

Evaluation of DIBR-synthesized videos and methodologies comparison

POINT MARIE
SACHOT CLARA

Outline

Introduction

1 The original experiment

- The dataset
- The methodology

2 Our experiment

- The dataset
- The methodology

3 Objective metrics analysis

4 Subjective metrics analysis

5 Comparison

Conclusion

Introduction

Continuation of the QoE mini-project

→ Comparing subjective and objective metrics for a video QA task

On this project

→ Collecting our own subjective datas using another methodology, and comparing the results

Will the SAMVIQ results differ from the ACR-HR ones for this precise task, and in what way?

The original experiment

The dataset



HAL
open science

Perceived quality of DIBR-based synthesized views
Emilie Bosc, Romuald Pépion, Patrick Le Callet, Martin Köppel, Patrick
Ndjiki-Nya, Luce Morin, Muriel Pressigout

Conference :

SPIE. OPTICS+
PHOTONICS

Journal :

IEEE J-STSP (2011) Journal of
Selected Topics in Signal
Processing

The original experiment

The dataset

IRCCyN/IVC DIBR Videos:

- SRC

- 102 video sequences of 6 seconds each
- 1024x768
- Between 15 and 30 Hz
- 3 different multiview plus depth sequences

- HRC

- 7 DIBR (Depth Image-Based Rendering) algorithms

Each DIBR algorithms processed each three sequences to generate 4 new viewpoints

SRC	Source name	Preview	Description
01	Book Arrival		<u>Name</u> : Book Arrival <u>Description</u> : A man in an office standing up to welcome another man. <u>Author</u> : Heinrich-Hertz-Institut. <u>Link</u> <u>Copyright</u> . <u>Video frame rate</u> : 15Hz <u>Number of frames</u> : 100
02	Lovebird1		<u>Name</u> : Lovebird1 <u>Description</u> : A couple walking in front of a temple. <u>Author</u> : Electronics and Telecommunications Research Institute (ETRI). <u>Link</u> <u>Copyright</u> . <u>Video frame rate</u> : 30Hz <u>Number of frames</u> : 150
03	Newspaper		<u>Name</u> : Newspaper <u>Description</u> : A man and a girl reading the newspaper around a table when a man coming behind them. <u>Author</u> : Gwangju Institute of Science and Technology (GIST). <u>Link</u> <u>Copyright</u> . <u>Video frame rate</u> : 30Hz <u>Number of frames</u> : 200

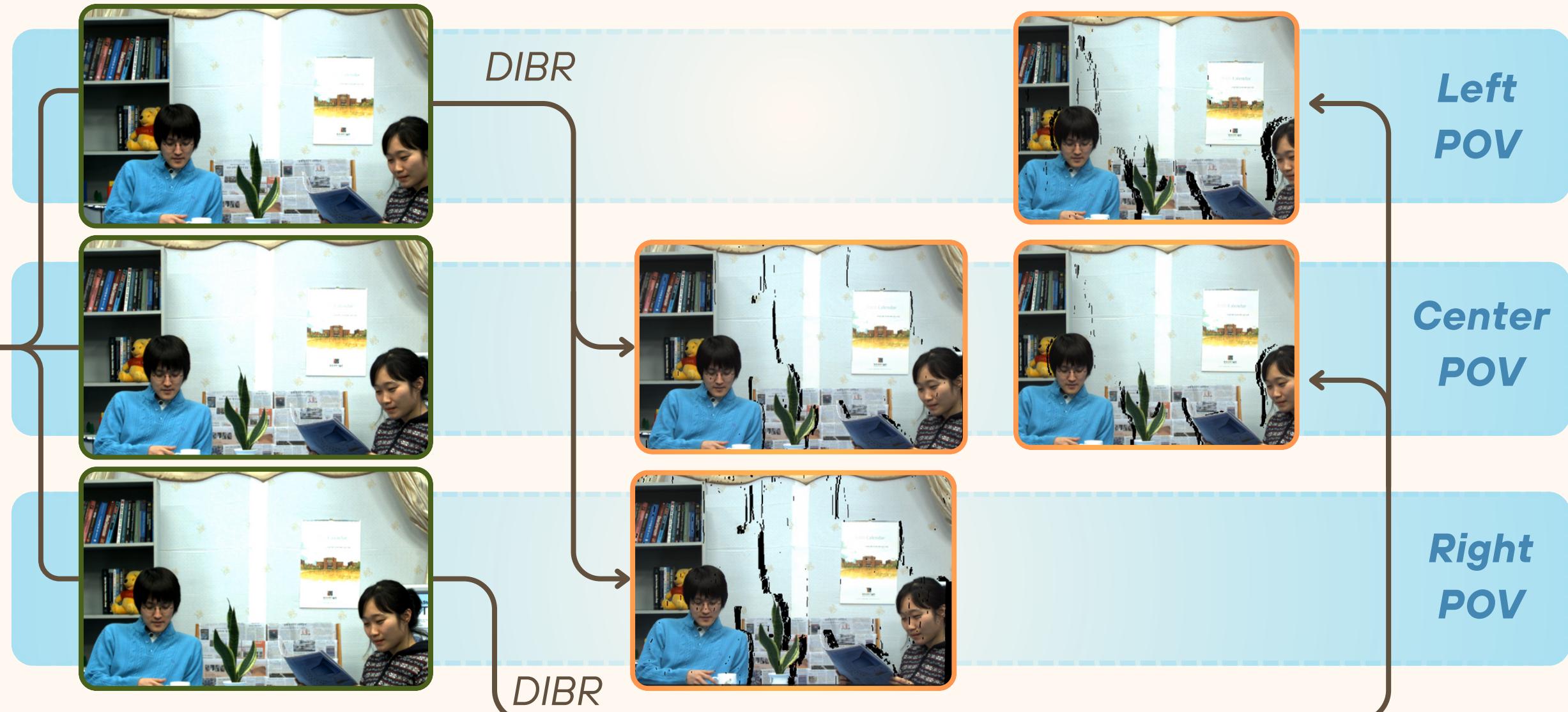
The original experiment

The dataset



**3 Sources
(SRC)**

x7 DIBR algorithms



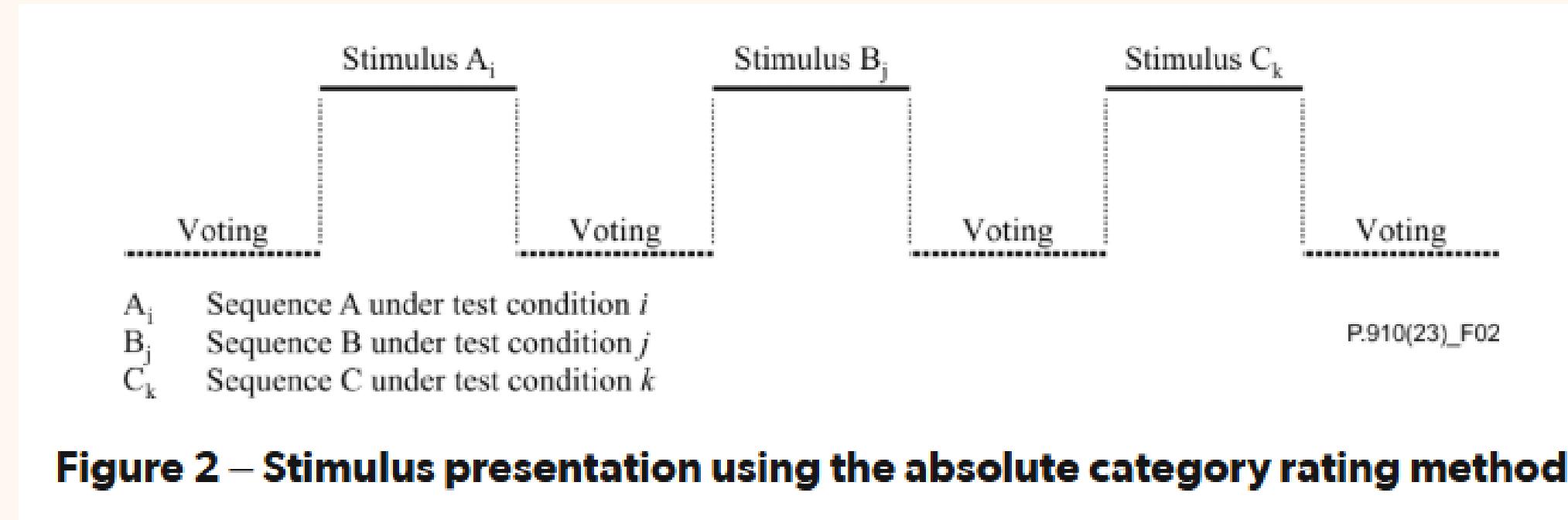
**3 Viewpoints
per source :
left, center
and right POV**

**4 generated
viewpoints per source
and per algorithm
(PVS)**

The original experiment

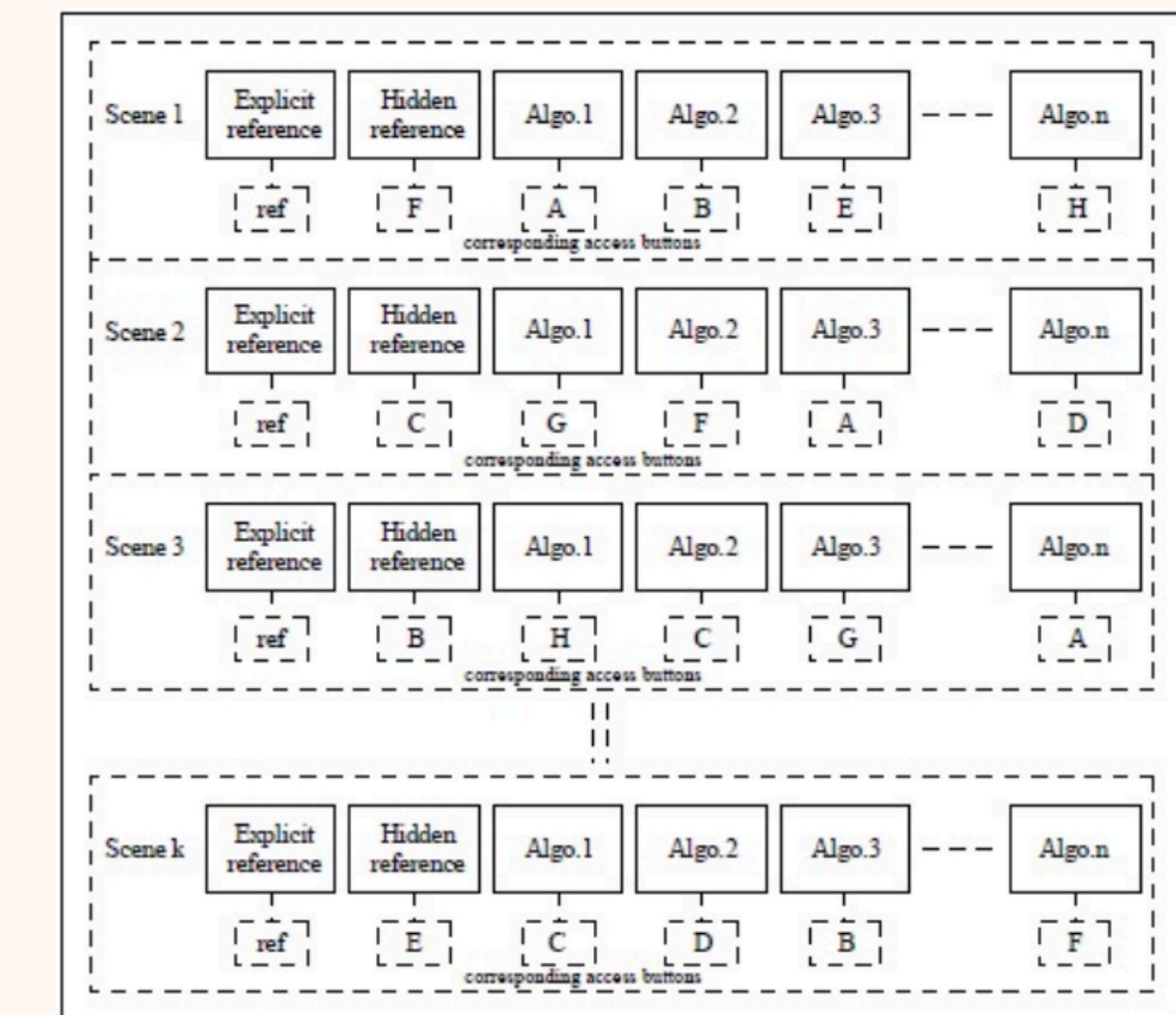
The methodology

- Test procedure : ACR-HR (Absolute Category Rating with Hidden Reference):
The participant watch a sequence once and rate its quality right after
- 32 participants, each saw and rated each PVS



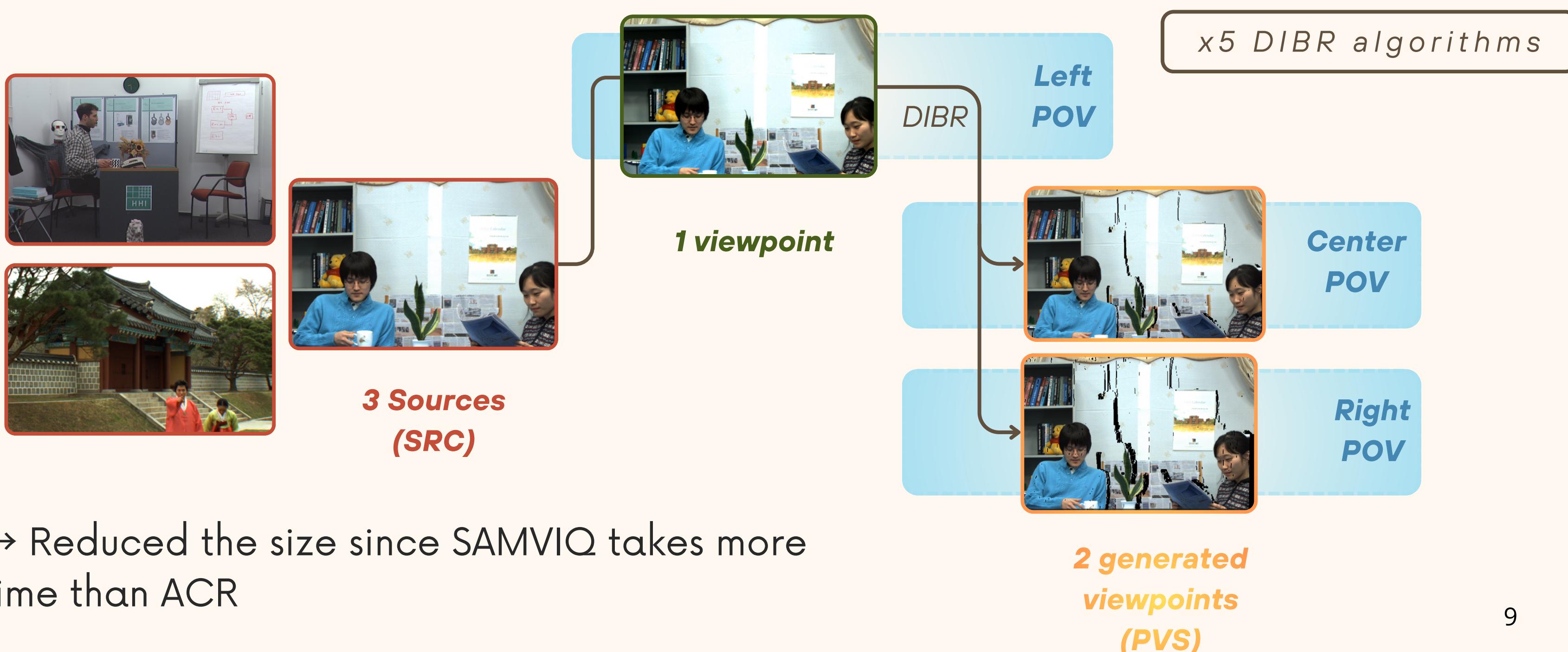
Our experiment

Replicate the experiment, but with a
SAMVIQ (Subjective Analysis
Methodology for Video Quality)
procedure



Our experiment

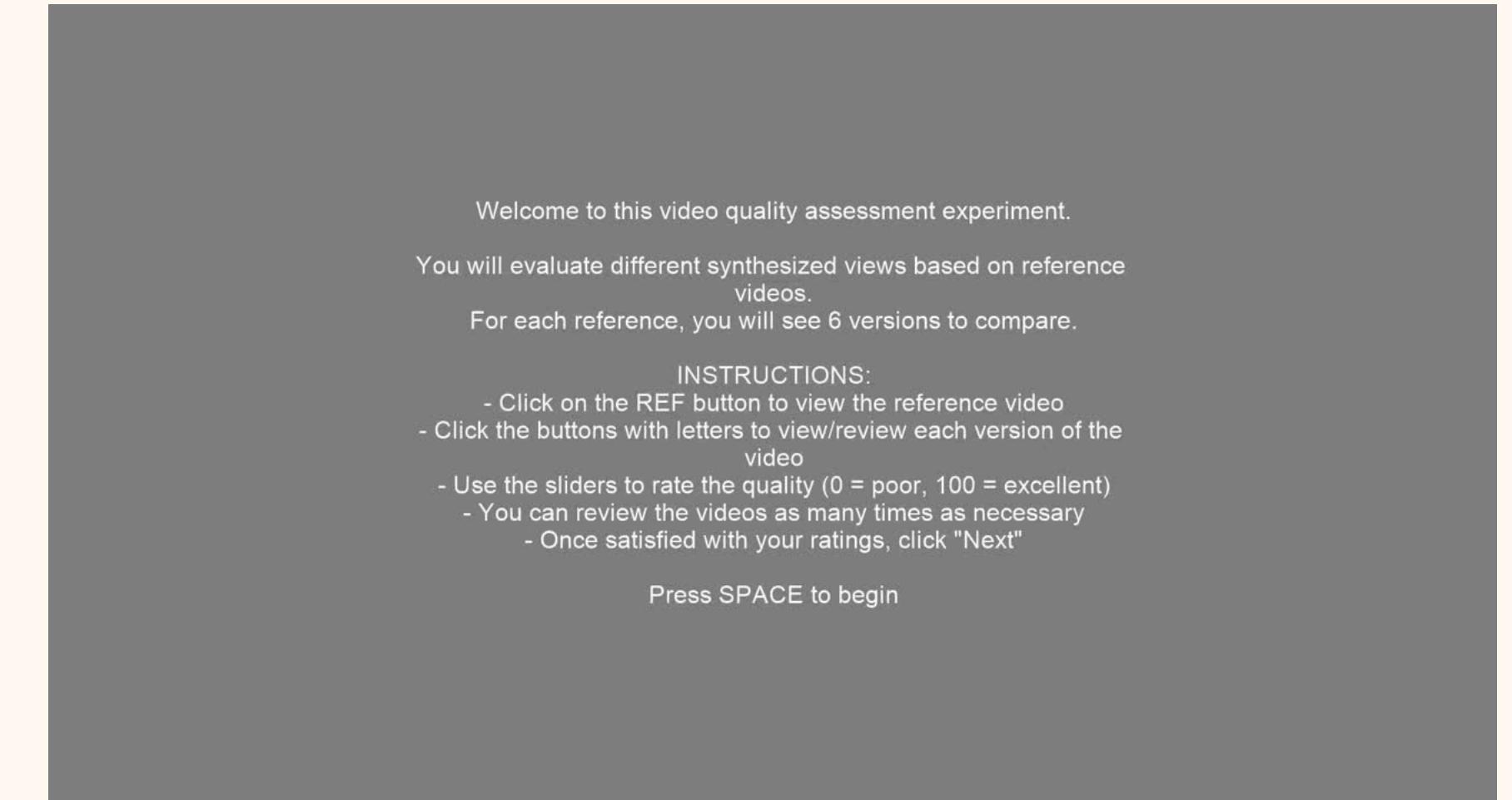
The dataset



Our experiment

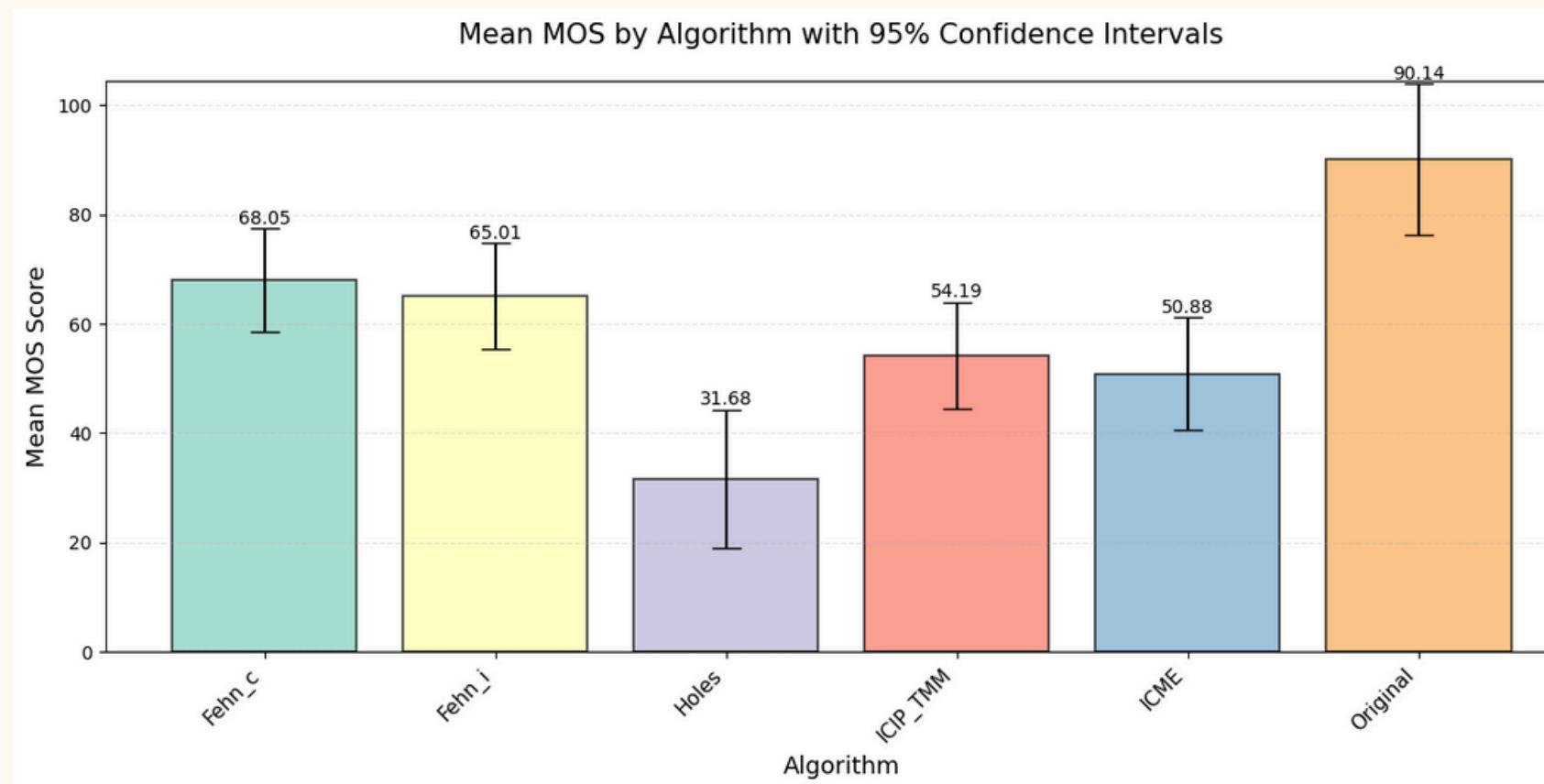
The methodology

- Written rules
- Training phase
- Test phase
 - 6 sources x 5 synthesized views
 - + 1 hidden reference
 - Rating with a sliding bar (0-100)

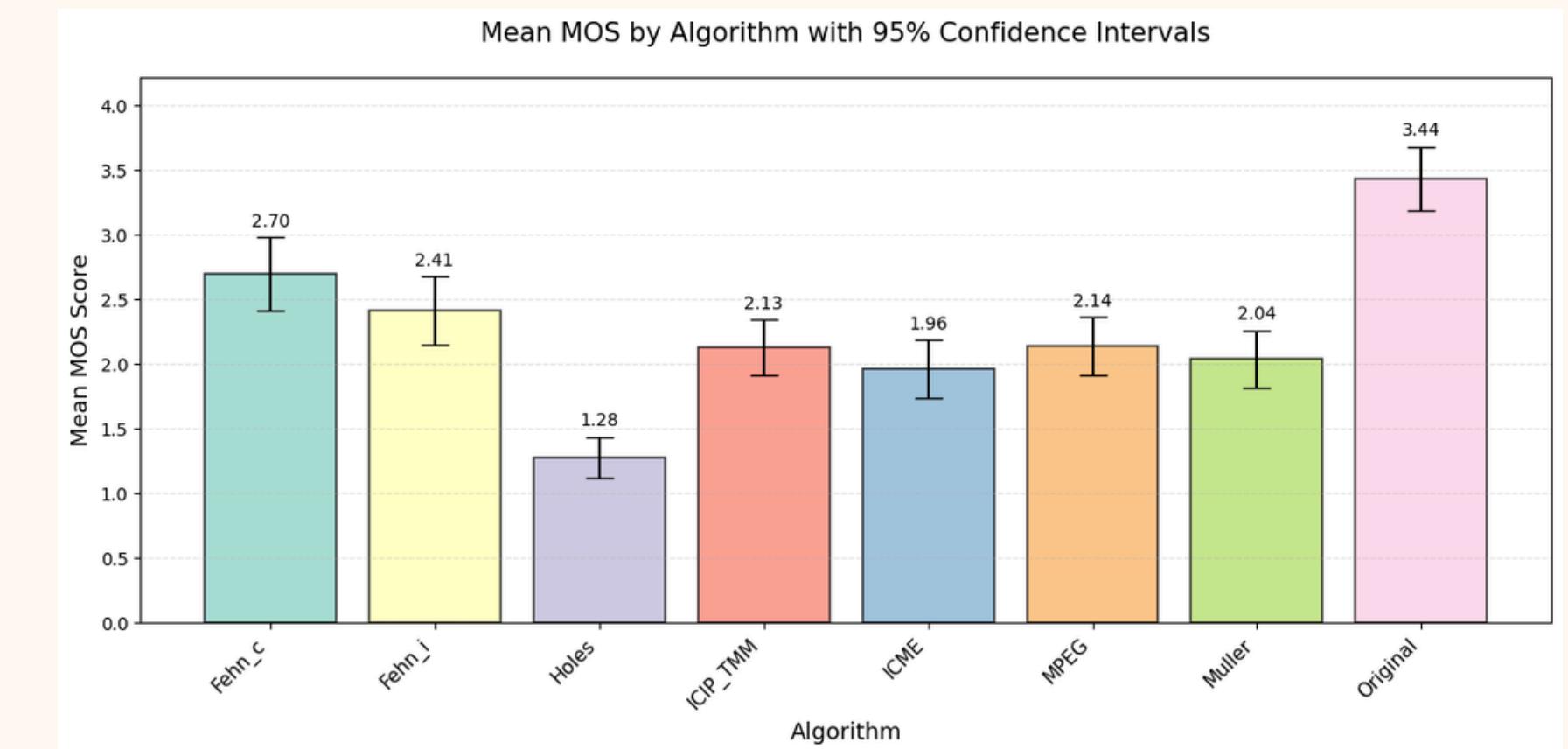


→ 15 participants (avg. age : 22), each saw and rated each PVS

Subjective metrics



Mean MOS and CI per HRC on our experiment



Mean MOS and CI per HRC on the original experiment

Subjective metrics

One-way ANOVA

P-value: 9.63e-12

→ There are statistically significant differences between algorithms

Post-hoc : Dunnett test

Comparison of each algorithm with the Original (Control)

Difference of perceived quality for each algorithm is statistically significant compared to original

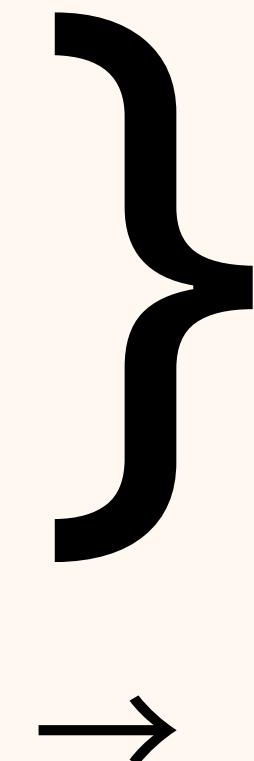
→ Not the case for Fehn_c in the original experiment

SUMMARY: Comparison of Each Algorithm vs. Original (Control)				
Algorithm	Mean Diff	95% CI	p-value	Significant?
Fehn_c	-22.5444	[-8.578, -36.511]	0.0004	YES
Fehn_i	-25.2778	[-11.311, -39.245]	0.0001	YES
Holes	-59.1778	[-45.211, -73.145]	0.0	YES
ICIP_TMM	-36.6889	[-22.722, -50.656]	0.0	YES
ICME	-40.0889	[-26.122, -54.056]	0.0	YES

Objective metrics

4 metrics :

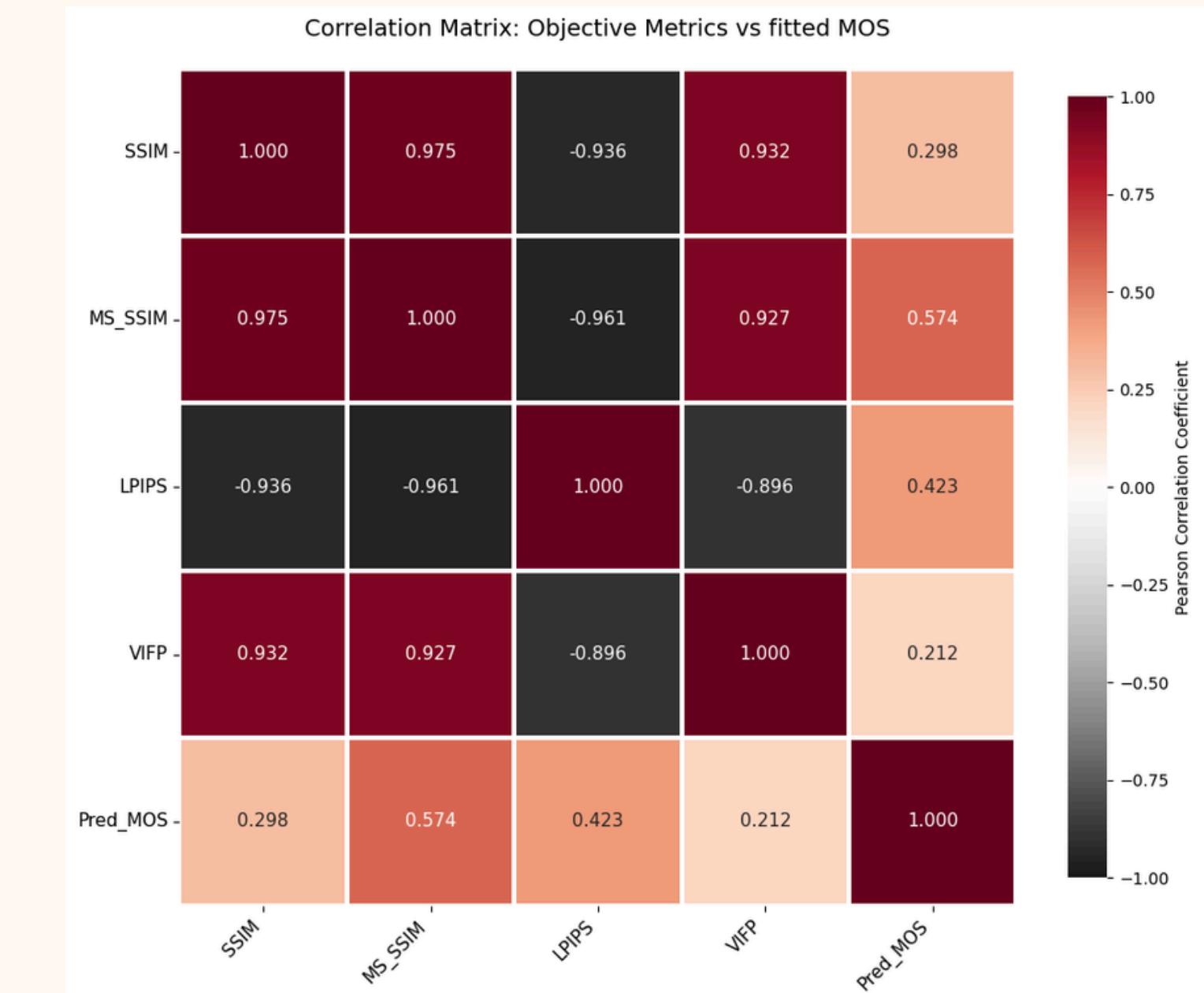
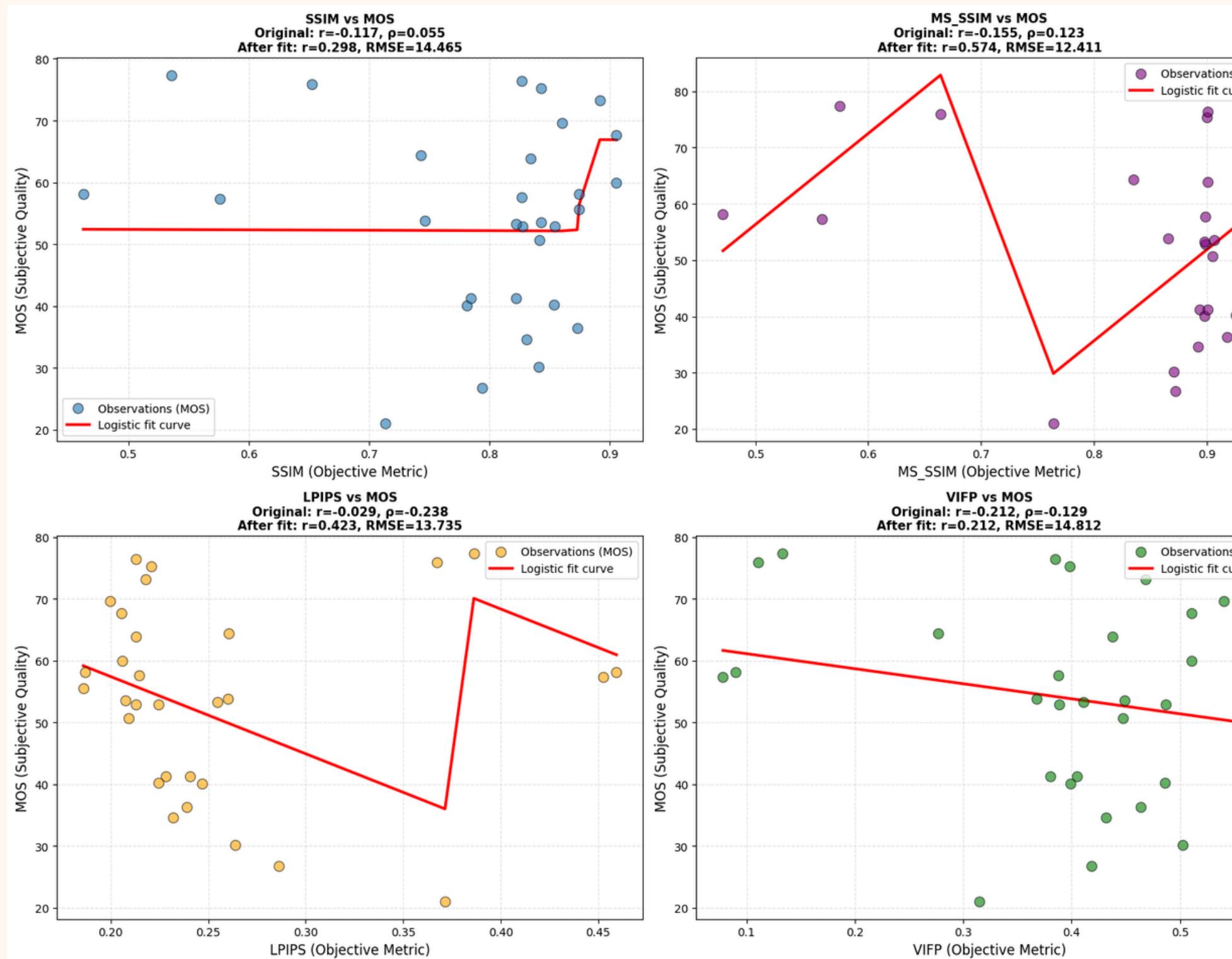
- **SSIM** (Structural Similarity Index) : widely used on State-of-the-art
- **VIFP** (Visual Information Fidelity Pixel-wise) : measure image fidelity by quantifying the amount of visual information preserved after image processing
- **MS_SSIM** (Multi-Scale SSIM)
- **LPIPS** (Learned Perceptual Image Patch Similarity) : calculates perceptual similarity between two images



Also present in
the original
paper

interesting measure
for our task

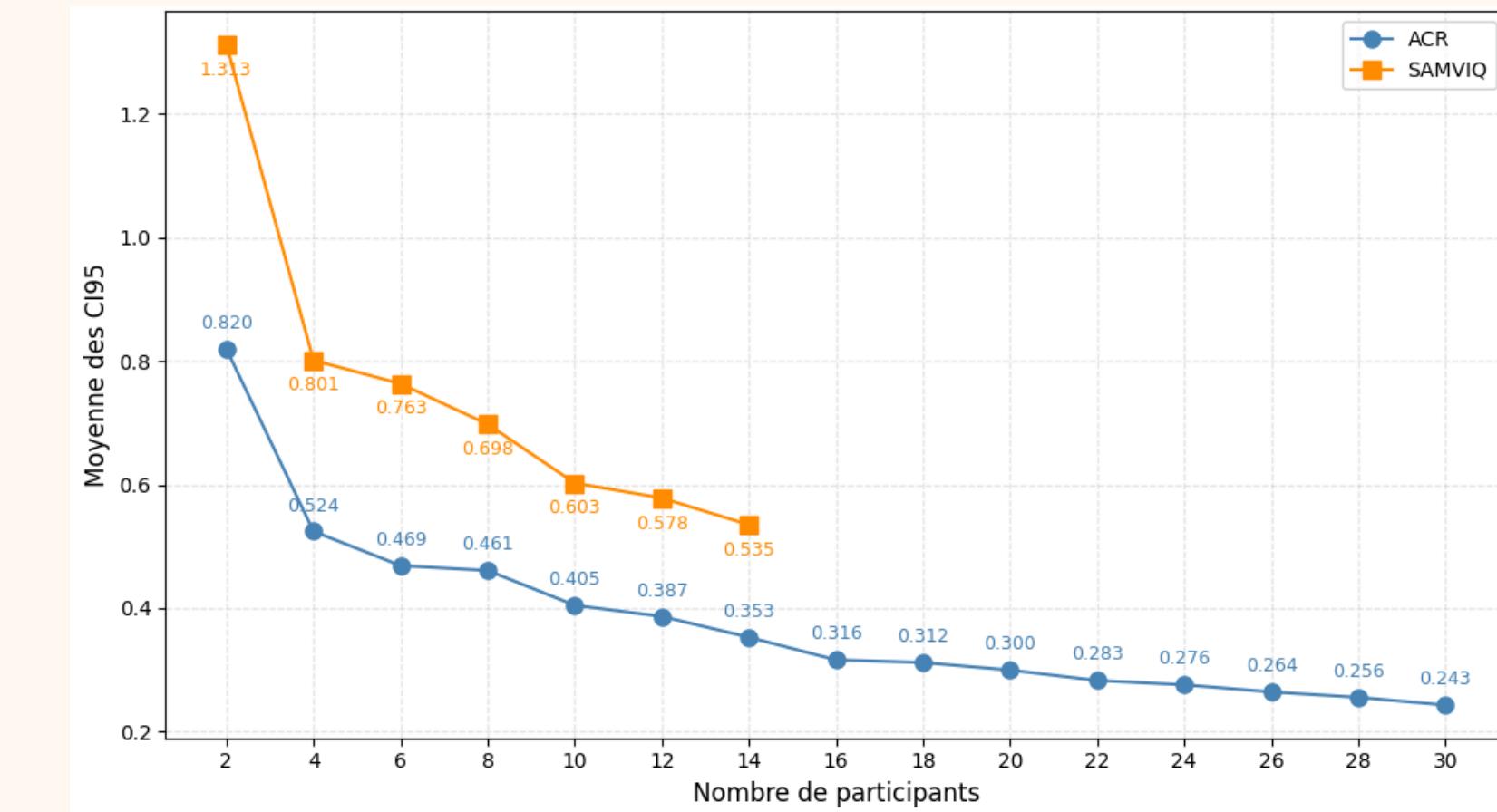
Objective metrics



Comparison ACR vs. SAMVIQ

	ACR-HR	SAMVIQ
SSIM	0.452	0.298
VIFP	0.48	0.212
MS_SSIM	0.27	0.574

Correlation with MOS



Mean CI according to the number of users

Discussion

Points to improve :

- **The sample of participants:** have more participants and more various
- Get the exact **same material and test conditions**
- Test on the **whole dataset**
- **The fitting function of the MOS:** use the same fit than in the original paper
- **The objective metrics:** run the metrics on more frames / use metrics adapted to artefacts specific to DIBR synthesized views

Conclusion

With our project, we :

- gathered our own subjective data
- experimented SAMVIQ and compared it with ACR-HR

The results that we analysed are :

- less accurate than the original results of the paper (our Confidence Intervals are wider)
- MS-SSIM ($r=0.574$) and LPIPS ($r=0.423$) gave better predictions for our subjective data than the other metrics, but still they are not enough

BYOE

29/01/2026

Thank you

POINT MARIE
SACHOT CLARA