工業技術研究院
Industrial Technology
Research Institute

AI-Driven Visual Intelligence Applications
- Person Analytics
- Markerless 3D Motion Capture

**Information and Communications
Research Laboratories (ICL)**

Yi-Yu Su
2025/09

Person Analytics

§ Flow Analysis

§ Cross Camera Person Re-ID
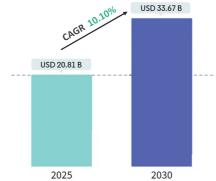
工業技術研究院
Industrial Technology
Research Institute

Rapid Market for AI Surveillance Applications

§ Global surveillance camera market to reach $43.56B in 2025, growing to $62.57B by 2029 (CAGR 9.5%).

§ Asia-Pacific leads market growth, projected at $20.81B in 2025 and $33.67B by 2030 (CAGR 10.10%).

Global surveillance camera market
Source：The Business Research Company (2025)

Asia-Pacific surveillance camera market
Source：Mordor Intelligence (2024)

Commercial Space Application Scenarios
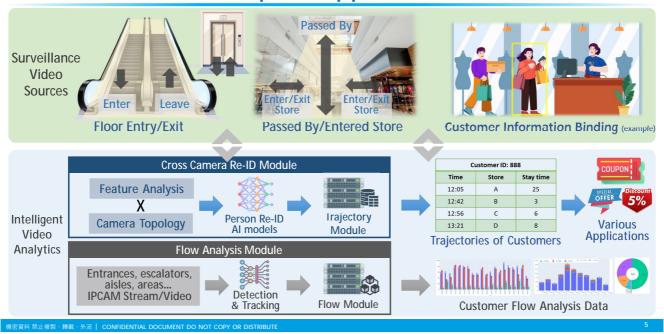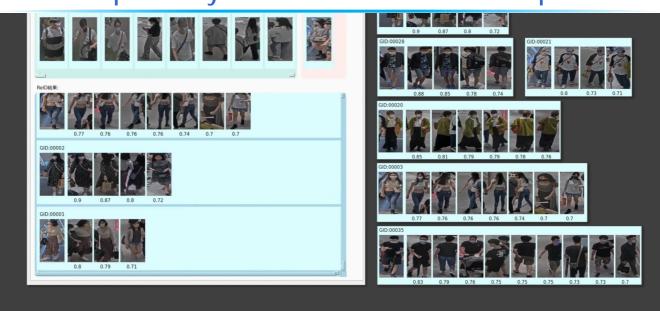

Surveillance Video Sources

Floor Entry/Exit — Enter / Leave

Passed By/Entered Store — Passed By, Enter/Exit Store, Enter/Exit Store

Customer Information Binding (example)

Intelligent Video Analytics

Cross Camera Re-ID Module
- Feature Analysis X Camera Topology
- Person Re-ID AI models
- Trajectory Module

Flow Analysis Module
- Entrances, escalators, aisles, areas... IPCAM Stream/Video
- Detection & Tracking
- Flow Module

Customer ID: 888

Time	Store	Stay time
12:05	A	25
12:42	B	3
12:56	C	6
13:21	D	8

Trajectories of Customers

Various Applications

Customer Flow Analysis Data

People Analytics – Tracking & Positioning

Total Number 10   Floor Plan
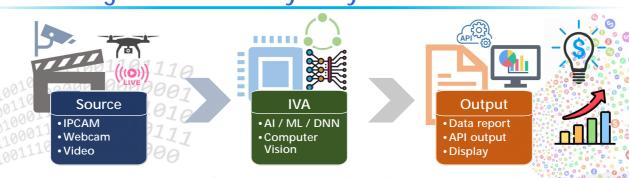
and spatial localization within
a common commercial setting.

People Analytics – Re-ID in Commercial Space

91% high accuracy rate of re-identification.

Intelligent Video Analysis System Architecture

Source
- IPCAM
- Webcam
- Video

IVA
- AI / ML / DNN
- Computer Vision

Output
- Data report
- API output
- Display

Various Sources Support
- IPCAM/WebCAM
- RTSP live stream / video files
- H.264 / H.265 video codec
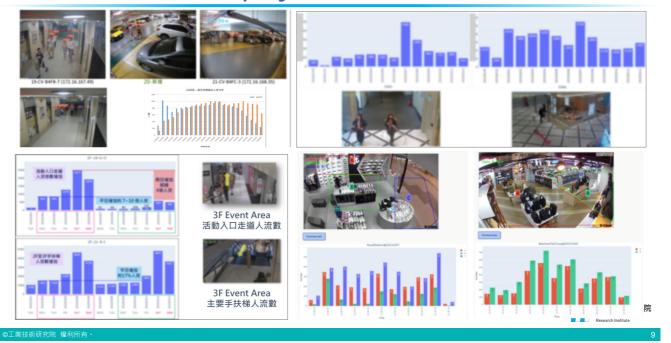
Flexible System & AI Models
- Flexible system architectures: Cloud/On-premise/Hybrid
- AI Modules: Person Detection/ Tracking and Localization, Gender/ Age Recognition, Cross-Camera Person Re-identification, etc.

Seamlessly Integrable
- Cloud/Edge/Hybrid integration
- Comprehensive API (e.g., Restful)
- Enterprise Account Integration (e.g., LDAP)
- Web-based Data Management and System Maintenance Interface

Markerless
3D Motion Capture

Motion Capture Technologies & Challenges

§ Traditional Motion Capture (MoCap): Optical & Inertial
- Requiring suits/markers, and complex setup / operate.
- Mainly for professional filming / sports science / gaming...

§ Marker-less Motion Capture
- Utilizing general cameras with image analysis to obtain 2D/3D human pose
- No suits/markers are needed, flexible with setup and portability.
- Challenges：
  - Depth ambiguity
  - Complex poses and occlusions
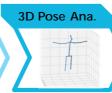 - Motion stability and smoothness
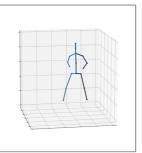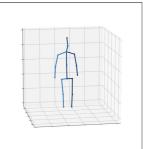 - Multi-person tracking

MoCap applied to virtual characters

Optical MoCap Inertial MoCap

Marker-less MoCap

Single-Camera Operation (2D to 3D HPE)

LIVE Capture | 2D Pose Est. | 3D Pose Ana. | Processing | Integration

Video Test – Fitness and Exercise (source - YouTube) Video Test – Mobility Assessment (source - YouTube)

HPE : Human Pose Estimation
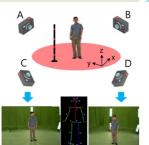
Markerless 3D Motion Capture

2D Motion Capture	3D Pose Estimation	3D Data Fusion	3D Data Output

Multi-camera 3D calibration and positioning

2D stream → 3D keypoints

3D pose data fusion

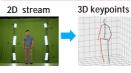**3D pose data smoothing**

Standard 3D data format

BVH FBX

MAYA blender

unity MOBU

- Multi-camera synchronization and calibration system : rapid, accurate, and flexible.
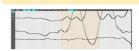- 2D keypoints filtering : corrects occlusions, intersections, or false detections.

- Spatiotemporal 3D reconstruction algorithm : combining spatial features intersection and temporal analysis.
- Optimization of joint angles to ensure realistic poses.

- 3D pose keypoints data processing mechanism, integrating kinematics to reconstructs 3D pose and motion information.
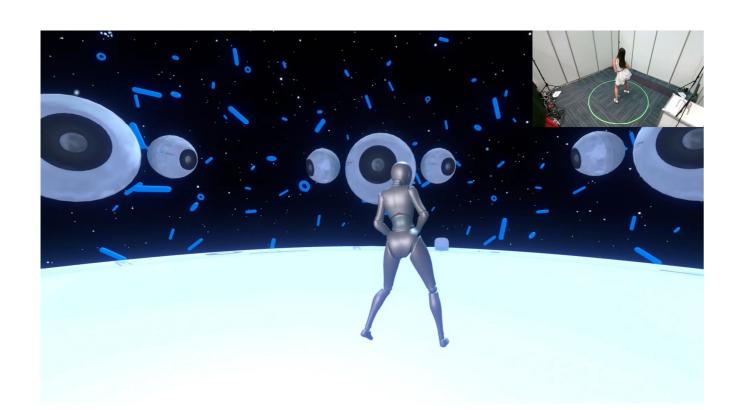- Motion smoothing mechanism to reproduce natural movement.

- Conversion to a standard 3D skeletal data format, compatible with most 3D animation software for applying virtual and augmented reality applications.
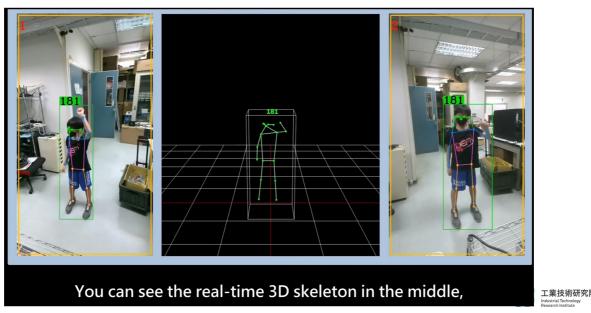
工業技術研究院
Industrial Technology
Research Institute

A Lighter Front-Facing 2 Cameras Setup

You can see the real-time 3D skeleton in the middle,

Technical Features and Application Scenarios

Analysis	Single-Camera	Multi-Camera
Method	Single-cam image → 2D human skeleton → 3D estimation model → 3D skeleton data.	Multi-cam images → Multiple 2D skeleton → 3D fusion/reconstruction → 3D skeleton data.
Pros	• Requires less setup time. • Easy to operate and lower overall cost.	• High 3D skeleton precision & movement stability. • Excellent for multi-person analysis.
Cons	• Lower precision and may not perform well with specific movements. • Best suited for single-person applications.	• Requires more setup time, more complex system installation and operation • Demands more space, and higher overall cost.
Features	• Suitable for applications with general precision requirements and lower budgets. • Fitness tracking or activity/energy measurement, etc.	• Suitable for applications that require high precision and movement fidelity. • Sports science, professional performances, etc.
Equipment	1 camera (webcam) + a main computer.	4 cameras (webcams) + a main computer with a GPU card.

工業技術研究院
Industrial Technology
Research Institute