

The ISO/IEC 30137 series addresses the annotation of human beings. It is not intended to provide for annotation of non-human objects such as cars, animals, or luggage.

Records conformant to this document can be produced from video in either of the following ways:

— automatically, in which software analyses video and estimates quantities defined in this document, or

— manually, in which human reviewers annotate video with a goal of producing ground truth video annotation, which can be used by a receiving system (i.e. any service or device that decodes, interprets and uses standardized data).

This supports several applications, including:

— People counting:

  — stating of the number of people present in a location,

  — stating of the number of people traversing a given point or volume,

  — stating of population density (e.g. in crowds),

  — measurement of crowd densities,

  — performance of crowd behavioural analyses.

— Automated detection and tracking:

  — automated enrolment (addition) of subjects to a watchlist, exhaustively or after behavioural analysis,

  — detection of subjects, and parts of subjects (e.g. faces),

  — tracking of subjects through time, e.g. following motion in a single video,

  — tracking of subjects appearing through camera networks, including cases where a subject is viewed simultaneously by different cameras, and cases where the subject appears sequentially before several cameras,

  - re-identification, the process of connecting an identity of a subject across two or more video sequences.
- Automated identification:
  - in law enforcement, looking for subjects of interest present on watchlists (negative identification, blacklists),
  - in law enforcement, applications in review of post-event VSS video from one or multiple cameras against watchlists,
  - in private commercial settings, looking for individuals to be given preferential service,
  - identification of cooperative enrolled subjects (positive access control, whitelists).

This document includes annotation of the following information:

- Imaging type: single camera, sequential cameras, stereo cameras, combination, camera capture spectrum.
- When the subject appears in the video (start time) and when they leave (end time).
  - Brief description of the subject (what can be seen in the video?).
- Where and when the face of the subject appears.
  - Brief description of the face (pose, orientation, expression, occlusion).
- Intermediate tracking points between the start and end times, for subject and face.
- Absolute description of the subject:
  - estimated age, sex,
  - hair and eye colour,
  - estimated height and corpulence,
  - clothing and clothing colour,
  - glasses/hat,
  - best subject image or best subject face image.
- Subject interactions with other subjects and groups.
- Subject interactions with other video elements (bag, car, etc.).
- Known identity of the subject.
- The presence of other subjects who are not annotated.
- Regions of interest, outside of which an algorithm or receiving system would not operate.
- Absence: Where items of interest, including subjects, are known to be absent.

Standardized annotation supports evaluation, research and development, and operational deployment.

Information technology — Use of biometrics in video surveillance systems —

Part 4: Ground truth and video annotation procedure

1 Scope

This document establishes requirements for the annotation of humans, human faces and other body parts, and arbitrary objects appearing in imagery. It specifies the following:

— metadata to be inserted in a video stream;

— encoding of full and partial spatial and temporal ground truth information for:

  — objects present in a video, and

  — objects absent in a video;

— procedures for different annotation of known and unknown subjects.

This document does not specify:

— encoding of video data.

2 Normative references

There are no normative references in this document.