

ICS 35.240.15
CCS L 71



**NATIONAL STANDARD OF THE
PEOPLE'S REPUBLIC OF CHINA**

GB/T 44261.4-2024

**Information technology - Use of biometrics in video
surveillance systems - Part 4: Ground truth and video
annotation procedure**

——对全部和部分空间和时间的真值信息进行编码：

(English Translation)

Issued on: October 26, 2024

Implemented on: May 1, 2025

**Issued by: State Administration for Market Regulation, China National
Standardization Administration**

Table of Contents

Preface
Introduction
1 Scope
2 Normative references
3 Terms and Definitions
4 Abbreviations
5 Compliance
6 Supports information encoding of annotations
6.1 Overview
6.2 Area Marking
6.3 Encoding of Object Class Information
6.4 Object Information Encoding
6.5 Marking Code
6.6 Frame Timestamp Encoding
6.7 Frame and Interval Coding
6.8 Trajectory Coding
6.9 Imaging System Information Coding
7 Labeling video sequences
7.1 Overview
7.2 Track Annotation in Video Sequences
7.3 Missing Annotations in Video Sequences
7.4 Annotation of Count Information
Appendix A (Normative) XSD Schema for this document
Reference

Foreword

This document is in accordance with the provisions of GB/T 1.1-2020 "Guidelines for standardization work Part 1: Structure and drafting rules for standardization documents" Drafting.

This document is Part 4 of GB/T 44261 "Information technology - Application of biometric identification technology in video surveillance systems".

GB/T 44261 has published the following parts.

--- Part 1: System design and specifications;

--- Part 4: Ground truth and video annotation procedures.

This document is equivalent to ISO/IEC 30137-4:2021 "Application of information technology biometric recognition technology in video surveillance systems" Part 4: Ground truth and video annotation procedures.

Introduction

GB/T 44261 "Application of Information Technology Biometric Identification Technology in Video Surveillance Systems" stipulates the application of information technology biometric identification technology in video surveillance systems.

How to better use and evaluate biometric recognition technology, including an introduction to the overall architecture, description of testing and reporting, data formats, and real-time GB/T 44261 is intended to consist of four parts.

--- Part 1: System design and specifications. The purpose is to standardize and guide the overall application of biometric recognition technology in video surveillance systems.

It is applicable to the design and use of video surveillance systems with biometric recognition functions.

--- Part 2: Performance testing and reporting. The purpose is to clarify the performance evaluation indicators of biometric recognition technology in video surveillance systems.

The criteria and report contents are applicable to the evaluation of the use of biometric recognition technology in video surveillance systems.

--- Part 3: Data format. The purpose is to standardize the related data generated when using biometric recognition technology in video surveillance systems.

The data recording format is suitable for the reuse and exchange of biometric identification related data between different video surveillance systems.

--- Part 4: Truth value and video annotation procedures. The purpose is to standardize the annotation method of video data, namely XSD (XMLSchema the performance of automatic face recognition (AFR) has been significantly improved, making it widely used in applications such as automatic border control and civil aviation airport video surveillance.

Widespread use. In these applications, the facial image encoded in the e-passport is compared with the face presented by the traveler at the control point.

The success of the first generation of AFR systems has encouraged vendors to consider other application scenarios where the subject does not have to know the biological the use of feature comparison, and the environment in which images are collected may be far from

optimal, therefore, in this recognition application scenario with low controllability, Poor performance may require greater involvement of trained personnel.

GB/T 44261 provides guidance on the application of biometric recognition technology (mainly AFR) in various scenarios in video surveillance systems (VSS).

The GB/T 44261 series includes camera selection and GB/T 44261 uses the term VSS to replace the older but more commonly used term - closed circuit Television (CCTV).

The main purpose of this paper is to address the problem of labeling human objects, rather than providing annotations for non-human objects (such as cars, animals, or luggage).

Provides marking method.

Records conforming to this document may be generated from video by any of the following means.

--- Automatically generate and use software to perform video analysis and quantity estimation on the annotation content defined in this document;

---Manual generation, the reviewer aims to generate the true value of the video annotation, and the result can be used by the receiving system (i.e. decoding, any service or device that parses and consumes standardized data).

This document supports a variety of application scenarios, including the following.

---Personnel statistics.

- The number of people present at a particular location;
- The number of people moving within the designated points or areas;
- Population density (e.g., crowds);
- Measure crowd density;
- Analysis of crowd behavior.

---Automatic detection and tracking.

- Automatically register (add) the subject to a watch list, either directly or after behavioral analysis;
- Detecting subjects and certain parts of them (e.g. faces);
- Tracking a subject over a certain timeframe, for example, tracking the movement of a subject within a single video;
- Track the subject by networking the cameras, including the subject being observed by different cameras at the same time, or the subject being observed by different cameras at the same time.

Appearing in front of multiple cameras;

- Re-identification, which is the process of associating the identity of the same subject in two or more video sequences.

---Automatic identification.

- In the process of law enforcement, search for relevant entities on the watch list (denied identification, blacklist);

- During law enforcement, review post-event VSS video from one or more cameras against a watch list;

- In an individual-oriented business environment, identify individuals who need priority service;

- Identify registered partners (positive access control, whitelist).

This document includes the following information.

---Imaging type. single camera, sequential camera, stereo camera, combined camera, spectral camera.

---The time when the subject appears in the video (start time) and leaves (end time).

- A brief description of the subject (what can be seen in the video).

---When and where the subject's face appears.

- A brief description of the face (pose, orientation, expression, occlusions).

---The tracking points between the subject and the face from the start to the end.

---Structured description of the subject.

- Assess age and gender;

- Hair and eye color;

- Estimate height, weight;

- Clothing and clothing colors;

- Glasses, hats;

- Best subject image, best subject face image.

---Interaction of the subject with other subjects and groups.

---Interaction of the subject with other video elements (bags, cars, etc.).

---The known identity of the subject.

---There are other unmarked entities.