

Recommendation **ITU-T F.747.14 (06/2024)**

SERIES F: Non-telephone telecommunication services

Multimedia services

Requirements and capability framework of the multimodal fusion system for vision

ITU-T F-SERIES RECOMMENDATIONS
Non-telephone telecommunication services

TELEGRAPH SERVICE	F.1-F.109
Operating methods for the international public telegram service	F.1-F.19
The gentex network	F.20-F.29
Message switching	F.30-F.39
The international telemesssage service	F.40-F.58
The international telex service	F.59-F.89
Statistics and publications on international telegraph services	F.90-F.99
Scheduled and leased communication services	F.100-F.104
Phototelegraph service	F.105-F.109
MOBILE SERVICE	F.110-F.159
Mobile services and multideestination satellite services	F.110-F.159
TELEMATIC SERVICES	F.160-F.399
Public facsimile service	F.160-F.199
Teletex service	F.200-F.299
Videotex service	F.300-F.349
General provisions for telematic services	F.350-F.399
MESSAGE HANDLING SERVICES	F.400-F.499
DIRECTORY SERVICES	F.500-F.549
DOCUMENT COMMUNICATION	F.550-F.599
Document communication	F.550-F.579
Programming communication interfaces	F.580-F.599
DATA TRANSMISSION SERVICES	F.600-F.699
MULTIMEDIA SERVICES	F.700-F.799
ISDN SERVICES	F.800-F.849
UNIVERSAL PERSONAL TELECOMMUNICATION	F.850-F.899
ACCESSIBILITY AND HUMAN FACTORS	F.900-F.999

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T F.747.14

Requirements and capability framework of the multimodal fusion system for vision

Summary

With the rampant development of big data and artificial intelligence some new modal types data for vision have emerged, such as light detection and ranging (LiDAR) data, radar data 3D data, etc. Multimodal fusion for vision is one new technology for higher quality human eye presentation or higher performance machine vision tasks. The multimodal fusion system for vision will be popular among the future vision systems. The multimodal fusion system focuses on specific methods of multimodal data fusion.

Recommendation ITU-T F.747.14 specifies the requirements and capability framework to accommodate the existing and newly developed visual facilities for multimodal fusion for vision applications including autonomous vehicles, autonomous robot navigation and weather visibility prediction. This Recommendation defines the related requirements, capability framework and reference use cases for the multimodal fusion system for vision.

History*

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T F.747.14	2024-06-13	16	11.1002/1000/15908

Keywords

Capability framework, multimodal fusion, requirements, vision.

* To access the Recommendation, type the URL <https://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents/software copyrights, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the appropriate ITU-T databases available via the ITU-T website at <http://www.itu.int/ITU-T/ipr/>.

© ITU 2024

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Table of Contents

	Page
1 Scope	1
2 References.....	1
3 Definitions	1
3.1 Terms defined elsewhere	1
3.2 Terms defined in this Recommendation.....	2
4 Abbreviations and acronyms	2
5 Conventions	2
6 Overview of multimodal fusion for vision	2
7 Requirements of multimodal fusion system for vision.....	3
7.1 Data acquisition	3
7.2 Data transmission	4
7.3 Data fusion.....	4
7.4 Data analysis.....	4
7.5 Security and personally identifiable information protection	5
8 Capability framework of multimodal fusion system for vision.....	5
8.1 Capabilities for data acquisition	5
8.2 Capabilities for data transmission	5
8.3 Capabilities for data processing.....	6
8.4 Capabilities for data fusion.....	6
8.5 Capabilities for data analysis.....	6
Appendix I – Use cases of multimodal fusion system for vision.....	7
I.1 Use case of autonomous vehicles	7
I.2 Use case of autonomous robot navigation.....	7
I.3 Use case of weather visibility prediction.....	8
Bibliography.....	9

