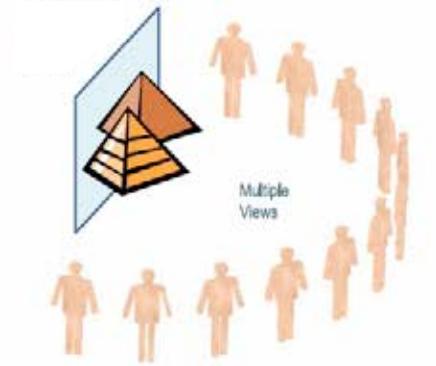
FVV Live: A real-time free-viewpoint video system

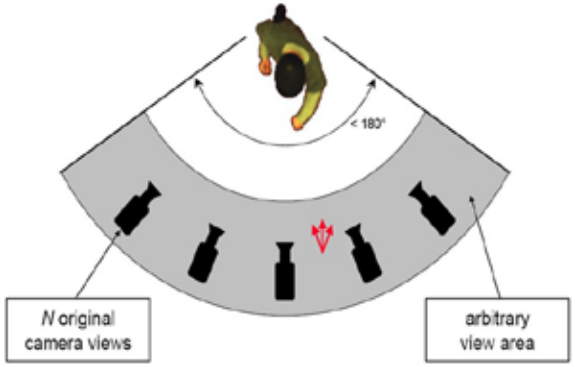
Daniel Corregidor, Pablo Carballeira, Julián Cabrera, César Díaz, Daniel Berjón, Francisco Morán, and **Narciso García**

narciso@gti.ssr.upm.es

VQEG London 2021

Multiview → Free-Viewpoint Video (FVV)



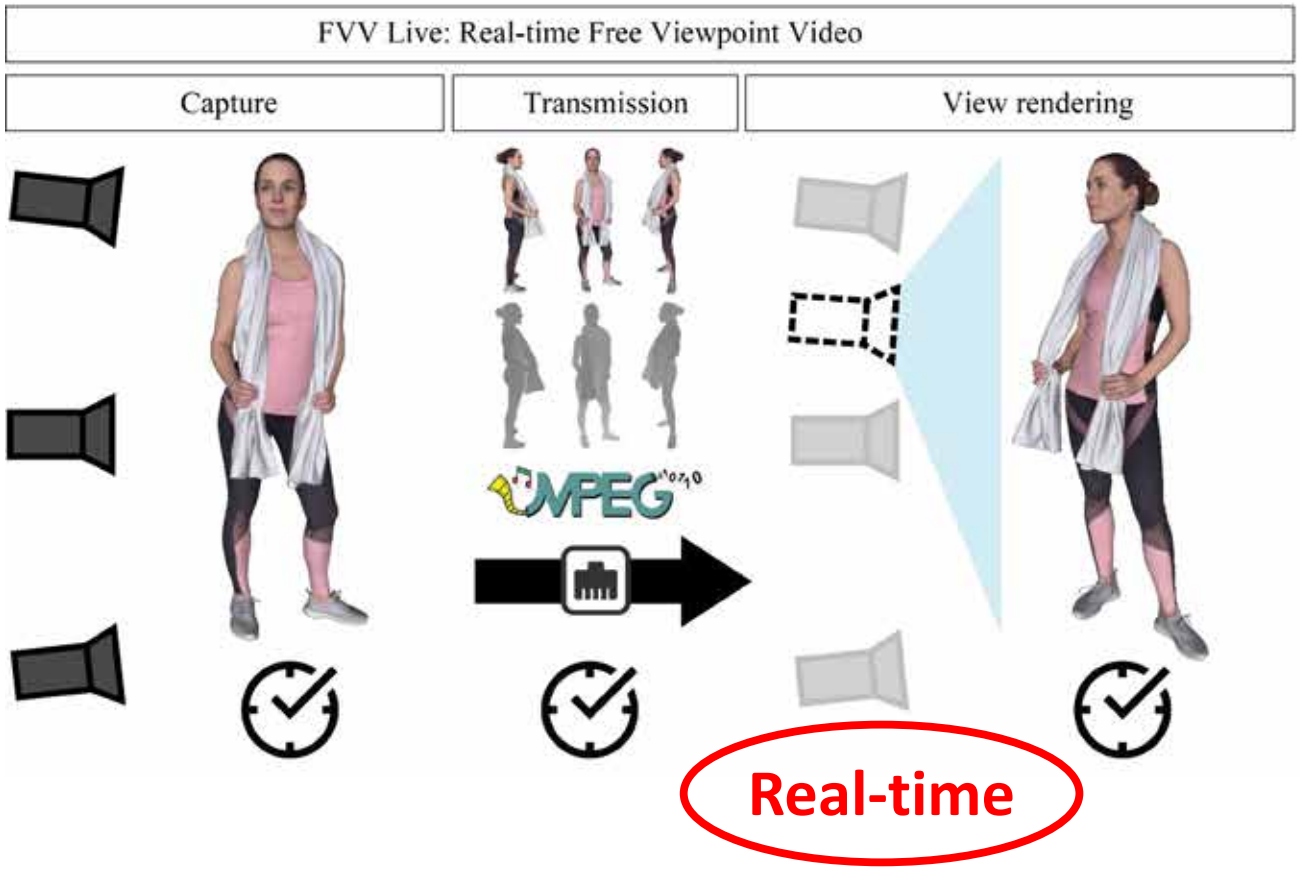


FVV Live

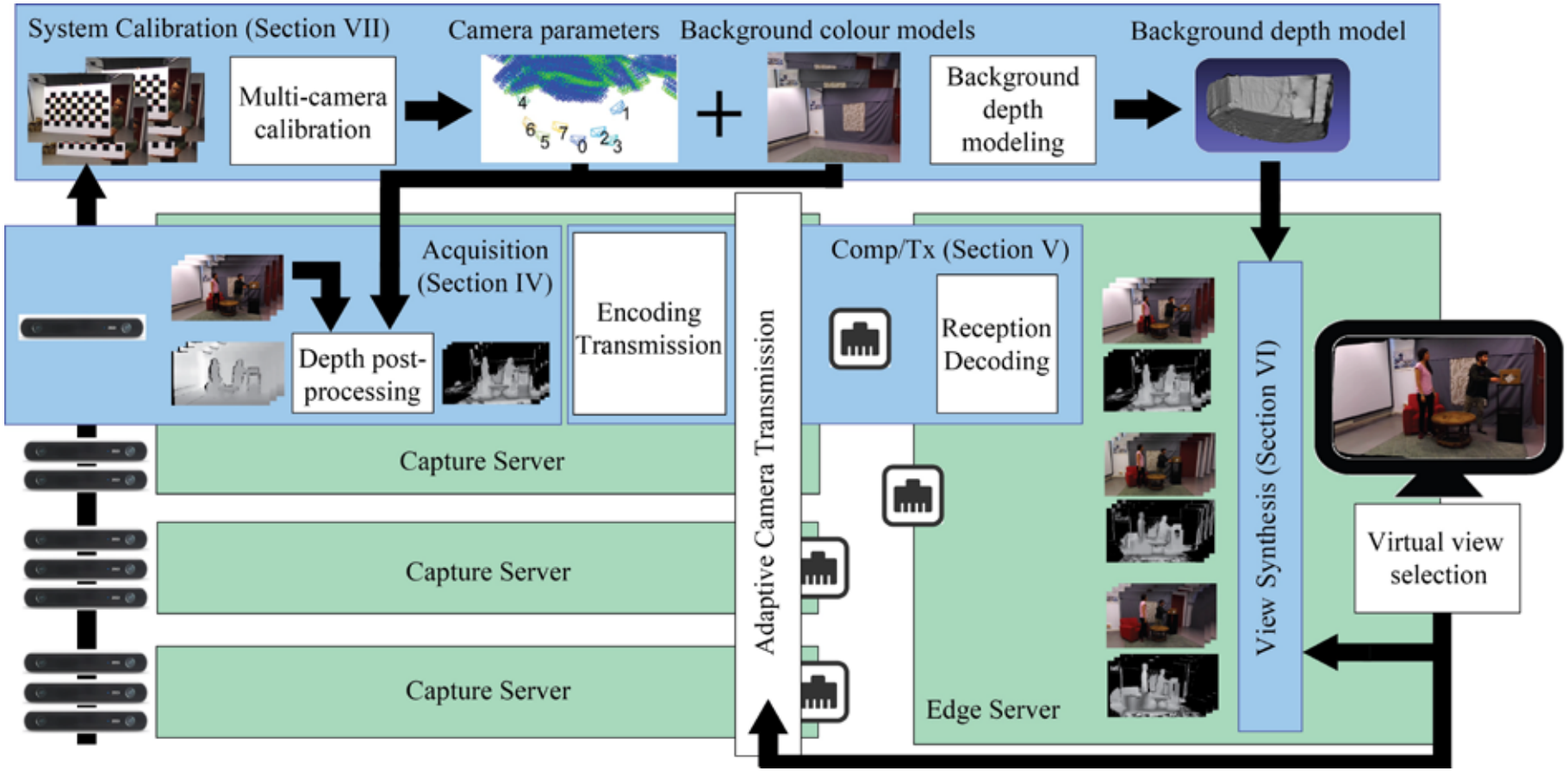
- FVV Live is a novel end-to-end free-viewpoint video system, designed for real-time operation, using consumer-grade cameras and hardware, which enables low deployment costs and easy installation for immersive event-broadcasting or videoconferencing.
- System design to maximize perceptual video quality.

- FVV Live is able to operate in real time at 1920x1080p resolution at 30 fps using off-the-self hardware (cameras and processing hardware), with an end-to-end delay of around 360 ms, and a mean motion-to-photon latency around 155 ms.
- These low delay values allow a natural navigation of the scene, and bilateral immersive communications.

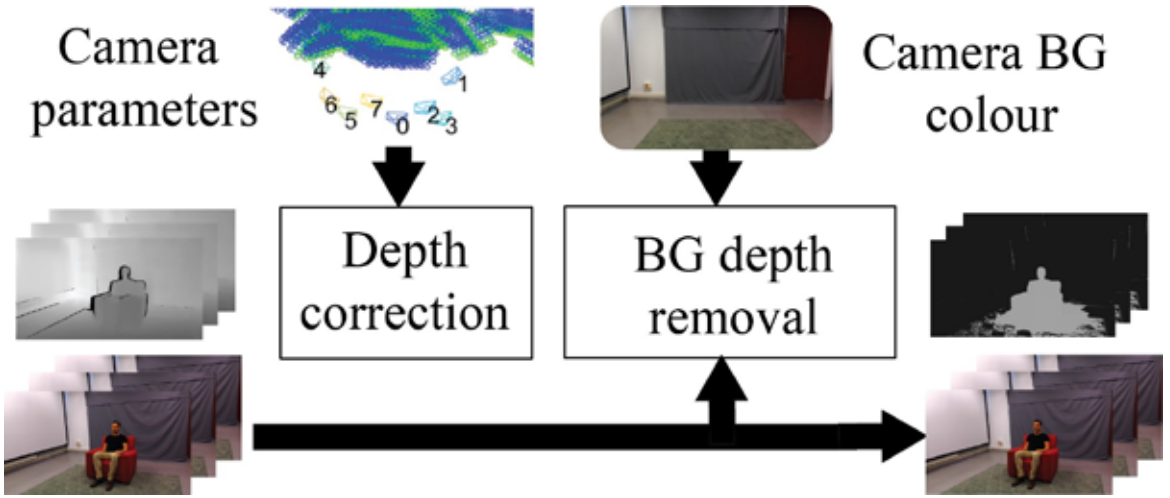
Concept of the FVV Live system



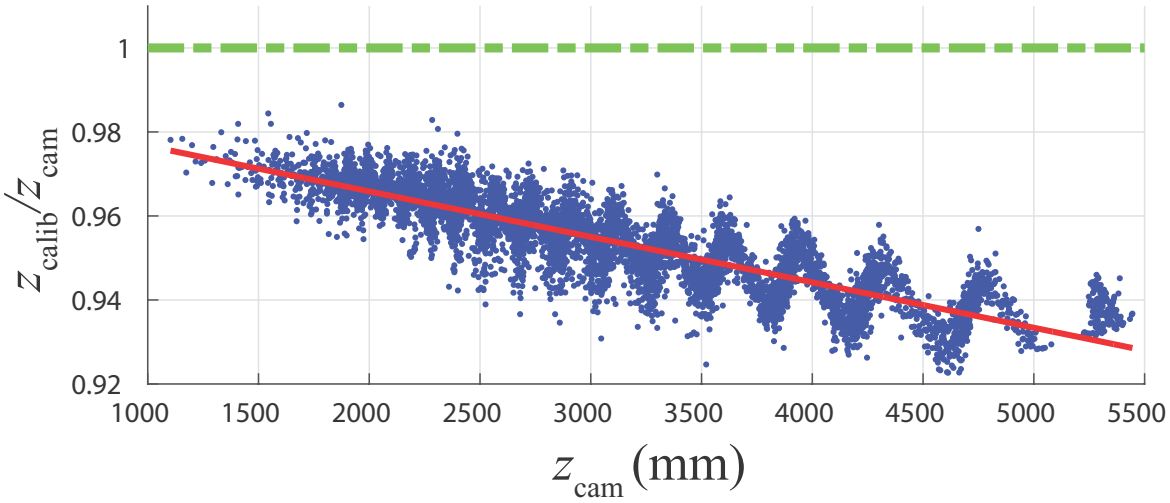
Block scheme of the FVV Live system



Scheme of the acquisition block for one camera

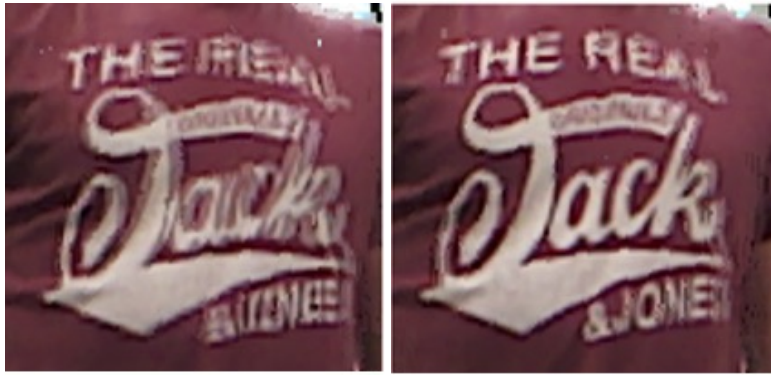


Depth error model



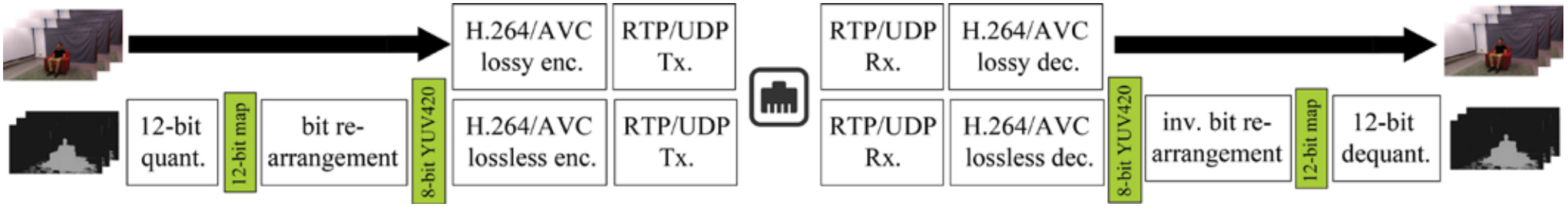
The green dashed line indicates the ideal relation between both estimates in the absence of error.
The red line models the systematic error due to a stereo pair calibration error.

Depth correction block

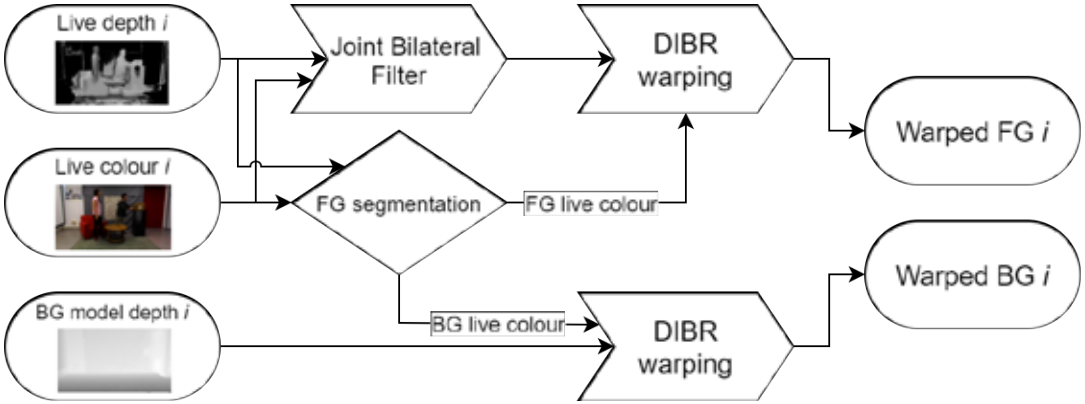


Example of double-image artifacts in virtual views corrected by the depth correction block.
Left: detail of virtual views obtained from uncorrected depth data.
Right: detail of virtual views obtained from corrected depth data.

Compression and transmision of colour and depth streams

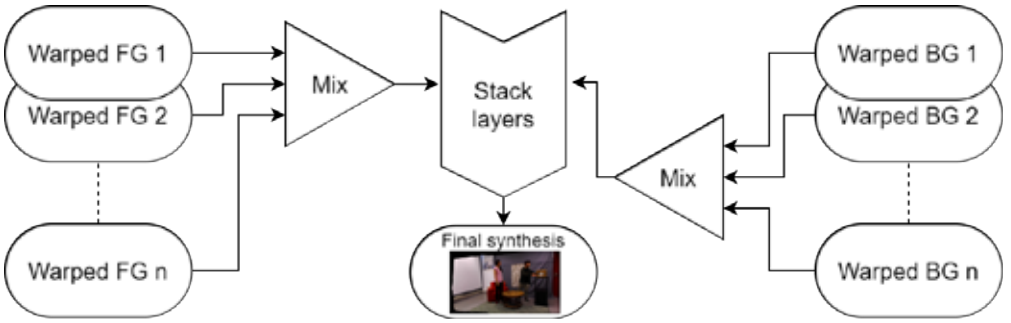


View synthesis: FG and BG processing



Per-camera contributions

Virtual view composition



Implementation details

GPU load tests in capture servers
Only viable configuration in bold

Camera-GPU configuration	Max. Avg. Frame processing time (ms)
2@Depth GPU + 1@Encoding GPU	**19.47**
3@Depth GPU + 1@Encoding GPU	39.09
2@Depth GPU + 2@Encoding GPU	49.43

Parameter	Colour stream	Depth stream
Compression	Lossy AVC	Lossless AVC
Rate control	one-pass VBR	one-pass VBR
Target bitrate	5-15 Mbps	-
GOP	IPPPP	IPPPP
IDR period	30 frames	30 frames

Coding parameters for colour and depth setreams

Empirical depth bit-rates of BG pixels

Scenario		Simple	Medium	Complex
Mean ratio of BG pixels (%)		80	75	57
Depth bit rate (Mbps)	w/o. BG depth rem..	96-112	92-114	99-121
	w. BG depth rem.	34-74	43-79	79-99

End-to-end and motion-to-photon delays

- t_{CS-ES} = 252 ms → time span between the capture of a set of frames and the synthesis of the corresponding virtual frame at the edge server.
- t_{ES-mob} = 109 ms → time span between the encoding of a synthetized frame and its display on the smartphone screen.
- Thus, below the one-way delay threshold of 400 ms for bidirectional video comms.

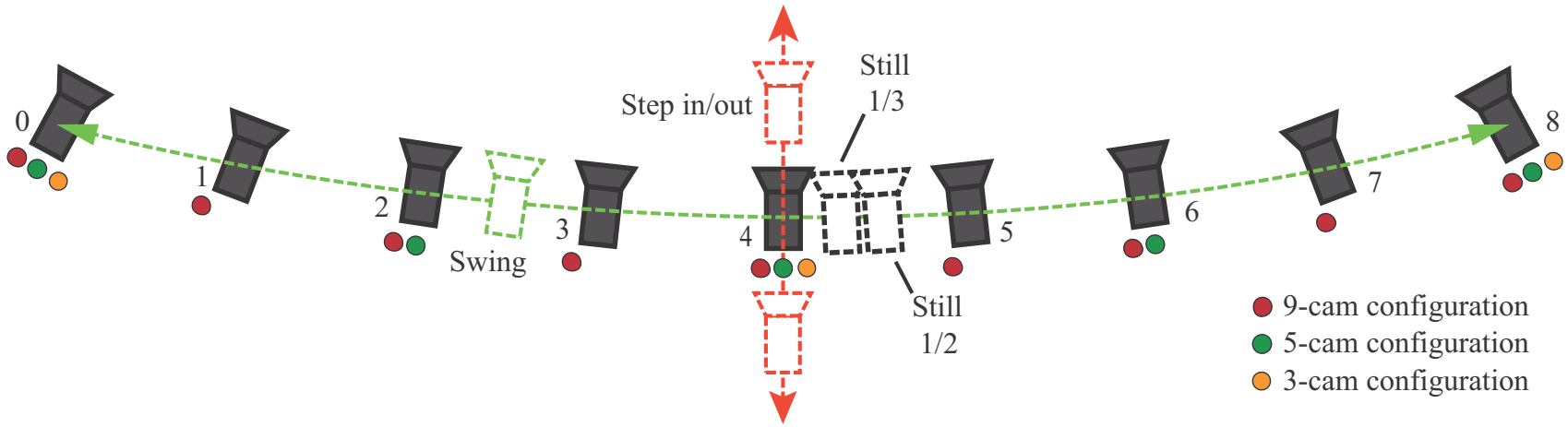
Use of FVV Live for immersive videoconferences.

- The motion-to-photon (MTP) delay (time span between a viewpoint update command and its result) is in the 30-63 ms range (mean value 47 ms) for edge server control, and 139-172 ms range (mean value 156 ms) for smartphone control.

Responsive virtual view navigation in both cases.

FVV Live real-time low-delay demo

Assessment of the Visual Quality of FVV Live



- See “Considerations on FTV quality assessment”, VQEG_IMG_2018_140, Mountain View, Mountain View, USA, Nov. 2018.

Additional information

- Awards:
  - 2019 IET's Vision and Imaging Technology Award (First Prize in the General Category) to the best technology in the field of Vision and Imaging.
  - 2020 IEEE International Conference on Multimedia and Expo ICME 2020 Best Demo Award.

- FVV Live paper published in the IEEE Transactions on Multimedia (Early access):
  - https://ieeexplore.ieee.org/document/9431676
  - https://doi.org/10.1109/TMM.2021.3079711

FVV Live: A real-time free-viewpoint video system

Questions – Discussion – Debate - …

Grupo de Tratamiento de Imágenes (GTI)
Universidad Politécnica de Madrid

This work has been partially supported by the Ministerio de Ciencia, Innovación y Universidades (AEI/FEDER) of the Spanish Government under project TEC2016-75981 (IVME)