

ETSI TR 126 985 V18.0.0 (2024-05)



**5G;  
Vehicle-to-everything (V2X);  
Media handling and interaction  
(3GPP TR 26.985 version 18.0.0 Release 18)**



---

**Reference**

RTR/TSGS-0426985vi00

---

**Keywords**

5G

**ETSI**

650 Route des Lucioles  
F-06921 Sophia Antipolis Cedex - FRANCE

---

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B  
Association à but non lucratif enregistrée à la  
Sous-Préfecture de Grasse (06) N° w061004871

---

**Important notice**

The present document can be downloaded from:

<https://www.etsi.org/standards-search>

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format at [www.etsi.org/deliver](http://www.etsi.org/deliver).

Users of the present document should be aware that the document may be subject to revision or change of status.

Information on the current status of this and other ETSI documents is available at

<https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx>

If you find errors in the present document, please send your comment to one of the following services:

<https://portal.etsi.org/People/CommitteeSupportStaff.aspx>

If you find a security vulnerability in the present document, please report it through our  
Coordinated Vulnerability Disclosure Program:

<https://www.etsi.org/standards/coordinated-vulnerability-disclosure>

---

**Notice of disclaimer & limitation of liability**

The information provided in the present deliverable is directed solely to professionals who have the appropriate degree of experience to understand and interpret its content in accordance with generally accepted engineering or other professional standard and applicable regulations.

No recommendation as to products and services or vendors is made or should be implied.

No representation or warranty is made that this deliverable is technically accurate or sufficient or conforms to any law and/or governmental rule and/or regulation and further, no representation or warranty is made of merchantability or fitness for any particular purpose or against infringement of intellectual property rights.

In no event shall ETSI be held liable for loss of profits or any other incidental or consequential damages.

Any software contained in this deliverable is provided "AS IS" with no warranties, express or implied, including but not limited to, the warranties of merchantability, fitness for a particular purpose and non-infringement of intellectual property rights and ETSI shall not be held liable in any event for any damages whatsoever (including, without limitation, damages for loss of profits, business interruption, loss of information, or any other pecuniary loss) arising out of or related to the use of or inability to use the software.

---

**Copyright Notification**

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.

The content of the PDF version shall not be modified without the written authorization of ETSI.

The copyright and the foregoing restriction extend to reproduction in all media.

© ETSI 2024.  
All rights reserved.

---

# Intellectual Property Rights

## Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: "*Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards*", which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (<https://ipr.etsi.org/>).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

## Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

**DECT™**, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™** and **LTE™** are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM®** and the GSM logo are trademarks registered and owned by the GSM Association.

---

# Legal Notice

This Technical Report (TR) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities. These shall be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between 3GPP and ETSI identities can be found under <https://webapp.etsi.org/key/queryform.asp>.

---

# Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the [ETSI Drafting Rules](#) (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

# Contents

Intellectual Property Rights .....	2
Legal Notice .....	2
Modal verbs terminology.....	2
Foreword.....	5
Introduction .....	6
1 Scope .....	7
2 References .....	7
3 Definitions, symbols and abbreviations .....	8
3.1 Definitions .....	8
3.2 Symbols.....	9
3.3 Abbreviations .....	9
4 V2X overview .....	9
5 Media use cases in V2X .....	9
5.1 Support for remote driving .....	9
5.2 Information sharing for high/full automated driving.....	10
5.3 Video data sharing for assisted and improved automated driving (VaD).....	10
5.4 Teleoperated support (TeSo).....	10
5.5 Video composition .....	10
6 System architecture .....	10
6.1 General .....	10
6.2 System.....	10
6.3 Terminal .....	11
6.4 Procedures .....	12
6.4.1 Support for remote driving .....	12
6.4.1.1 Description.....	12
6.4.1.2 Leveraging 5G QoS framework.....	13
6.4.1.3 Potential requirements.....	14
6.4.2 Information sharing for high/full automated driving .....	15
6.4.3 Video data sharing for assisted and improved automated driving (VaD) .....	15
6.4.4 Teleoperated support (TeSo).....	15
6.4.5 Video composition.....	15
6.4.6 Data collection from in-vehicular sensors for HD map updates .....	15
6.4.6.1 Description .....	15
6.4.6.2 Potential requirements.....	16
7 Protocols.....	16
7.1 General .....	16
7.2 Uplink.....	16
7.3 Downlink.....	17
7.4 Direct link.....	17
8 Media consideration .....	17
8.1 General .....	17
8.2 Video.....	17
8.2.1 Resolution, Framerate, and Bitrate .....	17
8.2.1.1 Resolution Aspects.....	17
8.2.1.2 Frame Rate .....	18
8.2.1.3 Bitrates / quality .....	19
8.2.2 Adaptation.....	19
8.2.2.1 General.....	19
8.2.2.2 Level of Adaptation Control .....	20
8.2.3 Operational Variance .....	20
8.2.4 Video Configuration Requirements .....	20

9 Sensor sharing of object information .....21

9.1 General .....21

9.2 On-going work .....21

9.2.1 ETSI-ITS .....22

9.2.2 5GAA.....22

9.2.3 SAE.....22

9.3 Example objects .....23

9.4 Considerations .....23

9.4.1 Liability .....23

9.4.2 Multi-mode media .....23

9.4.3 Message frequency .....23

9.4.4 Device Considerations .....24

9.5 Conclusion.....25

10 Conclusion.....26

**Annex A: Change history .....27**

History .....28

---

# Foreword

This Technical Report has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

- x the first digit:
  - 1 presented to TSG for information;
  - 2 presented to TSG for approval;
  - 3 or greater indicates TSG approved document under change control.
- y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.
- z the third digit is incremented when editorial only changes have been incorporated in the document.

In the present document, modal verbs have the following meanings:

- shall** indicates a mandatory requirement to do something
- shall not** indicates an interdiction (prohibition) to do something

The constructions "shall" and "shall not" are confined to the context of normative provisions, and do not appear in Technical Reports.

The constructions "must" and "must not" are not used as substitutes for "shall" and "shall not". Their use is avoided insofar as possible, and they are not used in a normative context except in a direct citation from an external, referenced, non-3GPP document, or so as to maintain continuity of style when extending or modifying the provisions of such a referenced document.

- should** indicates a recommendation to do something
- should not** indicates a recommendation not to do something
- may** indicates permission to do something
- need not** indicates permission not to do something

The construction "may not" is ambiguous and is not used in normative elements. The unambiguous constructions "might not" or "shall not" are used instead, depending upon the meaning intended.

- can** indicates that something is possible
- cannot** indicates that something is impossible

The constructions "can" and "cannot" are not substitutes for "may" and "need not".

- will** indicates that something is certain or expected to happen as a result of action taken by an agency the behaviour of which is outside the scope of the present document
- will not** indicates that something is certain or expected not to happen as a result of action taken by an agency the behaviour of which is outside the scope of the present document
- might** indicates a likelihood that something will happen as a result of action taken by some agency the behaviour of which is outside the scope of the present document

**might not** indicates a likelihood that something will not happen as a result of action taken by some agency the behaviour of which is outside the scope of the present document

In addition:

**is** (or any other verb in the indicative mood) indicates a statement of fact

**is not** (or any other negative verb in the indicative mood) indicates a statement of fact

The constructions "is" and "is not" do not indicate requirements.

---

## Introduction

With connected car often considered as one of key applications of 5G, 3GPP has been working on the features and required technologies of vehicle-to-everything (V2X), in which coordinated efforts are made by vehicles and networks to realize advanced driver assistance systems (ADAS) or self-driving vehicles. The initial works on V2X focused on the exchange or delivery of short control messages to avoid collision, and did not consider the functionalities required for advanced connected cars. However, the use cases for advanced driving of eV2X, specified in TR 22.886, include various scenarios including media, such as video and sensor data.

Although more advanced transmission capabilities are required, media can overcome the limitations of coded messages that might not be able to describe arbitrary situations on the road, and can be a more intuitive source of information for the decision of driver or artificial intelligence, e.g., by providing visual information that cannot be provided by the on-board cameras or sensors of vehicles with limited coverage or accuracy. Strong tolerance to the variance of weather is another advantage of media-based V2X that can exploit the point-to-point transmission or multicasting/broadcasting capability of 3GPP radio access.

---

# 1 Scope

In the present document, media-related use cases of V2X described in TR 22.886 and their requirements in TS 22.186 are analysed in detail, to clarify the operation of advanced driving using networked visual information. The required procedures for media capture, compression, and transmission, based on the mechanisms for media handling and transportation, such as MTSI, MBMS, and FLUS, are investigated. Operation over the PC5 interface is currently outside of the scope of this document.

---

# 2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

- [1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".
- [2] 3GPP TR 22.885: "Study on LTE support for Vehicle to Everything (V2X) services".
- [3] 3GPP TR 22.886: "Study on enhancement of 3GPP Support for 5G V2X Services".
- [4] 3GPP TS 22.185: "Service requirements for V2X services".
- [5] 3GPP TS 22.186: "Enhancement of 3GPP Support for V2X Scenarios".
- [6] 3GPP TR 23.785: "Study on architecture enhancements for LTE support of V2X services".
- [7] 3GPP TS 23.285: "Architecture enhancements for V2X services".
- [8] 3GPP TR 23.786: "Study on architecture enhancements for EPS and 5G System to support advanced V2X services".
- [9] 3GPP TS 23.468: "Group Communication System Enablers for LTE (GCSE\_LTE); Stage 2".
- [10] 3GPP TR 33.885: "Study on security aspects for LTE support of Vehicle-to-Everything (V2X) services".
- [11] 3GPP TS 33.185: "Security aspect for LTE support of Vehicle-to-Everything (V2X) services".
- [12] 3GPP TR 36.885: "Study on LTE-based V2X services".
- [13] SAE International, "SAE J3016 Taxonomy and Definitions for Terms Related to On-road Motor Vehicle Automated Driving Systems".
- [14] P. Gomes, F. Vieira and M. Ferreira, "The See-Through System: From implementation to test-drive," Vehicular Networking Conference (VNC), 2012 IEEE, Seoul, 2012, pp. 40-47.
- [15] 5G-PPP-White-Paper-on-Automotive-Vertical-Sectors.
- [16] void.
- [17] 3GPP TS 26.238 "Uplink Streaming" (FLUS).
- [18] 3GPP TR 26.939 "Guidelines on the Framework for Live Uplink Streaming (FLUS) ".
- [19] Automotive Edge Computing Consortium (AECC) whitepaper on "General Principle and Vision", 25<sup>th</sup> April 2018, available [https://aecc.org/wp-content/uploads/2018/02/AECC\\_White\\_Paper.pdf](https://aecc.org/wp-content/uploads/2018/02/AECC_White_Paper.pdf).



- [20] 3GPP TS 26.114: "IP Multimedia Subsystem (IMS); Multimedia Telephony; Media handling and interaction".
- [21] 3GPP TS 26.346: "Multimedia Broadcast/Multicast Service (MBMS); Protocols and codecs".
- [22] ETSI TS 102 894-1: "Intelligent Transport Systems (ITS); Users and applications requirements; Part 1: Facility layer structure, functional requirements and specifications".
- [23] ETSI TS 102 894-2: "Intelligent Transport Systems (ITS); Users and applications requirements; Part 2: Applications and facilities layer common data dictionary".
- [24] ETSI TS 103 324: "Intelligent Transport Systems (ITS); Specification of the Collective Perception Service".
- [25] ETSI TR 103 562: "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Analysis of the Collective Perception Service (CPS); Informative Report for the Collective Perception Service"
- [26] EU Project "Sustainable Intelligent Mining Systems (SIMS)", part of EU Research and Innovation Horizon 2020, <https://www.simsmining.eu/>
- [27] Ingemar Johansson, Siddarth Dadhich, Ulf Bodin, and Tomas Jönsson, "Adaptive Video with SCReAM over LTE for Remote-Operated Working Machines", 2<sup>nd</sup> August 2018, available <https://www.hindawi.com/journals/wcmc/2018/3142496/>
- [28] IETF RFC 8298: "Self-Clocked Rate Adaptation for Multimedia", Dec 2017, <https://tools.ietf.org/html/rfc8298>
- [29] ITU-T Recommendation H.264 (04/2013): "Advanced video coding for generic audiovisual services"
- [30] ITU-T Recommendation H.265 (04/2013): "High efficiency video coding".
- [31] SAE J2735 - Dedicated Short Range Communications (DSRC) Message Set Dictionary.

---

## 3 Definitions, symbols and abbreviations

### 3.1 Definitions

For the purposes of the present document, the terms and definitions given in 3GPP TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in 3GPP TR 21.905 [1]. The following definitions of V2X are from [2], [3].

**Road Side Unit:** A stationary infrastructure entity supporting V2X applications that can exchange messages with other entities supporting V2X applications.

**NOTE:** RSU is a term frequently used in existing ITS specifications, and the reason for introducing the term in the 3GPP specifications is to make the documents easier to read for the ITS industry. RSU is a logical entity that combines V2X application logic with the functionality of an eNB (referred to as eNB-type RSU) or UE (referred to as UE-type RSU).

**V2I Service:** A type of V2X Service, where one party is a UE and the other party is an RSU both using V2I application.

**V2N Service:** A type of V2X Service, where one party is a UE and the other party is a serving entity, both using V2N applications and communicating with each other via LTE network entities.

**V2P Service:** A type of V2X Service, where both parties of the communication are UEs using V2P application.

**V2V Service:** A type of V2X Service, where both parties of the communication are UEs using V2V application.

**V2X Service:** A type of communication service that involves a transmitting or receiving UE using V2V application via 3GPP transport. Based on the other party involved in the communication, it can be further divided into V2V Service, V2I Service, V2P Service, and V2N Service.

## 3.2 Symbols

For the purposes of the present document, the following symbols apply:

-

## 3.3 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

RSU	Road Side Unit
V2I	Vehicle-to-Infrastructure
V2N	Vehicle-to-Network
V2P	Vehicle-to-Pedestrian
V2V	Vehicle-to-Vehicle
V2X	Vehicle-to-Everything
SAE	Society of Automotive Engineers
LIDAR	Light Detection and Ranging

---

## 4 V2X overview

3GPP started a feasibility study on V2X as a message-based mechanism for improved safety and group driving (platooning), defining a variety of use cases including the interaction of vehicles, networks, and pedestrians, and their potential requirements [2]. As new needs for facilitating more advanced applications such as autonomous driving became elevated, the study continued to investigate the aspects of handling higher bit-rate sensor data and media [3]. Basic requirements set from these studies were specified in [4], [5].

From the outcomes of these feasibility studies, further studies on the network architecture that could meet the requirements of V2X followed [6], [7]. Initial consideration of V2X was in the operating environments of LTE but its applications were extended to 5G [8]. In addition to the conventional Uu interface between UE (vehicle) and eNodeB/gNB, PC5, a new interface enabling direct communication between vehicles, were defined, which re-uses the functionalities of group communication system enablers defined for mission-critical applications [9]. Based on the V2X architectures defined, security aspects were studied, and potential vulnerability of new interfaces were investigated [10], [11]. Finally, advanced features of radio access technologies are under study for higher performance [12].

---

## 5 Media use cases in V2X

The first five use cases are specified in TR 22.886 [3].

### 5.1 Support for remote driving

In this use case, specified in clause 5.4 of TR 22.886 [3], a vehicle is controlled remotely by either a human operator or cloud computing.

**NOTE 1:** The assumption is that each H.265/ HEVC stream is up to 10 Mbps and outside-vehicle video streams showing the lane situations around a vehicle are delivered to a remote driver. In addition, outside-vehicle audio streams may be delivered to a remote driver for conveying the noises and horn sounds from other vehicles.

NOTE 2: At least an inside-vehicle video stream and an inside-vehicle audio stream, whose delay requirements can be more relaxed than those of outside-vehicle video and audio, may be delivered to a remote driver for monitoring the situations.

## 5.2 Information sharing for high/full automated driving

This use case, specified in clause 5.11 of TR 22.886 [3], is interpreted as an automated driving at the level of e.g. SAE Level 4 and Level 5 automation [13]. Each vehicle shares its high resolution perception data (e.g., camera, LIDAR, occupancy grid) and/or detailed planned trajectory with other vehicles, directly or via the network.

An RSU may capture high resolution perception data around a corner or an obstacle, or an intersection. Each RSU shares its high resolution perception data with vehicles A, B, and C, via point-to-point transmission or multicast/broadcast.

## 5.3 Video data sharing for assisted and improved automated driving (VaD)

The visual range of the driver is in some road traffic situations obstructed, for instance by trucks driving in front [14], [15]. In this use case specified in clause 5.16 of TR 22.886 [3], video data sent from one vehicle to the other can support drivers in these safety-critical situations. Video data may also be collected and sent through a capable UE-type RSU.

## 5.4 Teleoperated support (TeSo)

Teleoperated Support (TeSo), specified in clause 5.21 of TR 22.886 [3], enables a single human operator to remotely control autonomous vehicles for a short period of time. TeSo enables efficient road construction (control of multiple autonomous vehicles from a single human operator), e.g., snow plowing.

## 5.5 Video composition

In this use case, specified in clause 5.25 of TR 22.886 [3], UEs have a camera and they take a video of the environment, and send this video to a server. The server can be in the cloud or in the near the UE point of attachment (i.e., Mobile Edge Computing (MEC)). The server/MEC will then post-process the videos received and combine the information in order to create a single video of the environment.

The videos may also be supplied by RSUs when there are UEs not sufficient for the composition. The UEs location and direction information, allows the server to accurately represent the location, relative speed and distance of vehicles, pedestrians, and any objects in that area.

---

# 6 System architecture

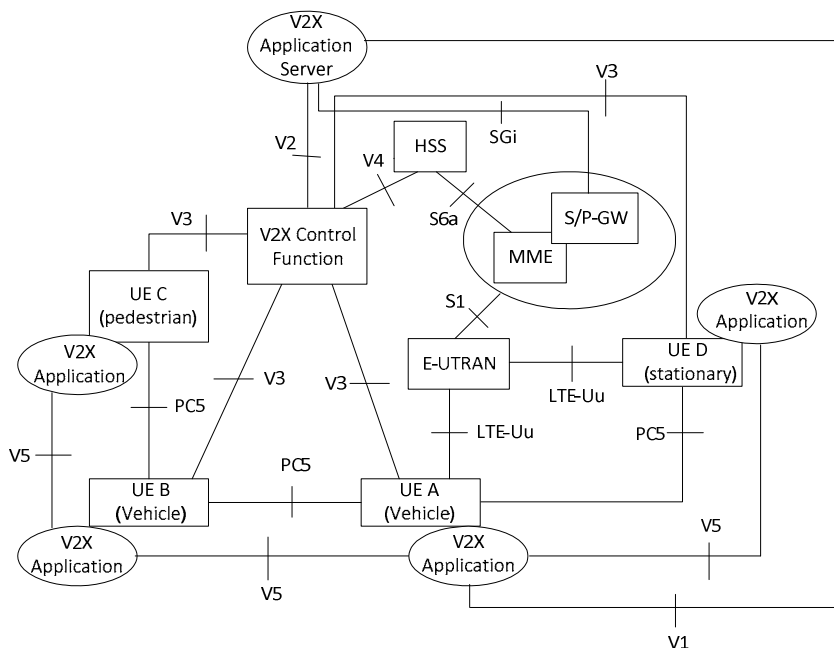
## 6.1 General

Many use cases of V2X including uplink and downlink transmission of media can be realized with the LTE-Uu interface and existing media services including MTSI, MBMS, or FLUS, without any architectural changes.

There are two modes of operation for V2X communication, i.e., over the PC5 and LTE-Uu interfaces. LTE-Uu provides the conventional point-to-point or point-to-many transport services at high bit-rate or QoS. PC5, compared to the Uu interface, is limited in reachable range but is generally better at localization of the communications, has shorter delays, and is less impacted by coverage issues.

## 6.2 System

Figure 6.2-1 shows a high level view of the non-roaming architecture for V2X system based on the PC5 and LTE-Uu interfaces. Functionalities of the network entities and reference points in this architecture can be found in [7].

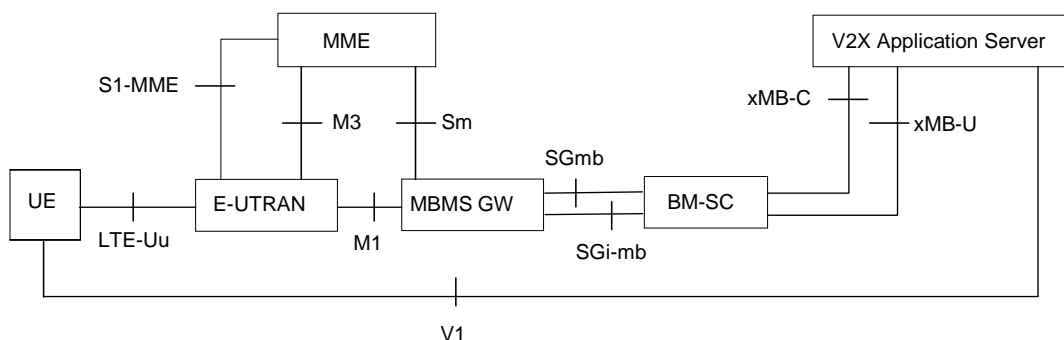


**Figure 6.2-1: Non-roaming reference architecture for PC5 and LTE-Uu based V2X communication**

NOTE 1: RSU is not an architectural entity but an implementation option that may be achieved by collocating a V2X application logic/server with some entities of the 3GPP system.

NOTE 2: V2X application servers in different domains can communicate with each other for the exchange of V2X messages but the interface between V2X application servers and methods of the message exchange is out of the scope of this document.

Figure 6.2-2 shows a high level view of the reference architecture with MBMS for LTE-Uu based V2X communication. The xMB reference point provides functionality overall for any content and supports security framework between content provider and BM-SC. The content in this context can be the information on traffic or road situation. A similar reference architecture using MB2 reference point can be found in [7].



**Figure 6.2-2: Reference architecture for MBMS for LTE-Uu based V2X communication via xMB**

### 6.3 Terminal

Figure 6.3-1 [7] shows three types of services, V2V, V2P, and V2I, a V2X system may provide. Depending on the use cases, a terminal in V2X may be a vehicle with Universal Integrated Circuit Card (UICC) embedded and capable of directly communicating with the network, a vehicle connected to the network via a smartphone, or a typical smartphone a pedestrian is carrying, which can transmit or receive V2X-related media or data.

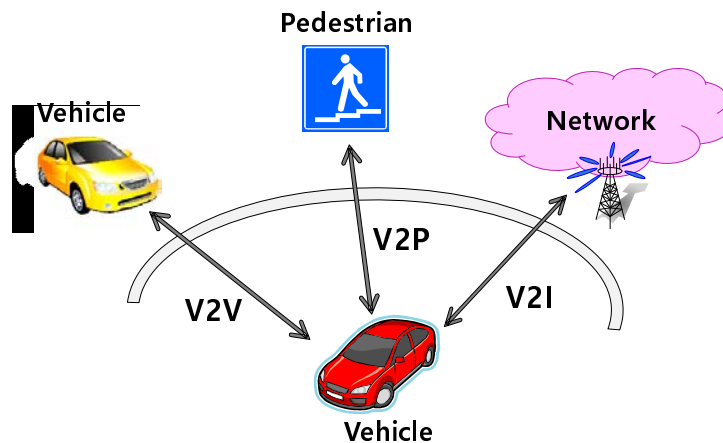


Figure 6.3-1: Types of V2X (V2V, V2P, and V2I)

## 6.4 Procedures

This clause outlines the strategies to realize V2X use cases with codecs and protocols.

### 6.4.1 Support for remote driving

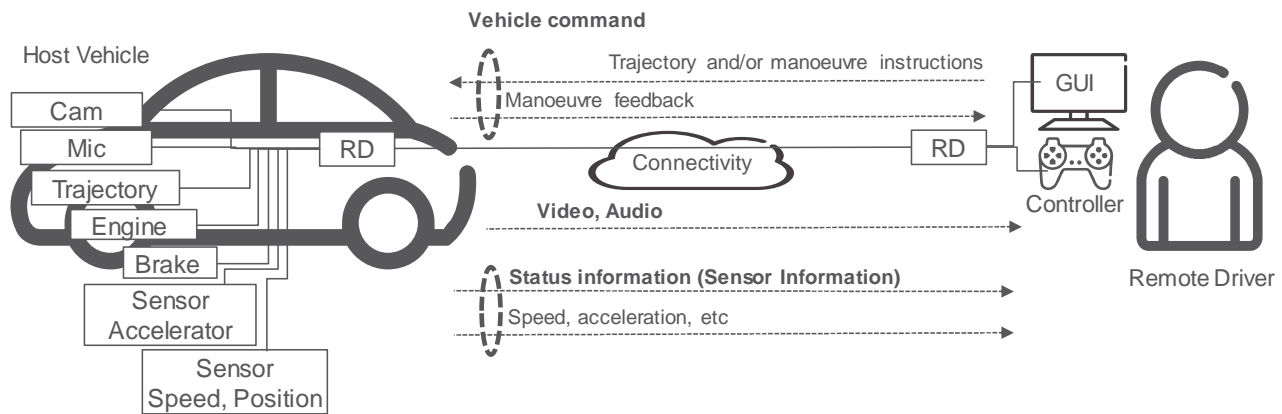
In this basic use case, media streams including the audio-visual information of scenes in several directions around the vehicle are transmitted in the uplink. It would be essential to maintain the QoS during the operation, e.g., by using MTSI or FLUS that supports IMS. Each stream may be assigned a bit-rate or a resolution that depends on its relative importance in the remote control of vehicle, which may also depend on the roadside situation.

For example, streams corresponding to the side or backward directions may be considered more important than the stream for the front when moving a car stopped in the first lane to the roadside. The inside-vehicles video may cover a wide angle or be made up of by stitching images captured by multiple cameras, as they are not used for controlling purposes. In contrast, the outside-vehicle videos would have to be captured as conventional images, to avoid the distortion that may compromise the accuracy of remote driving.

#### 6.4.1.1 Description

TS 22.186 [5] contains a use-case around Remote Driving (Clause 5.4 in TR 22.886 [3]). The use-case is subdivided into several sub cases. TR 22.886 lists a human as remote driver or a "cloud" as possible remote driver. In case of a "Cloud" based driver, a remote driving application server is deployed using cloud computing technologies.

Here, we focus on the case of a human as a remote driver. Figure 6.4-1 depicts an illustration of the setup. The Remote Driving (RD) application on the UE side collects information from various sensors and also connects to the engine, steering and braking system for command execution (automation system actuators). Note, this is an example list of potential sensors and actuators.



**Figure 6.4-1: Application flows in Remote Driving (RD) concept**

On the remote driver side, a Remote Driving (RD) application dispatches the incoming information to appropriate rendering devices, such as display or sound system. The RD is also converting the controller device actions into a protocol, which provides the manoeuvre instructions to the vehicle.

It is assumed that the vehicle has no (or very limited) additional local advanced driver assistance systems (ADAS), so the vehicle is a Level of Automation (LoA) 0 type of vehicle, which is remotely operated. Accordingly, there is a high dependency on the network performance, e.g. reliability, latency and bitrate.

It can be observed, that the architecture setup for sending video, audio and sensor data to the remote driver is very similar to the Uplink Streaming architecture (FLUS) in TS 26.238 [17] and TR 26.939 [18].

In some cases, a bi-directional speech channel can be available, which allows the remote driver to speak with vehicle passengers. This speech channel is separated from the video & audio traffic flows.

### 6.4.1.2 Leveraging 5G QoS framework

A number of different application traffic flows can benefit from the 3GPP QoS framework. There are different possibilities to provide QoS support for the different service flows.

When QoS support is available from the network, it is beneficial to separate the above described traffic flows due to different importance and QoS requirements.

- The uplink video & audio flows should be separated from the uplink sensor data flows to secure a different traffic handling priority of the uplink sensor data. Potentially, audio and video should be further separated, where video has a higher priority than audio (external microphone).
- When the sensor information is needed for haptic type of feedback, then the sensor data may have a higher priority than video and audio data.
- The manoeuvre instructions are sent on the downlink and should be separated from the other traffic due to importance.
- The conversational speech channel should be separated as usual conversational service traffic.

The traffic characteristics and the network performance expectation are:

**Manoeuvre instructions** (actor instructions) are remote control instructions from the remote driver to the vehicle to e.g. accelerate or brake. Manoeuvre instructions might be generated and sent when triggered by a control event, like a brake instruction. The vehicle acknowledges the reception of the instruction (e.g. implicitly via TCP or an explicit UDP acknowledge). Packet losses will delay the execution of the manoeuvre instruction, when the instruction needs to be retransmitted.

Key traffic characteristics:

- Event based instructions, e.g. instruction to brake triggered by the remote driver.

- Instructions are sent on the downlink (from remote driver to the vehicle). Manoeuvre instructions should be acknowledged (uplink).
- Instructions should be provided reliably and at low latency. TCP or UDP can be used.
- Sporadic traffic, sometimes bursty, depending on manoeuvre instructions.
- Message sizes may be small, and the instruction stream may be low in bitrate.

Expectation on the system:

- Low Latency (i.e. Fast instruction execution and fast loss detection).
- Low loss rate (i.e. RAN is preferably handling error recovery).

The **video (camera) and audio (microphone) sensor** data (maybe also radar or lidar sensor data) may be very high in bitrate and very bursty. There may be one or more video cameras facing front. Additional rear video cameras may provide a backward view. Microphones are recording the external sound, e.g. for early and non-visual identification of emergency vehicles and detection of honking. Audio & video traffic flows are mostly unidirectional from vehicle to remote driver. Acknowledgements for error detection (e.g. for audio or video) or other downlink traffic (e.g. rate adaptation may not be needed) is either not present or at low bitrate.

Key traffic characteristics:

- Traffic is high bitrate (multiple cameras and microphones), unidirectional from the vehicle to the remote driver.
- Potentially no reception acknowledgement and therefore no error detection in media transport.
- Potentially no rate adaptation commands.
- Video traffic is typically very bursty.

Expectation on the system:

- Low Latency
- Low loss rate, bounded by latency (low latency is of higher importance than loss rate)

Potentially, audio and video should also be separated. The video flow is more important than the external microphone (audio) flow.

The **vehicle status sensor** (Acceleration, Speed, direction, Position, etc) data may be sent with a fixed interval, e.g. 10Hz or with a vehicle-dependent frequency. A sensor reading is typically only some few bytes. The status sensor data flow is unidirectional, from the vehicle to the remote driver. Acknowledgements, etc are FFS.

Key traffic characteristics:

- Moderate bitrate, unidirectional from the vehicle to the remote driver
- Reception acknowledgement for error detection and error recovery (some sensor streams might be essential).
- Constant bitrate and continuous traffic flow

Expectation on the system:

- Low Latency, potentially tactile requirements
- Low loss rate, bounded by latency

When a **conversational speech** channel exists to allow direct communication between the remote driver and vehicle passengers, the speech channel is a bi-directional channel, with same (symmetric) characteristics in both directions.

### 6.4.1.3 Potential requirements

The system may support high quality at low latency video in uplink direction. The re-use of the Framework for Live Uplink Video (FLUS) [17] can be considered.

The system may support decent quality at low latency audio in uplink direction. The re-use of the Framework for Live Uplink Video (FLUS) [17] can be considered.

The system may support a separate speech communication channel, which is separated from other, driving-related audio, video and data channels.

The system may support undefined quality at low latency sensor data in uplink direction, such as speed, acceleration, etc. Carrying appropriate sensor data formats can be studied e.g. for FLUS.

The system may use 3GPP defined QoS Framework e.g. as discussed in TR 26.939 [18]. Several QoS flows may be activated for the remote driving application.

The system may be rate adaptive so that e.g. the video quality can adjust to changing network conditions.

The system may be latency adaptive so that e.g. the video compression can be adapted to the latency requirement, for achieving a suitable coding latency vs. compression efficiency trade-off, depending on the current latency requirement.

The system may support low latency, low bitrate command distribution scheme to send vehicle control commands from the Remote driver to the vehicle.

## 6.4.2 Information sharing for high/full automated driving

This use case includes both the uplink and downlink, and may also include the sidelink in the sharing of information. Conventional media information from cameras can be compressed with typical video codecs, and other types of sensor information such as LIDAR output or detailed planned trajectory often has the form of multidimensional arrays that include a large amount of redundant information, which can easily be compacted with data compression techniques and controlled depending on the status of network and vehicles. However, fall of accuracy below acceptable levels and their impact on the safety needs to be avoided.

## 6.4.3 Video data sharing for assisted and improved automated driving (VaD)

This use case, often called the see-through [15], assumes the use of sidelink (PC5) interface whose capability for handling real-time video is under discussion.

NOTE 1: LTE PC5 cannot support 700 Mbps, which is required as [R.5.4-008] in [5] assuming a stream of raw video in RGB format with a resolution of 1280x720 at a frame rate of 30 Hz, in Rel-15.

NOTE 2: Sidelink V2X communication in Rel-14/15 LTE does not support channel establishment over PC5 and maintenance of one-to-one connection over PC5.

## 6.4.4 Teleoperated support (TeSo)

This use case may be realized as similarly as the remote driving use case.

## 6.4.5 Video composition

The uplink part of this use case may be realized as a form of network-based stitching, which can be enabled by FLUS, and the downlink part may be realized via point-to-point transmission or multicast/broadcast.

NOTE 3: For a proper composition, video streams in the uplink need to be aligned in both the temporal and spatial domains, and also captured in steady positions.

## 6.4.6 Data collection from in-vehicular sensors for HD map updates

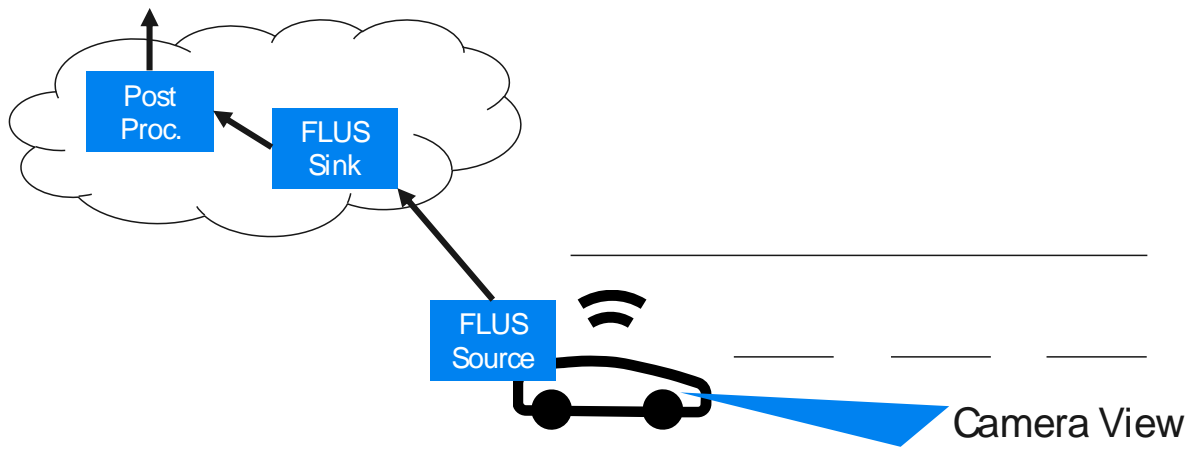
### 6.4.6.1 Description

TS 22.186 contains a use-case around 3D video composition for V2X scenario (TR 22.886, Clause 5.25). In other consortia (such as AECC [19], Section 3.2 and Table 1), this type of mechanisms is often named “High Definition Maps” (HD Map). The HD Map contains much more detailed information compared to a regular map. Regular maps are



used by today's navigation systems. HD Maps are used by new assistance and/or automated driving systems. The HD Map can contain sensor reading interpretations e.g. specific landmark descriptions to increase the positioning precision and is also used for obstacle detection (i.e. objects, which are not supposed to be on the road). Note, there are various different names for HD map like information and there is no standard defining the format.

In order to keep HD Maps up to date, information from the vehicle mounted sensors (such as video camera, scans from LIDAR) can be upstreamead or uploaded. Further, the video or LIDAR images from vehicles may be leveraged for other purposes than HD Map updates, for instance for road condition supervision and maintenance scheduling.



**Figure 6.4-2: Uplink video from vehicle-mounted cameras (and other sensors)**

The above depicted scenario sketch depicts the setup considering the usage of the Framework for Live Uplink Video (FLUS) [17], [18]. A vehicular mounted front-looking camera is capturing a video. The vehicle includes a FLUS source function and upstreams the video to the cloud deployed FLUS sink function. The FLUS sink forwards the stream to the post-processing function, which is then extracting relevant information for the HD Map.

### 6.4.6.2 Potential requirements

The system may support high quality at configurable latency video in uplink direction. The re-use of the Framework for Live Uplink Video (FLUS) [17] should be considered.

The system may support undefined quality at configurable latency sensor data in uplink direction, such as speed, acceleration, etc. Carrying appropriate sensor data formats should be studied e.g. for FLUS.

The system may use 3GPP defined QoS Framework e.g. as discussed in TR 26.939 [18].

The system may be rate adaptive so that e.g. the video quality can adjust to changing network conditions.

---

## 7 Protocols

### 7.1 General

V2X use cases require a diversity of media compression and transmission schemes that cannot be served by a single architecture of codecs and protocols. Depending on the nature of information or data to be handled, many existing protocols may be re-used or extended to meet the technical needs of these vehicular applications.

### 7.2 Uplink

As shown in figure 4.3 of [20], media codecs on top of protocol stacks, as in MTSI or FLUS, may be used to transport compressed media and sensor data to RSU or eNodeB with QoS established and maintained via IMS. Some use cases require smaller latencies than previously available, e.g., those supported by QCI 1 or 2. Depending on the use cases, other transport protocols may be used for higher performance.

## 7.3 Downlink

In the downlink requiring point-to-point transmission from RSU or eNodeB to vehicles, protocol architecture shown in figure 4.3 of [20] may be used. When there is a need for the vehicles to share the same information, MBMS protocol architecture shown in figure 9 of [21] can be considered if this can meet the end-to-end latency requirements of the service. Since the media information is likely to be consumed in a different fashion from conventional applications, some protocols may not have to be present, e.g., like those for content protection, when the traffic information needs to be shared by all vehicles in the cell.

## 7.4 Direct link

NOTE: use of PC5 link for video is not in the scope of this study item.

# 8 Media consideration

## 8.1 General

In some V2X use cases such as the information sharing for high/full automated driving, decoded media is not used for human perception but for the understanding of machines [13]. In these situations, conventional approaches for evaluating the handling of media, e.g., for communications, entertainment, or surveillance, may not be relevant.

## 8.2 Video

### 8.2.1 Resolution, Framerate, and Bitrate

#### 8.2.1.1 Resolution Aspects

There are several dashboard-mounted cameras on the market which support different wide-angle field of view (FoV). However, there is a dependency between the FoV, the video capture sensor resolution and the distance of objects of interest.

As example, we calculate here the capture size of a pixel (in meter). The larger the distance between the camera (capture sensor) and a captured object, the larger the quantization noise. In other word, the larger the distance between camera and the object, the larger the area which is captured by a single pixel.

In the following, we show some examples of the dependency.

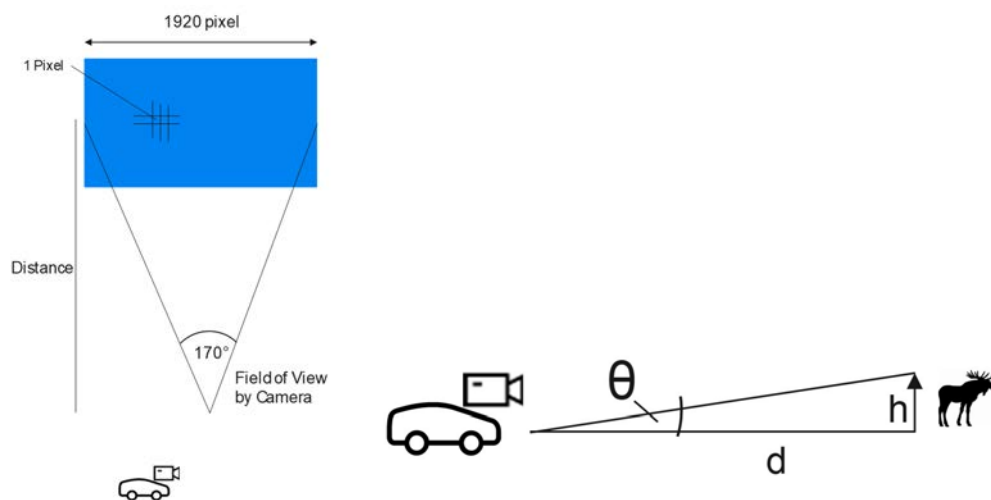


Figure 8.2-1: Determining the pixel size

$d$  is the distance between the vehicle mounted camera and the object.  $h$  is here the width of an object, which would be captured by a single pixel.

$$\theta = \frac{170^\circ}{1920}$$

$$h = \tan \theta * d$$

$$h = \tan \frac{170^\circ}{1920} * d$$

**Table 8.2-1: Pixel resolution based on distance, camera resolution, and field of view**

Row	Distance [m]	Camera Resolution (Horizontal)	Camera Field of View	Object width, captured by one pixel [m]
1	10	1920	170	0,015453448
2	50	1920	170	0,077267242
3	100	1920	170	0,154534484
4	10	4096	170	0,007243799
5	50	4096	170	0,036218997
6	100	4096	170	0,072437994
7	10	4096	40	0,001704423
8	50	4096	40	0,008522116
9	100	4096	40	0,017044231

The table shows a pixel resolution for different camera sensors, different distances and field of view. Row 3, for example, indicates the size of a captured object in a pixel for 100 m distance. An object of width 1.5 meter would be captured into 10 pixels. A blob of 10 pixel might be too few for an object detection. Or, the camera resolution / field of view should depend on the distance target of object detection.

The vehicle at 90km/h needs 4 seconds to drive a distance of 100 m.

The camera resolution and the camera field of view (lens) may depend on the use-case for the camera.

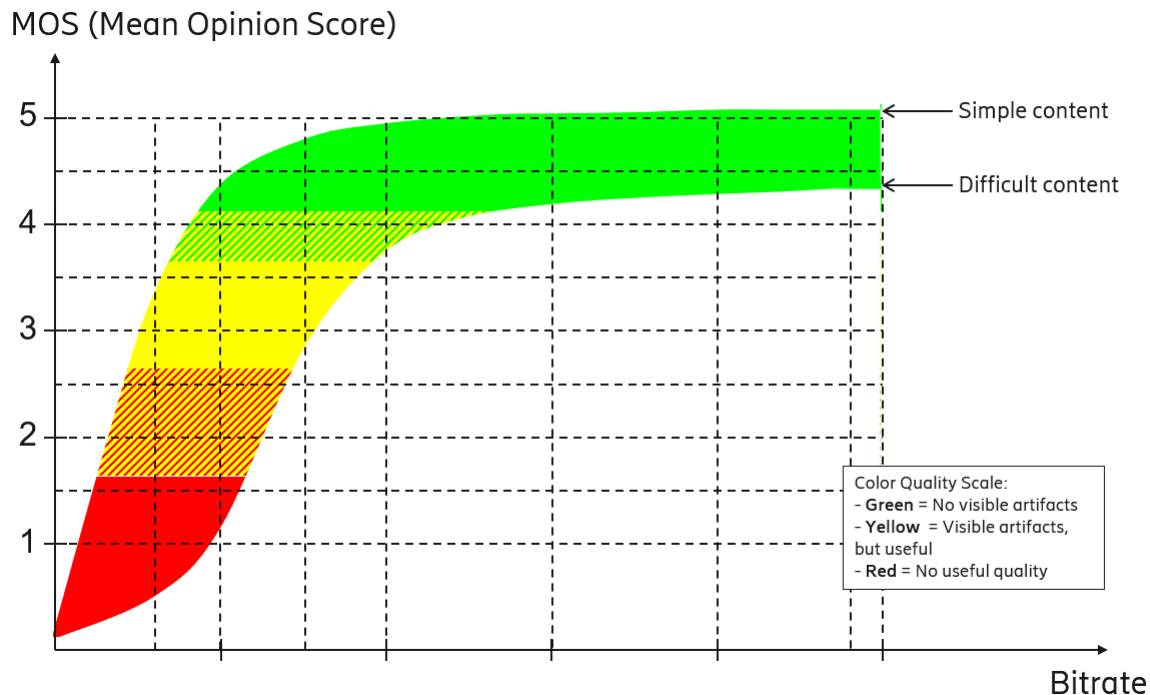
### 8.2.1.2 Frame Rate

Many available consumer cameras capture frames at 30 frames per second. Some commercial dash cams already support 60 frames per second to increase the sharpness of captured images.

When a 25 fps camera is mounted to a vehicle, which is driving 90 km/h, then the vehicle moves 1 m for each frame duration (90 km/h == 25 m/s).

### 8.2.1.3 Bitrates / quality

For consumer video, the dependency between bitrate and quality (MOS) has been well studied. The figure below is shown for illustration purposes only to demonstrate that the variation in quality and bitrate generally follows a certain trend. The needed bitrate for the same quality depends on the viewing distance, content, the codec used, the codec profile, and certain implementation features like look-ahead for motion estimation / compensation.



**Figure 8.2-2: Mean Opinion Score over bitrate**

Certain codec features have an impact on the end-to-end latency and may not be used.

Several use-cases such as remote driving or see-thru consider that the video from the vehicle mounted camera is viewed by a human. In cases where the video is viewed by a machine, the quality requirements are still to be studied.

## 8.2.2 Adaptation

### 8.2.2.1 General

Streaming of video puts extra requirements on the access as the bitrates are typically high and can easily exceed the available network capacity. Bitrates up to 12 Mbps are not uncommon for Full HD resolutions. Addition of several cameras for full 360 (several video screens or stitched for VR goggles) view can lead to bitrates up to 80 Mbps. NR and 5G deployment can in the future provide with very high access bitrate in some areas, while other areas have lower access bitrate, to further complicate things, the networks may be loaded differently, depending on time of day or other special events. The use of MBR/GBR bearers can provide certain guarantees that a GBR bitrate can be offered to a given V2X application and this can help to ensure a minimal acceptable media quality also when network load is high.

Despite this, it is necessary to make V2X applications that use video for e.g. tele-remote control of vehicles rate adaptive.

- 1) Delay: Connections (a.k.a. radio bearers) over a cellular LTE or NR connection can deliver a throughput that depends on a number of factors such as radio quality, that can be dependent on distance to the radio base station; and network load that depends on number of simultaneous users and their traffic mix. QoS can be used to prioritize one type of users over others, for instance a GBR (guaranteed bitrate) can be configured that gives priority for the amount of data per unit time fitting within the agreed GBR. If there is still spare capacity, data beyond that will also be transmitted but not prioritized. A guaranteed bitrate can be offered as long as there is available capacity in the network or there are other, less prioritized users to take capacity from. A V2X

application that tries to send at a higher rate than what is possible, will experience an increasing delay because packets will start to become queued up in the radio interface.

- 2) Packet loss: Packet loss occur for two reasons, they are either related to congestion or more of a stochastic nature. In LTE, the packet losses that occur are predominantly congestion related. Thus, packet losses occur mainly when queues are filled up too much.

Increased delay because of queuing should be avoided in tele-remote applications, because long delays can make the remote control sluggish, with decreased usability and potential increased security risks as a result. In the worst case a video playout may become choppy, slow down or even freeze completely. Packet losses are undesired, low levels of packet loss can be acceptable as it may only affect a small part of a reproduced video image for a limited time. Large levels of packet loss can give ghost images or potentially cause the video image to freeze completely for longer periods.

Another aspect of multi-camera video streaming in the presence of network congestion is that it can be desired that some cameras representing a more important field of view are more preserved, this means that less important cameras are sacrificed.

Congestion control for multimedia is in some respect a well-researched area. The MTSI based congestion control is described in TS 26.114. Lately, IETF chartered standardisation work for multimedia congestion control in the RMCAT working group.

SCReAM (RFC8298) [24], developed under the RMCAT charter in IETF, was initially devised for WebRTC but has been refined for tele-remote operation of industrial vehicles. For instance, SCReAM congestion control is being evaluated in the EU SIMS project [22] where SCReAM congestion control of multicamera video over LTE is demonstrated in a setup with four or six commercially available IP cameras. SCReAM was from the very beginning devised for good operation over cellular links such as LTE. The function in SCReAM and its applicability to tele-remote operation is described in [23]. SCReAM is, besides being sensitive to delay changes, also sensitive to packet loss. Furthermore, SCReAM has built-in ECN support.

One interesting feature with SCReAM is that it has built it stream prioritization, this makes it possible to prioritize e.g. one camera stream when the available network bandwidth becomes critically low. This can for instance be used in a tele-remote setup where a front camera is deemed more important than the other cameras. This built-in priority mechanism in SCReAM can be used instead of a network based QoS differentiation mechanism for instance in cases where a network based QoS differentiation mechanism is not practical to realize.

### 8.2.2.2 Level of Adaptation Control

In many use cases including video, such as the support of remote driving or the update of HD maps, the control server or the network may have to exercise a precise control over the operation of video encoders in the RSUs or vehicles. For example, key parameters such as the bit-rate or frame rate may have to be modified to new values within time limits and also within bounded ranges while maintaining certain quality targets.

In conversational services (e.g. TS 26.114 [20]), the UE or the conference server can control the operation of voice/audio encoders to a higher level than the video encoders, capable of changing bit-rate or audio bandwidth within a time window of a few frames, via in-band signalling such as CMR. On the other hand, the video encoders can be controlled via out-of-band, codec-independent signalling such as TMMBR. Contemporary video codecs such as H.264 [29] or H.265 [30] do not provide in-band signalling mechanisms in their Network Adaptation Layers (NAL), which play the role of RTP payload header that includes CMR in the 3GPP voice/audio codecs.

### 8.2.3 Operational Variance

In use cases such as the composition of video, performance of video encoders in the RSUs or vehicles need to be similar, to enable a consistent quality of video over entire directions. However, as typical standards for video codecs define only the structure of decoder and bit-stream, performance of video encoders are left to the discretion of implementation. As a result, even at the same profile and level, performance of video encoders may vary widely, generating at similar bit-rates video bit-streams with a diversity of quality, depending on factors such as the complexity or contents. This variance may limit the performance of V2X systems.

### 8.2.4 Video Configuration Requirements

The requirements in Table 8.2-2 were obtained from consultation with the 5GAA.

Table 8.2-2: Video Codec Requirements from 5GAA

Use case	resolution	Field of view	FPS	Data rates	Latency	Human/machine
High Definition Sensor Sharing – AV lane change		360				Machine
See-Through for Pass Manoeuvre	1280x720, colour depth 8 bit, 24 bit resolution, subsampling 4:2:2		30	15 Mbps NOTE 1	50 ms NOTE 2	Human
Tele-operated driving		4 cameras NOTE 3		15 Mbps per camera		Machine or human
Obstructed view assist	1280x720, colour depth 8 bit, 24 bit resolution, subsampling 4:2:2		30	15 Mbps NOTE 1		Human
Infrastructure Assisted Environment Perception				40 – 120 Mbps per camera		Machine or human

NOTE 1: 15 Mbps assumes video compression.

NOTE 2: Video delay caused by latency must be within the tolerance of the human driver. For example, a vehicle at 90 km/h moves 1.25 m within 50 ms.

NOTE 3: It is still to be determined whether the four cameras are producing a 360 video with same quality in all four directions, using two cameras facing front (e.g., with FoW of 120° and 40°, and same or different quality). It is also to be determined at what video quality/bitrate below which the service and vehicle need to stop.

---

## 9 Sensor sharing of object information

### 9.1 General

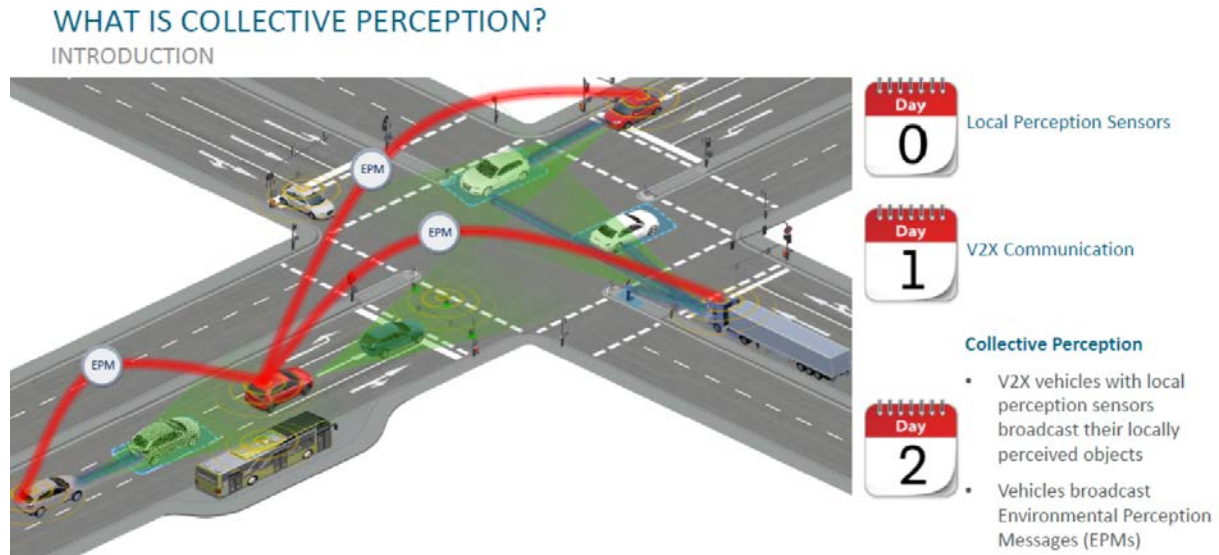
An alternative to sharing video feeds from vehicular sensors is to use processing at or near the sensor to classify and identify objects of interest and then transmit the object information.

### 9.2 On-going work

Other standard organizations have been working on sensor-based classification and identification of objects.

## 9.2.1 ETSI-ITS

Due to the high data rate and transmission frequency requirements of sending raw sensor data, ETSI-ITS has been working on object-based solutions. It has developed a Common Data Dictionary (CDD) in TS 102 894 [23] and a specification TS 103 324 [24] “Collective Perception (CP) Service” which specifies how an ITS station can inform other ITS stations about the position, dynamics and attributes of detected neighbouring road users and other objects. The CP service shares information with other ITS stations through the transmission of Collective Perception Messages (CPMs) and Environmental Perception Messages (EPMs) as illustrated in figure 9.1-1.



**Figure 9.1-1: Collective perception**

ETSI-ITS also developed a list of recommended requirements for the Collective Perception Service in TR 103 562 [25]. This identifies relevant issues for object-based sensor sharing such as:

- Conditions that generate CPMs with object information
- The frequency at which CPMs should be sent
- CPM object quality assessment
- CP message formats and data elements

## 9.2.2 5GAA

5GAA completed a work item in November 2017 to produce 5GAA SENSHA TR, A-170272 which investigates the following objectives:

- Identify sensor data sharing requirements from envisioned automated driving applications, including architecture options, requirements and implications
- Perform a gap analysis from ETSI ITS standards (at a minimum) regarding existing data objects
  - a. Identify what new objects and extensions of the existing objects are needed
- Perform the analysis of the existing interface messaging protocols to determine their suitability
- Perform the analysis of the security and privacy aspects

## 9.2.3 SAE

SAE currently has several related WIPs: J2945/8 on sensor sharing which is pending in the Vehicular Applications TC and the other on trajectory sharing which is pending in the C-V2X TC. It is suggested that 3GPP establish a liaison relationship with SAE so that we can coordinate the progress of the work across these TCs and 3GPP.

## 9.3 Example objects

Table 9.1 lists examples of the classification for which descriptions and parameters are being specified in the relevant standards bodies.

**Table 9.1: Example objects**

Object	Examples/Description
1D Object	Traffic lanes and road boundary
2D Object	Traffic sign, warning sign, guide signs, regulatory signs
3D Object	Pedestrian, cyclist, moped, motorcycle, passengerCar, bus, lightTruck, heavyTruck, trailer, specialVehicles, tram, roadSideUnit, traffic cones, "unknown point cloud", Object 3-axis distance, Object 3-axis speed, Object 3-axis acceleration, Object dimensions

## 9.4 Considerations

### 9.4.1 Liability

One of the concerns sometimes raised by automobile manufacturers about the object-based approach is the reliability of the information. There is a concern about liability in case there is a failure and the object information is generated and sent by a sensor on another vendor's vehicle.

One solution to this concern is to set performance and conformance requirements on the object identifier algorithms. ETSI-ITS and SAE include confidence values for some reported parameters (e.g., HeadingConfidence, SpeedConfidence). These are not part of an industry performance standard but do provide a measure of the accuracy of the reported data and might be leveraged for developing performance and conformance standards.

Relying purely on sensor video transmissions doesn't fully address the liability concerns as these sensors and their encoding algorithms (e.g., compressed video) would also need to meet some performance and conformance standards for them to be considered reliable. Currently there are no video encoder performance or conformance standards.

### 9.4.2 Multi-mode media

Aside from the significant savings in transmission bandwidth requirements, the object-based approach allows incorporating non-video sensor information into the object classification and identification (e.g., location using LIDAR) to provide richer and more completed information than a raw video feed.

### 9.4.3 Message frequency

For the primary use cases, the frequency of message transmission has a predicted periodicity of about 100 ms. However, to accommodate for event-driven messages, transmission intervals between 50-100 ms are also being considered. More advanced use cases can further increase the data exchange rate, e.g.:

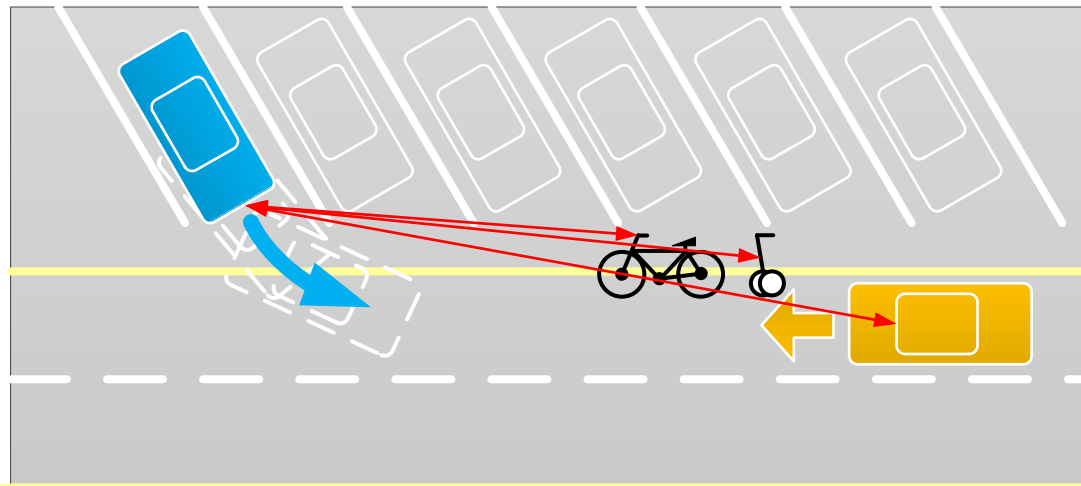
- Planned trajectory
- LIDAR
- Dynamic 3D local map sharing (e.g. 3D road model built based on LIDAR)



### 9.4.4 Device Considerations

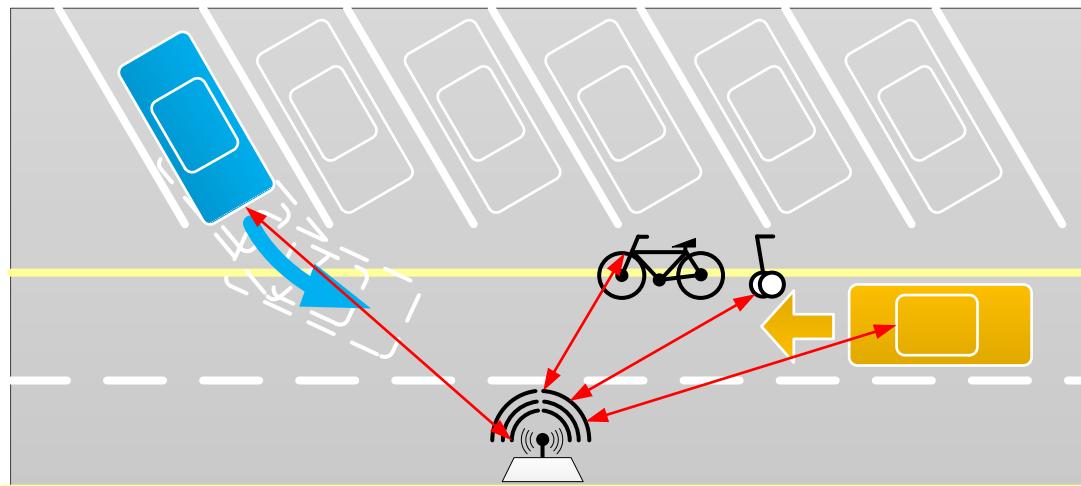
For C-V2X safety use cases there can be communication with devices used by Vulnerable Road Users (VRUs).

Figure 9.3-1 illustrates a scenario where the parked blue vehicle is planning to back out of a parking spot and sends an alert to nearby road users. The recipients of this alert can either take evasive action or notify the backing vehicle that they are in the vicinity, causing the vehicle to wait before backing out.



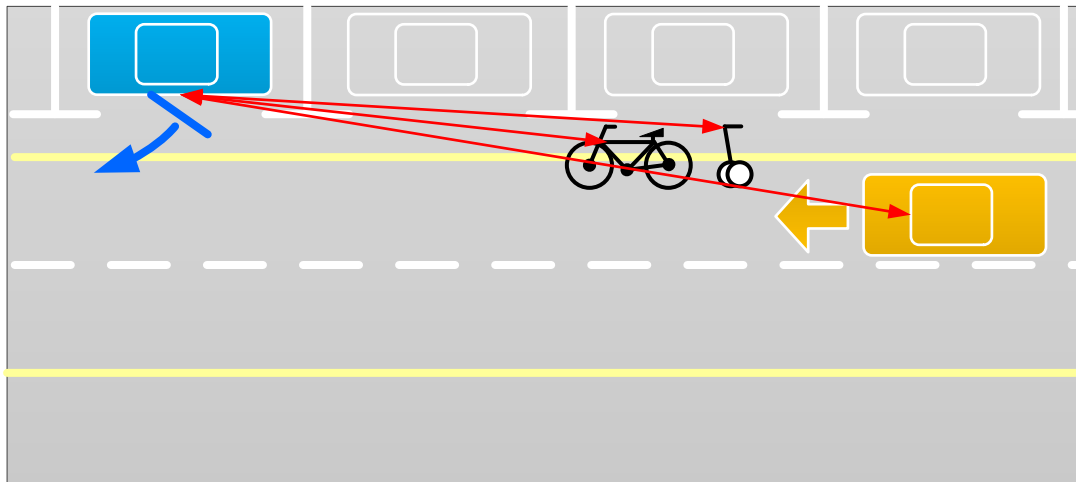
**Figure 9.3-1 Parking Spot Exit: V2V and V2VRU alert of neighboring road users**

Figure 9.3-2 illustrates a similar scenario where an RSU detects that the blue parked car is starting to back out of a parking spot. The RSU sends an alert to the nearby road users which then take evasive action. Alternatively, the RSU, having also detected the presence of the nearby road users, can alert the vehicle backing out so that it can wait for the road to clear.



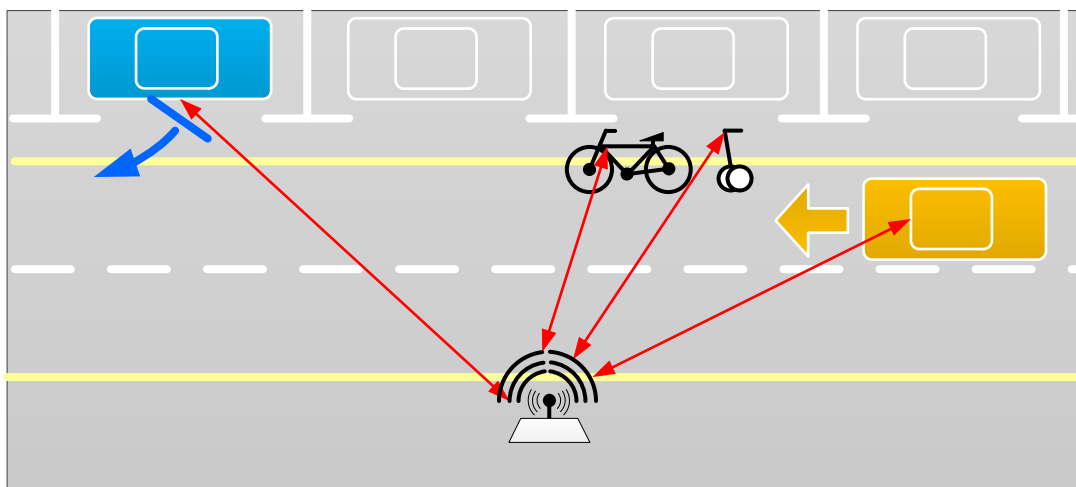
**Figure 9.3-2 Parking Spot Exit: V2I (V2 RSU) alert of neighboring road users**

Figure 9.3-3 illustrates a scenario where the parked blue vehicle is planning to open the side door and sends an alert to nearby road users. The recipients of this alert can either take evasive action or notify the vehicle that they are in the vicinity, causing the vehicle to wait before opening the door.



**Figure 9.3-3 Door Opening: V2V and V2VRU alert of neighboring road users**

Figure 9.3-4 illustrates a similar scenario where an RSU detects that the blue parked car is planning to open its door. The RSU sends an alert to the nearby road users which then take evasive action. Alternatively, the RSU, having also detected the presence of the nearby road users, can alert the vehicle so that it can wait for the road to clear before opening the door.



**Figure 9.3-4 Door Opening: V2I (V2 RSU) alert of neighboring road users**

In each of the above use cases, the alert could be a live video stream and/or an object identifier. Since some of the devices receiving and processing the alert can have smaller battery, processing, and form factors, such as a wearable or bicycle computer, they cannot all be expected to decode and process video streams to determine the nature of the alert. Therefore, while it may be useful to send video streams to some devices, there should always be a baseline form of communication, such as an object identifier, to guarantee that all devices can properly decode the essential information to take appropriate actions, e.g. alert user.

## 9.5 Conclusion

Object-based approaches can serve as alternative solutions for the FS\_mV2X use cases. These approaches, which are being standardized in other SDOs, use lower bitrates than sending video media, which is useful in scenarios that have bitrate limitations, and are necessary when communicating with smaller VRU devices such as wearables. Still, there may be use cases where sending video media is preferred and can be used if the necessary throughput is supported.

It is important for SA4 to understand these use cases and investigate all potential solutions, including the setting of conformance/performance standards for sensors and their processing in 3GPP or other standards bodies. SA4 should also study which approaches are appropriate for the different use cases, including the possibility of combining the two approaches, i.e., sending video media with object data.

To support these studies and future collaboration with other SDOs in the relevant areas it is recommended that 3GPP establish an agreement to exchange information SAE. A liaison relationship is likely to be insufficient for this as SAE does not readily share its information with other SDOs. This may require 3GPP to establish a collaboration agreement with SAE to allow for a more open exchange of information.

---

## 10 Conclusion

In this study, a set of use cases was proposed and adopted from TR 22.886, in which media was involved in some fashion in the operation of vehicular services. These include the support for remote driving, teleoperation, information sharing for high/full automated driving, video composition, and data collection from in-vehicular sensors for HD map updates.

Based on the use cases, user plane protocol architectures and signaling procedures for the UE, the eNodeB, and the road side unit (RSU) were studied, and the requirements and procedures for media capturing, compression, and transmission (uplink and downlink) using the envisioned transmission opportunities of V2N and V2I scenarios were outlined. It was found that some of the use cases require a latency lower than even the conversational services such as MTSI. It was clarified that some of the Rel-15 LTE radio link interfaces could not support channel establishment, the maintenance of a one-to-one connection, and the bit-rate required for uncompressed HD video. Therefore, use cases requiring direct exchange of uncompressed HD video media between vehicles are not realizable in Rel-15.

For Automated Vehicle (AV) applications, object information is always needed as a baseline to guarantee operation (nlos, night) and interop with low-end devices receiving the information. Video media can be used as a supplement. For human-viewed applications, there are use cases where video media may be necessary. For some of the use cases identified, non-media based approaches (e.g., object-recognition based) were also studied and future work in this area may continue based on further input from the automotive industry.

---

## Annex A: Change history

Change history							
Date	Meeting	TDoc	CR	Rev	Cat	Subject/Comment	New version
12-2019	SA#86	SP-190994				Approved at TSG SA#86	16.0.0
2022-04	-	-	-	-	-	Update to Rel-17 version (MCC)	17.0.0
2024-03	-	-	-	-	-	Update to Rel-18 version (MCC)	18.0.0

---

# History

<b>Document history</b>		
V18.0.0	May 2024	Publication