Tittle:
Immersive media quality assessment

Advisors:

Supervisor:
Pr. Luce Morin, ☎ (+33)2 23 23 87 57 luce.morin@insa-rennes.fr
Co-supervisor:
Ass. Pr. Lu Zhang, ☎ (+33)2 23 23 88 12 lu.ge@insa-rennes.fr

Laboratory: VAADER Team, IETR Laboratory, UMR CNRS 6164, Rennes, France

Advisors short bio:
Prof **Luce MORIN** is a full professor at the National Institute of Applied Sciences (INSA), Rennes, France, where she teaches computer science, image processing, and computer vision.
She received the M.S. degree from ENSPS in Strasbourg in 1989, after a 6 month internship at the NASA Goddard Space Flight Center in Washington D.C.
She then prepared a PhD thesis in Computer Vision with Pr Roger Mohr in the LIFIA laboratory, INP-Grenoble. From 1993 to 2008, she was an associate-professor at University of Rennes and a member of the Temics team in the IRISA-INRIA Rennes laboratory. Since 2008, she is a member of the IETR UMR CNRS 6164 laboratory. She has been the head of the 40 people IETR VAADER research team since 2016. She is the co-author of 4 book chapters, 19 international journal papers, and more than 70 international conference papers. She has supervised 20 PhD students, 17 of which have defended their thesis. She has also been a partner in several national and international projects, such as European Network of Excellence (NOE) SIMILAR, National ANR-PERSEE, and RUBI3 projects. She has had scientifique collaborations with industrial partners such as Orange Labs, Envivio-Ericsson, Aviwest. Her research activities deal with 2D and 3D video compression and 3D modeling from videos. In the last years, she has been working on research topics closely related to the proposed PhD topic, namely quality assessment for synthesized views and representation and synthesis of numerical holograms from video plus depth data, which are very similar to light-fields..
Personal Website: http://lmorin.perso.insa-rennes.fr/

Dr **Lu ZHANG** is an associate professor at the National Institute of Applied Sciences (INSA) of Rennes, France. She is also a member of the Institute of Electronics and Telecommunications of Rennes (IETR), UMR CNRS 6164.
She received the M.S. degree from Shanghai Jiaotong University in 2007. Then she participated in the Engineering Leadership Program (ELP) in National Instruments (NI) at Shanghai for two years. From October 2009 to November 2012, she was a phD student at the University of Angers, and at laboratories LISA (renamed as LARIS now) and IRCCyN (renamed as LS2N now) in France. Her thesis topic was “Numerical observers for the objective quality assessment of medical images”. Then she worked on the Quality of Experience (QoE) in Telemedicine as a post-doc before she joined INSA and IETR in September 2013. Her PhD thesis was awarded (“prix de thèse”) by IEEE France Section, SFGBM, AGBM & GdR CNRS-Inserm Stic-Santé.
Since 2010, Dr. Lu ZHANG is the co-author of 20 international journal papers, 33 international conference papers and 5 french conference papers. She co-supervised 9 PhD students, 5 of

them have already defended or will defend their theses before 2021. She got the PEDR (Research and Doctoral supervision prize) with a global rank A (2019-2023). Her team got the 1st prize in ICME 2018 Grand Challenge “Prediction of Head+Eye Saliency for 360 Images”. She was invited to give seminars by several chinese universities or research institutes several times. She is the project leader of an ANR (France National Agency for Research) ASTRID (Specific Support for Defence Research Projects and Innovation) project from 2018 to 2020.
Personal Website: http://luzhang.perso.insa-rennes.fr/

Thesis topic:

With coming era of immersive media, the image quality assessment (IQA) appears as **crucial** for improving the quality of the end-users' immersive experience and reducing the side effects during the users' observation, whereas they have been **little exploited until now**. As the end-users are humans, the most reliable IQA methods are based on collecting human judgments (e.g. giving a rating on a scale from 1 to 5), measured through Mean Opinion Score (MOS), called *subjective IQA*. However, this technique is time-consuming and not suitable in practical applications. Therefore, the *objective Image Quality Metric (IQM)* capable of predicting the MOS is in great demand. The IQMs can be divided into three categories relying on the amount of reference information used in the metric: full-reference (FR) metrics assess the quality of the distorted image by using the original undistorted image; reduced-reference (RR) metrics use extracted features from the original undistorted image; no-reference (NR) metrics can evaluate the quality of the distorted images without access to the original undistorted image. The FR/RR metrics normally are easier and have better performances (because of richer information), but their major limitation is that they always need a reference that is unavailable in many circumstances. NR metrics are more useful but more difficult, however AI technologies bring new possible and promising solutions to this challenge.

Both of the international standardization groups, JPEG Pleno and **MPEG-I Visual** (working on the standard framework for the LFI representation and compression, respectively) have **mentioned the lack but the importance of quality metrics for immersive medias** at CLIM2019 workshop[1]. A good IQM correlated with the MOS can also be used directly for the evaluation, optimization and improvement of compression, representation and other processing methods for immersive medias.

Light field imaging has emerged as an immersive media technology allowing capturing richer visual information from our world. On the one hand, this higher dimensional representation of visual data offers powerful capabilities for scene understanding, and substantially **improves the performance of traditional computer vision problems** such as depth sensing, post-capture refocusing, segmentation, video stabilization, material classification, etc. On the other hand, **the high-dimensionality of light fields also brings up new challenges in the data processing chain (capture, compression, content editing, and display…)**. Super multi-view light-field displays require input views acquired at high angular resolution, covering a large field of view. Acquiring content with a large number of real cameras is often economically and technologically prohibitive, so interpolating intermediate views from those captured using sparse camera arrays has been suggested to achieve the required high view density. However, views from real cameras can be distorted at each stage of the processing chain and synthesized views suffer from several artifacts, which can severely affect the perceived quality. Concerning the IQA for LFIs, Tamboli et al. [1] were the first who investigated the distortions in LFIs and proposed an IQM combining spatial information from each constituent image and angular information from consecutive images. Later they extended this IQM to videos by integrating optical flow values. Another work of these authors [2] focused on the usage of traditional IQMs (originally proposed for 2D natural images) for the evaluation of the quality of synthesized views rendered by VSRS, a depth-image-based rendering (DIBR) algorithm. The big limitation is that they only used 2 scenes with very simple background (uniform wall). Kiran Adhikarla et al. [3] tried to include more scenes and more types of distortion in their database. Using the database [3], Fang et al. [4] then proposed an IQM by extracting local features from LFIs and global features from epipolar plane images. However, the view synthesis methods in [3] were too simple and out of date. Viola & Touradj [5] proposed a VALID database focusing on the compression distortions using 5 real-world scenes. A more recent work [6] also focused on the compression and noise distortions, and proposed an IQM based on depth map feature and tested on 7 synthetic scenes with ground-truth depth maps.

The objectives of the PhD thesis will be: 1) Construct a large-scale database, based on subjective test where human observers will assess the quality of LFIs with different contents and rendered by different combinations of state-of-the-art depth estimation and view synthesis algorithms, on both 2D display and Light Field Display (LFD). This will be the first database (providing the MOSs on LFIs' quality) large and complex enough to be used for the objective quality metrics' design based on AI (Artificial Intelligence) techniques and for the validation of these metrics. 2) Propose NR quality metrics correlated to the MOSs (got from Objective 1) based on deep-learning and machine-learning techniques.

Our previous related studies:

Study 1: Synthesized view quality assessment – In the context of S. Tian's PhD thesis, we have proposed one database [7], 2 FR and 2 NR metrics ([8-11]) for the evaluation of DIBR algorithms for free viewpoint television (FTV) applications (where the baselines are large). Note that the view synthesis is easier for dense light fields with narrow baselines, but more difficult for sparse light fields with large baselines, where the quality evaluation may make a bigger difference. Though these are traditional metrics, the studies gave us useful knowledge on the distortion specificities that may appear in DIBR synthesized views, thus are helpful for new metrics' design (even based on learning).

[1] http://clim.inria.fr/workshop.htm

Study 2: Database with ground-truth depth maps and Depth estimation algorithm based on deep network for LFIs – Recently, collaborated with ULB, we proposed together a deep end-to-end network for LFI depth estimation, which is 3 times smaller and 3 times faster than the current top-performing method Epinet [12]. The two existing synthetic databases HCI and CVIA-HCI imitate images captured by micro-lens array based LF cameras. While LF camera arrays can produce images with a much higher spatial resolution, no synthetic database simulating this type of images has been proposed before. To fill this gap, we constructed such a database including 30 synthetic scenes. One paper on this work has been submitted to a journal.

[1] Tamboli, Roopak R., et al. "Super-multiview content with high angular resolution: 3D quality assessment on horizontal-parallax lightfield display." Signal Processing: Image Communication, 2016.
[2] Tamboli, Roopak R., et al. "Objective Quality Assessment of 2D Synthesized Views for Light-Field Visualization." International Conference on 3D Immersion (IC3D), 2018.
[3] Kiran Adhikarla, Vamsi, et al. "Towards a quality metric for dense light fields." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017.
[4] Fang, Yuming, et al. "Light Filed Image Quality Assessment by Local and Global Features of Epipolar Plane Image." IEEE Fourth International Conference on Multimedia Big Data (BigMM), 2018.
[5] Viola, Irene, and Touradj Ebrahimi. "VALID: Visual quality assessment for light field images dataset." Tenth International Conference on Quality of Multimedia Experience (QoMEX), 2018.
[6] Paudyal, P., Battisti, F., & Carli, M. "Reduced Reference Quality Assessment of Light Field Images. IEEE Transactions on Broadcasting", 65(1), 152-165, 2019.
[7] S. Tian, L. Zhang, L. Morin, O. Deforges. “A benchmark of DIBR Synthesized View Quality Assessment Metrics on a new database for Immersive Media Applications”. IEEE Transactions on Multimedia; October 2018.
[8] S. Tian, L. Zhang, L. Morin, O. Deforges. “NIQSV+: A No Reference Synthesized View Quality Assessment Metric”. IEEE Transactions on Image Processing; December 2017; 27(4):1652-1664.
[9] S. Tian, L. Zhang, L. Morin, O. Deforges. “NIQSV: A No Reference Image Quality Assessment Metric for 3D Synthesized Views”. ICASSP, March 2017, New Orleans, USA.
[10] S. Tian, L. Zhang, L. Morin, O. Deforges. “SC-IQA: Shift compensation based image quality assessment for DIBR-synthesized views”. Visual Communications and Image Processing (VCIP), December 2018, Taichung, Taiwan.
[11] S. Tian, L. Zhang, L. Morin, O. Deforges. “Performance comparison of objective metrics on free-viewpoint videos with different depth coding algorithms”. SPIE Optical Engineering + Applications, August 2018, San Diego, California, USA.
[12] Y. Li, L. Zhang, Q. Wang, G. Lafruit. “MANet: Multi-scale aggregated network for light field depth estimation”. ICASSP, May 2020, Barcelona, Spain.