A Realistic 3D Smorgasbord

Light Field Data Compression Methods and Quality of Experience for Augmented Telepresence

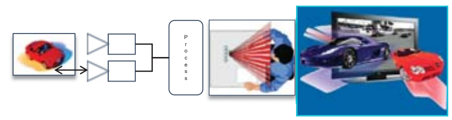
Prof. Mårten Sjöström, Realistic 3D Research Group,
Mid Sweden University

Image source: http://www.forsvenskad.se

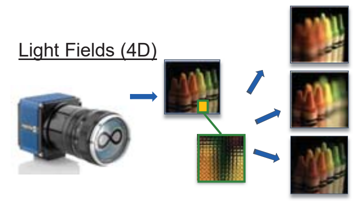
Realistic 3D Research Group – Four Research Foci

Our Smorgasbord is a mixture of technologies to represent a multi-dimensional parameter space in 3D geometrical space

Video supplemented by depth (3D)



Light Fields (4D)

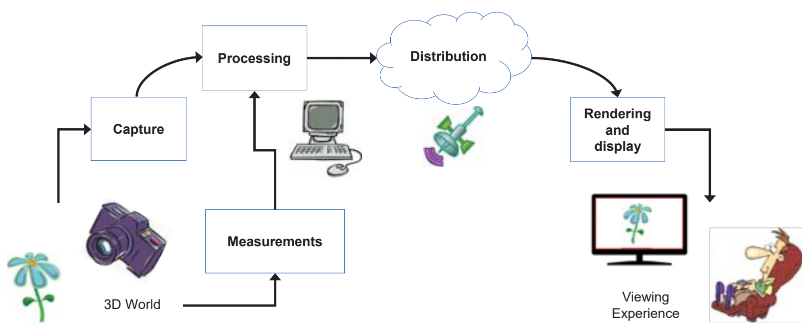


Visualization (nD → 2D)

Augmented Telepresence (AR & VR)

Imaging - from Capture to Presentation – a holistic approach

The whole processing chain influences the quality of the results



Today I Present You Two Delicacies:

Light Field Data Compression

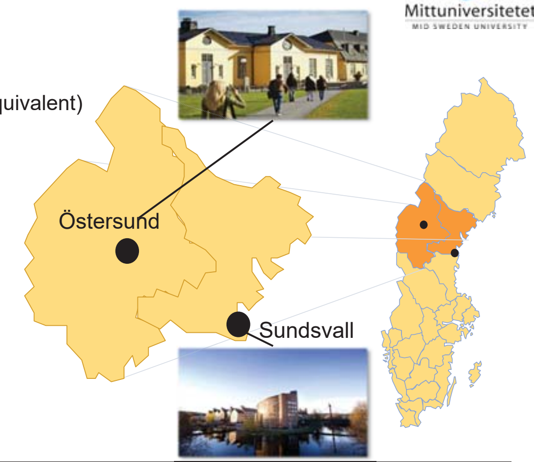
Quality of Experience for Augmented Telepresence

https://www.acreo.se/Master-thesis-Video-Quality-Experience

Mid Sweden University

Mid Sweden University

- ~ 13 600 students (7000 full-time equivalent)
- ~ 200 PhD students
- ~ 1000 employees
- ~ 90 professors
- Turnover SEK 928 Millions

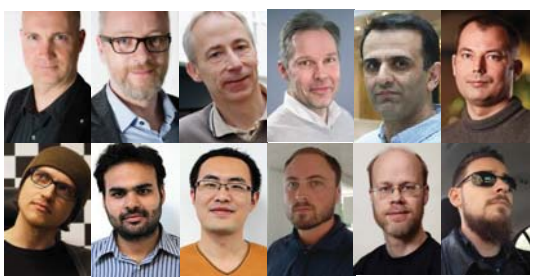


Multi-Dimensional Signal Processing and Imaging

Realistic 3D Research Group

Founded 2007

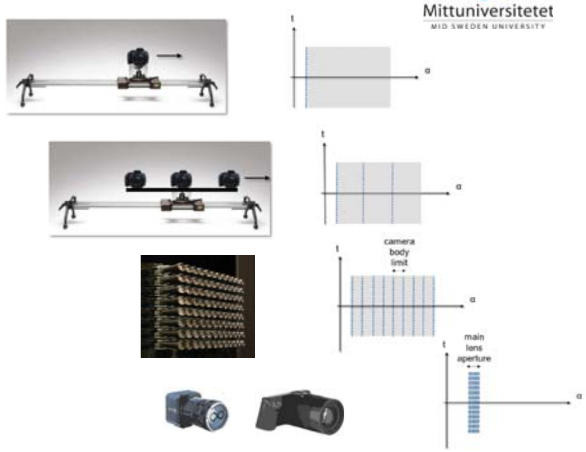
- 5 Senior Researchers
- 4 PhD students
- 3 Junior Lecturers (Adjunkter)
- 10 Alumni



Light Field Data Compression

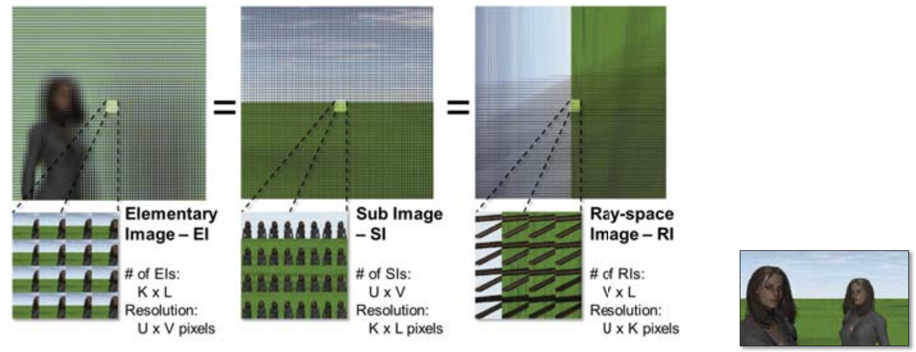
Data Structure and Sampling

- Moving camera
  - Requires a static scene
- Moving cameras
  - Requires a static scene
- Camera arrays
  - Dense spatial, Sparse angular
- Plenoptic camera
  - Sparse spatial, Dense angular
  - Type I and Type II



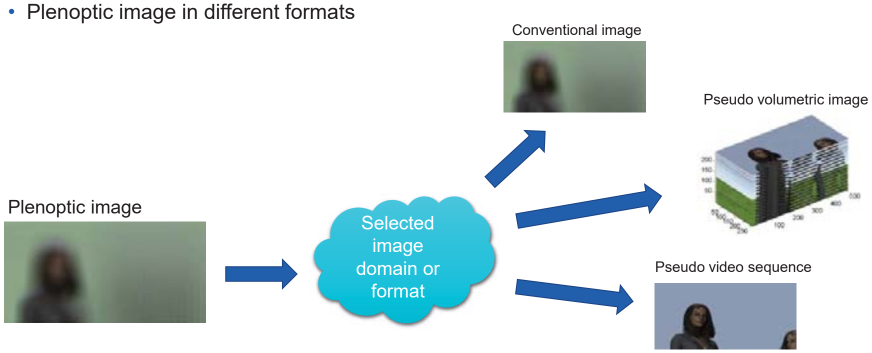
Data Format

- Plenoptic image in different domains



Data Format

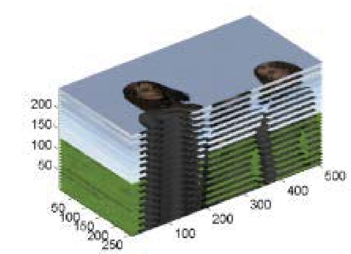
- Plenoptic image in different formats



3D Image Compression

Pseudo volumetric image

- Employ 3D image encoder
- Uses correlation in 3 dimensions
- 3D DCT
- JPEG 2000 3D
- Works well for high bit rates
  - Depending on contents

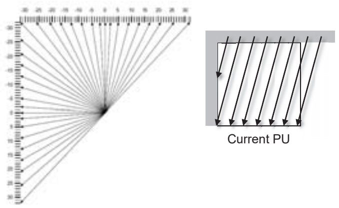


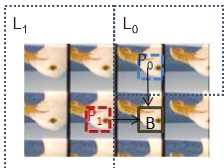
Aggoun A. (2006), A 3D DCT compression algorithm for omnidirectional integral images, In Proc. 2006 IEEE ICASSP, vol. 2, p. 517–520.

Olsson, R. (2008). Empirical rate-distortion analysis of JPEG 2000 3D and H.264/AVC coded integral imaging based 3D-images. In 2008 3DTV Conference.. pp. 93–96.

Image Compression – Intra Predictive Coding

- Intra prediction further enhances compression efficiency
  - HEVC intra coding achieves better RD-performance than JPEG 2000 for 2D images
- Multi-hypothesis prediction
  - Works both for Plenoptic camera type I & II





Conti C., Nunes P., and Soares L. D. (2012), New HEVC prediction modes for 3D holoscopic video coding," In Proceedings of IEEE International Conference on Image Processing (ICIP)

Li, Y. , Sjöström, M. , Olsson, R. & Jennehag, U. (2014). Efficient Intra Prediction Scheme For Light Field Image Compression. In Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

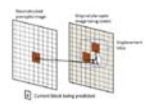
Li, Y. , Sjöström, M. , Olsson, R. & Jennehag, U. (2016). Coding of focused plenoptic contents by displacement intra prediction. IEEE Trans. Circ Sys for Video Tech., vol. 26: 7, pp. 1308-1319.

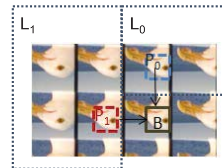
Li, Y. , Olsson, R. & Sjöström, M. (2016). Compression of Unfocused Plenoptic Images using a Displacement Intra prediction. In Proceeding IEEE International Conference on Multimedia & Expo Workshop (ICMEW)

Image Compression – Intra Predictive Coding

- Intra prediction further enhances compression efficiency
  - HEVC intra coding achieves better RD-performance than JPEG 2000 for 2D images
- Multi-hypothesis prediction
  - Works both for Plenoptic camera type I & II
- Also works for LF Video


HEVC intra Bi-directional prediction





Conti C., Nunes P., and Soares L. D. (2012), New HEVC prediction modes for 3D holoscopic video coding," In Proceedings of IEEE International Conference on Image Processing (ICIP)

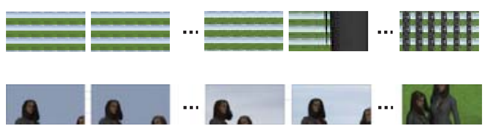
Li, Y. , Sjöström, M. , Olsson, R. & Jennehag, U. (2014). Efficient Intra Prediction Scheme For Light Field Image Compression. In Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

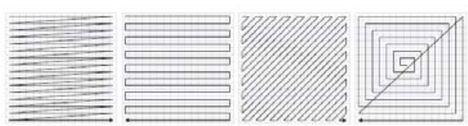
Li, Y. , Sjöström, M. , Olsson, R. & Jennehag, U. (2016). Coding of focused plenoptic contents by displacement intra prediction. IEEE Trans. Circ Sys for Video Tech., vol. 26: 7, pp. 1308-1319.

Li, Y. , Olsson, R. & Sjöström, M. (2016). Compression of Unfocused Plenoptic Images using a Displacement Intra prediction. In Proceeding IEEE International Conference on Multimedia & Expo Workshop (ICMEW)

Pseudo Video Sequence Compression

- Choose domain
  - 2D slices from 4D Light Field
  - Domain of normal images works best
- Organize series of images into a video
  - Decide image order for best use of codec
  - Circular order best
    - High correlation, high SNR

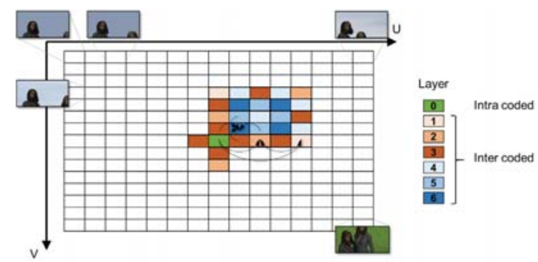




Olsson, R. , Sjöström, M. & Xu, Y. (2006). A combined pre-processing and H.264-compression scheme for 3D integral images. In *Proceedings International Conference on Image Processing (ICIP)*
Perra, C., and Assuncao, P. (2016). High-Efficient Coding of Light Field Images Based on Tiling and Pseudo-Temporal Data Arrangement. In *Proceeding IEEE International Conference on Multimedia & Expo (ICME)*

Pseudo Video Sequence Compression – 2D Hierarchical Structure

- 2D Hierarchical prediction structure
- Assign different QP-values
  - Low level layer: high quality
  - High level layer: lower quality
- Adapted to employ HEVC



Liu, D. et al. (2016). Pseudo-Sequence-based Light Field Image Compression. In *Proceeding IEEE International Conference on Multimedia & Expo (ICME)*

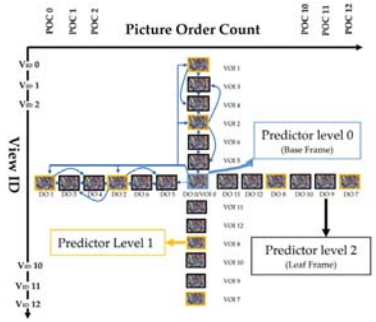
Li L. et al. (2017), Pseudo Sequence Based 2-D Hierarchical Coding Structure for Light-Field Image Compression, *Data Compression Conference (DCC), pp. 131 – 140.*

Li L. et al. (2017). Pseudo-Sequence-Based 2-D Hierarchical Coding Structure for Light-Field Image Compression, *IEEE J. Selected Topics in Signal Processing*, Vol. 11 (7), pp. 1107 - 1119

Pseudo Video Sequence Compression – 2D Hierarchical Structure

Mittuniversitetet
MID SWEDEN UNIVERSITY

- Adapted to employ *MV-HEVC*
- Assign different QP-values
  - Physical distance
  - Decoding order
- Also adapted to *any format*
  - Plenoptic camera
  - Camera arrays
  - Camera gantry



Ahmad, W. , Olsson, R. & Sjöström, M. (2017). Interpreting Plenoptic Images as Multi-View Sequences for Improved Compression. In ICIP 2017.. pp. 4557--4561.

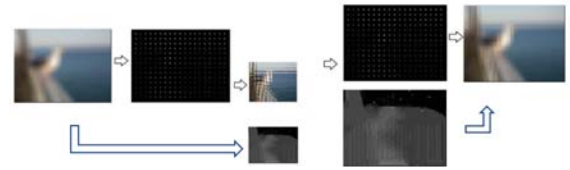
Ahmad, W. , Olsson, R. & Sjöström, M. (2018). Towards a generic compression solution for densely and sparsely sampled light field data. In Proceedings of 25TH IEEE International Conference On Image Processing.. pp. 654--658.

Model-based Compression

Mittuniversitetet
MID SWEDEN UNIVERSITY

Sparse Set + Disparity Map

- Sub-sample Microlens images
  - Displacement intra coding
- Estimate disparities (left and right)
- Transmit
  - Subsampled Microlens image
  - Disparity maps
  - Residual image

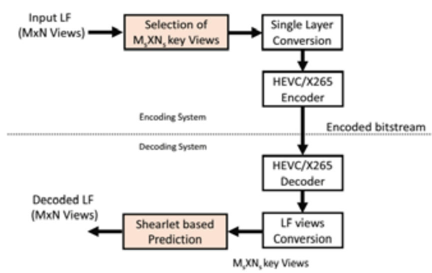


Li, Y. , Sjöström, M. , Olsson, R. & Jennehag, U. (2016). Scalable coding of plenoptic images by using a sparse set and disparities. IEEE Trans. on Image Proc, vol. 25(1):, pp. 80-91.

Intermediate View Prediction Methods

Mittuniversitetet
MID SWEDEN UNIVERSITY

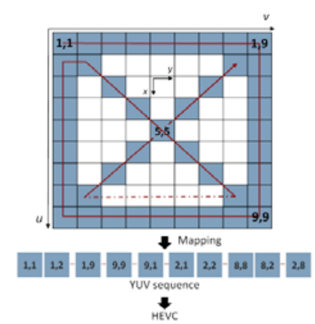
- Subsample Sub Aperture Image Grid
- Predict Removed SAIs
  - Many different options!
- Transmit Residual Image
  - For higher quality
- Many of these have good quality at low bit rate (without residual image)



Intermediate View Prediction Methods

Mittuniversitetet
MID SWEDEN UNIVERSITY

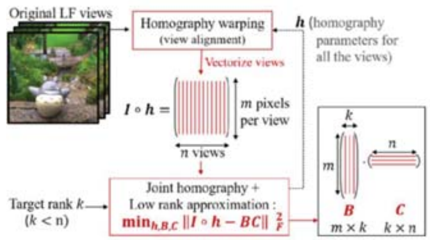
- Reconstruction of intermediate views by
  - **Sparse representation in the Fourier domain**
- Residual image encoded by **SHVC**
  - Scalability extension to HEVC
    Spatial and coarse grain SNR scalability)



Hawary, Fatma, et al. "Scalable Light Field Compression Scheme Using Sparse Reconstruction and Restoration." *ICIP 2017*. 2017.

Intermediate View Prediction Methods

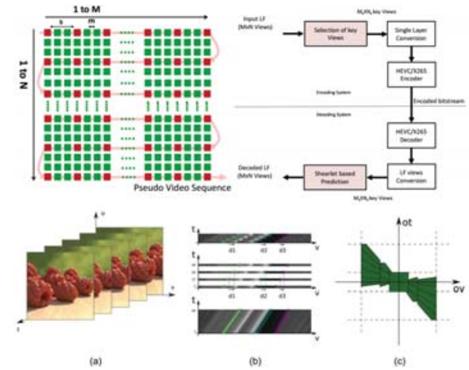
- Reconstruction of intermediate views by
  - **Homography-basd low rank approx.**
- Residual image encoded by **HEVC**



Jiang, Xiaoran, et al. "Light Field Compression With Homography-Based Low-Rank Approximation." IEEE Journal of Selected Topics in Signal Processing 11.7 (2017): 1132-1145.

Intermediate View Prediction Methods – Shearlet Transform

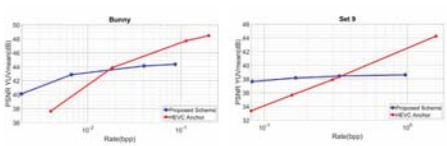
- Reconstruction of intermediate views by
  - **Shearlet transform prediction**
- Residual image encoded by **HEVC**

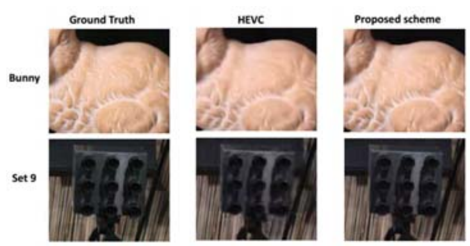


Ahmad, W., et al. "Shearlet Transform Based Prediction Scheme for Light Field Compression." Data Compression Conference (DCC 2018), Snowbird, Utah, US, March 27-March 30, 2018. 2018.

Intermediate View Prediction Methods – Shearlet Transform

- Reconstruction of intermediate views by
  - **Shearlet transform prediction**





Ahmad, W., et al. "Shearlet Transform Based Prediction Scheme for Light Field Compression." Data Compression Conference (DCC 2018), Snowbird, Utah, US, March 27-March 30, 2018. 2018.

Light Field Data Compression – Outlook

- Many new articles are appearing - Mainly intermediate view prediction
- Problems still to be solved
  - How can correlation in 4D ray-phase space be exploited further?
  - How to encode residual image efficiently?

Light Field Datasets

SMART: a Light Field image quality dataset

- *Selected image quality key-attributes*
  - Spatial information
  - Colorfulness
  - Texture key features
  - Depth distribution

Paudyal, P. , Olsson, R. , Sjöström, M. , Battisti, F. & Carli, M. (2016). SMART: a Light Field image quality dataset. In Proceedings of the 7th International Conference on Multimedia Systems, MMSys 2016.. pp. 374--379.

Light Field Datasets

Matching LF Datasets from Plenoptic Cameras 1.0 and 2.0

- For *comparison of processing on PC 1.0 and 2.0*

Fig. 5: A subset of LF images captured with Lytro Illum camera. The figure shows the central sub-aperture views.

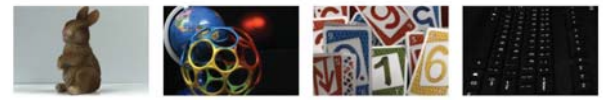


Fig. 6: The corresponding LF images captured with Raytrix R29 camera. The figure shows the total focus images.

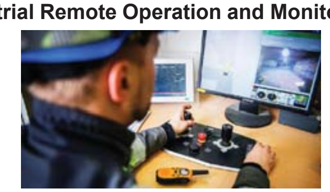
Ahmad, W. , Palmieri, L. , Koch, R. & Sjöström, M. (2018). Matching Light Field Datasets From Plenoptic Cameras 1.0 And 2.0. In Proceedings of the 2018 3DTV Conference.

Quality of Experience for Augmented Telepresence

Telepresence Applications

Airport of the Future

Industrial Remote Operation and Monitoring



https://bitzmagasin.se/garpenberg-okar-produktiviteten-med-fjarrstyrning/

Remote Health Care

https://www.idg.se/2.1085/1.614904/doktorn-kan-komma-via-internet

Real-time Surveillance

Image source: Alkit Communications AB

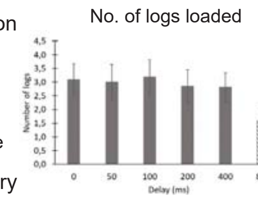
Augmented Telepresence

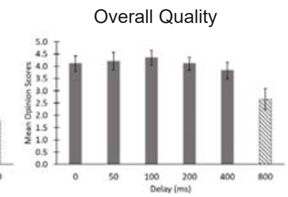
- Remote Operation and Monitoring in Real Time
- Augmented Reality to Support Operator
- Placement based on image and video analysis
- Presentation based on Information Visualization

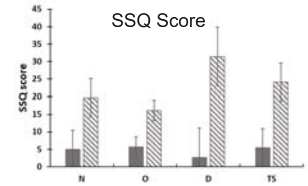
Image source: hiab.com + overlaid graphics by MIUN

Study on Quality of Experience – Simulator

- Varied delay display vs head motion
  - Not realistic – several aborted
- Varied the joystick delay
  - Number of loaded logs
  - Simulator Sickness Questionnaire
- Comparison naïve vs experts – very similar trends





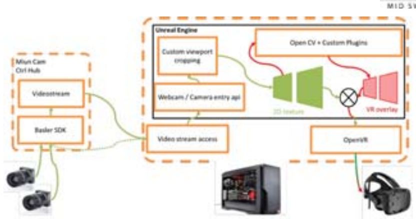


Nausea
Oculomotor
Disorientation
Total **S**core

Brunnström, K., Dima, E. Andersson M., Sjöström, Quereshi, T. & Johanson, M. (2019). Quality of Experience of visual-haptic interaction in a virtual reality simulator. In *Human Vision and Electronic Imaging (HVEI)*.

Augmented Telepresence Testbed

Light Field Evaluation System (LIFE)
- Balser machine vision cameras
- Jetson Single Board Computers
  - Near-camera video processing
  - Real-time video pair rectification
- Network connection



Stereo-pair setup

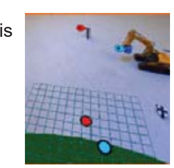
Head-mounted display

AR-enabled rendering system

Study on Quality of Experience in Laboratory Setting

Equipment:
- Augmented Telepresence – LIFE Testbed

1. Effect of camera view position
   - Task accomplishment significantly better
   - Viewpoint helpfulness significantly better
2. Effect of AR support (3 different)
   - Improved QoE
   - Improved task accomplishment for most cases
   - Viewing within Stereopsis



El. Dima; K. Brunnström; M. Andersson; M. Sjöström; J. Edlund; M. Johanson; T. Qureshi, View Position Impact on QoE in an Immersive Telepresence System for Remote Operation, 11th International Conference on Quality of Multimedia Experience (QoMEX), Berlin, Germany, 2019

Dima, E., Brunnström, K., Sjöström, M., Andersson, M., Edlund, J., Johanson, M., Qureshi, T. Effect of Position-Aiding AR and Camera Placement on Augmented Telepresence QoE, in manuscript.

Study on Quality of Experience in Industrial Setting

Equipment:

- HiVision Telepresence system
- Augmented Telepresence
- On-truck crane (but in laboratory)

1. Effect of camera view position
2. Effect of AR support (2 different)
   - Viewing on the limit of Stereopsis

Image source: hiab.com + overlaid graphics by MIUN

- To be continued…

Andersson, M., Edlund, J., Dima, E., Brunnström, K., Sjöström M., Johanson, M., Qureshi, T., Quality of Experience using AR for positioning assistance in an Augmented Telepresence system for remote operation of a crane, in manuscript.

Multi-Dimensional Signal Processing and Imaging

Realistic 3D Research Group

Prof Mårten Sjöström
STC Research Centre
Mid Sweden University, Sundsvall, Sweden
www.miun.se/stc/realistic3d