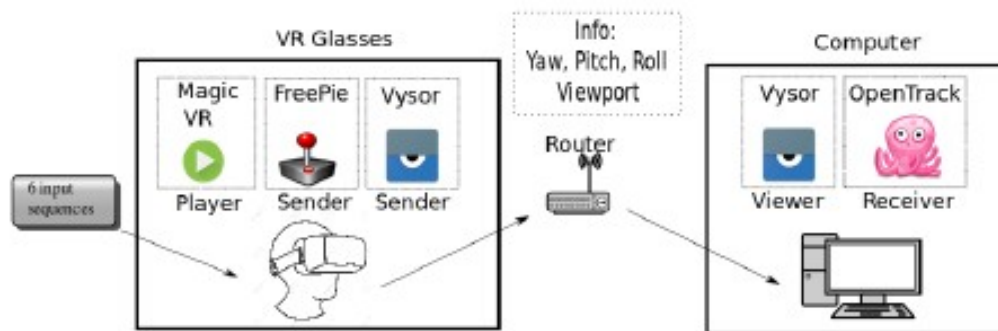


In this work, we created a platform testbed to automatically run the experiment and collect the data. The block diagram of this testbed platform is displayed in the figure bellow. The three most important elements in the testbed are the HMD, the router, and the server computer.



The participants watched the 360-degree videos using the **BOBOVR** X1 HMD [2]. BOBOVR has the following technical specifications: a modified Android 4.4, an octacore ARM 2GHz processor, 2GB DDR3 of RAM, sampling rate of 250 Hz, a Wifi 802.11 b/g/n connection, and a 5.5 inches screen display with a spatial resolution of 1920\_1080 and a FOV of 120 degrees.

The server computer has an Intel Core processor i5 7400@3GHz with 28 GB of RAM and runs Windows 10. The router is a Bandluxe, model R100 3.5G Wifi. The data flow goes from the HMD to the server, passing through the router. The router is responsible for linking the HMD and the computer server that stores head movements information.

The HMD applications are Magic VR, FreePie, and Vysor app. Magic VR is responsible for playing the 360-degree videos. FreePie records the yaw, pitch, and roll information, while the Vysor app records a copy of the viewport watched by the users. The data captured by the HMD is sent to the server. The softwares used in server are Vysor-for-desktop and OpenTrack. Vysor receives and plays the viewports so that the experimenter can see what each participant is watching. OpenTrack receives the head movement data and records it in a 'csv' format. This software also synchronizes the devices to make sure that the information is not out of sync.

Before the participants started watching the sequences (each participant watches a different set of sequence), the experimenter gave him/her instructions on how to explore the videos by moving his/her head.

To make sure that the dataset had a diverse content, representative of the types of videos available in the internet, we specified the following content categories:

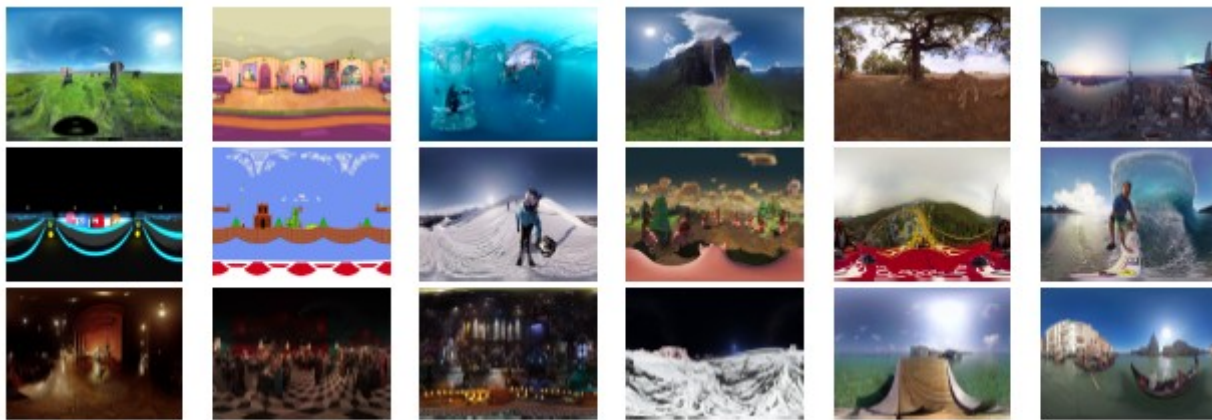
- Category 1: Low temporal activity and low spatial activity – Videos with little movement and little texture.
- Category 2: Low temporal activity and high spatial activity - Videos with low camera movement and scene textures, such as grass or buildings.
- Category 3: High temporal activity and low spatial activity - Videos with many moving objects, but with mostly uniform areas (e.g. animations).
- Category 4: High temporal activity and high spatial activity – Fast camera movements and scene textures.
- Category 5: Videos with people who interact with the viewer in the form of greetings, dialogues or body movement.
- Category 6: Contemplative videos of scenarios with little movement and no focus of attention.

Our goal here was to have videos in each of the above categories. Our premise is that the users' behavior varies according to the type of content (i.e. category). We selected eighteen 360-degree videos from Youtube, with 3 videos in each content category. The next Table shows a description of these 18 videos, including the category they belong to.

Index (video)	Video Title	Duration (sec)	SI	TI	Bitrate (Mbps) CRF = 17 resolution	Content Category resolution	Predicted Classes	Exp. Group
V1	Wild Elephants	60	129.34	3.04	36.4	CC-1	6	G1
V2	Omnon Stories	60	28.56	1.38	2.4	CC-1	1	G2
V3	Isle of Jaws	60	25.62	10.98	40.8	CC-1	1	G3
V4	Angel Falls	60	59.82	4.27	36.2	CC-2	6	G1
V5	Lions	60	99.62	2.09	42.4	CC-2	2	G2
V6	Manhattan Aerial	60	67.14	15.30	41.2	CC-2	2	G3
V7	Pac-Man	60	26.11	11.92	11.6	CC-3	3	G1
V8	Every Super Mario	60	114.03	1.25	2.1	CC-3	3	G2
V9	Ski	60	40.97	13.40	33.6	CC-3	3	G3
V10	Hog Rider	60	30.13	15.65	48.0	CC-4	4	G1
V11	RollerCoaster	60	70.93	10.36	55.0	CC-4	4	G2
V12	Tahiti Surf	60	30.44	12.50	54.3	CC-4	4	G3
V13	Circle of Life	60	31.05	7.42	21.4	CC-5	6	G1
V14	Fifty Shades Darker	60	32.10	2.57	18.5	CC-5	5	G2
V15	Mutant Ninja Turtles	58	69.86	2.24	8.6	CC-5	5	G3
V16	Pluto Frigid Heart	60	24.64	1.95	10.5	CC-6	6	G1
V17	Maldives	60	28.51	0.50	26.7	CC-6	6	G2
V18	Carnival in Venice	60	78.55	9.88	39.6	CC-6	6	G3

The videos were then divided into 3 groups, with each group having one video of each category. During the experimental sessions, each participant was assigned to a group of videos (1, 2, or 3), which contained one video of each category. Notice that 17 videos are 60 seconds long and 1 video is 58 seconds long. The videos were formatted in an equirectangular projection and, originally, had a 4K spatial resolution. To display these videos in the BOBOVR HMD, they were downsampled to a spatial resolution of 1920x1440 pixels and a temporal resolution of 24 frames per seconds (fps).

The figure bellow shows sample frames of the 360-degree videos in our dataset.



Given that a high exposure to immersive technologies can cause fatigue, participants did not watch all videos in our dataset. To make the experimental session more enjoyable to the user, the **18 videos were divided into 3 groups**, with each group having 6 videos of different content categories. Additionally, the order of the videos was shuffled to avoid a temporal effect.

More specifically, each participant watched one of the following six video sequences:

1. Sequence 1: Angel Falls, Surrounded by Wild Elephants, Pac-Man, Hog Rider Circle of Life, Pluto's Frigid Heart.
2. Sequence 2: Hog Rider Circle of Life, Pluto's Frigid Heart, Angel Falls, Surrounded by Wild Elephants, Pac-Man.
3. Sequence 3: Fifty Shades Darker, RollerCoaster, Omnon stories, Lions, Every Super Mario Game Ever, Maldives.
4. Sequence 4: Omnon stories, Lions, Every Super Mario Game Ever, Maldives, Fifty Shades Darker, RollerCoaster.
5. Sequence 5: Tahiti Surf, Isle of Jaws, Ski, Manhattan, Teenage Mutant Ninja Turtles, Carnival in Venice
6. Sequence 6: Manhattan, Teenage Mutant Ninja Turtles, Carnival in Venice, Tahiti Surf, Isle of Jaws, Ski.

After finishing the experimental session, the participants were asked to fill a questionnaire. The first questions of the questionnaire were about the name, age, gender, occupation of the participant, and if he/she used glasses or contact lenses to watch TV. Then, the following questions of the questionnaire were about his/her experience with virtual reality applications:

- It is the first time you wear virtual glasses ?
- Have you watched 360-degree videos?
- Have you felt any discomfort while watching the experiment?
- How would you rate your quality of experience during the experiment?
- How would you rate the visual quality of the content you have just watched?
- How would you rate the sound quality of the content have just you watched?
- How would you rate the immersion / spatial involvement of content?
- Has this immersion / spatial involvement contributed to the quality of your experience?
- Has the sound contributed to the quality of your experience?
- Describe your experience using 3 adjectives.

As total of **120 participants** participated in the experiment. Therefore, each sequence was watched 40 times. Due to network errors, 3 sequences were lost from Sequences 1, 4, and 6. This way, each video was watched at least 39 times.