August 19, 2019

SYSTEM AND METHOD TO IMPROVE HUMAN FACE QUALITY IN A TELEPRESENCE

HP INC

System and method to improve human face quality in a telepresence

This idea describes a system and method to improve human face quality in a telepresence using outlining and gaze detection.

During a telepresence session, the focus is on the primary subject. That primary subject may be a human or an object. The degraded network makes the video quality suffer in terms of pixels being passed, introduction of latency, etc. This invention describes a system and method to improve the quality of the primary object in a telepresence session. The idea is to segment and remove background to reduce the bandwidth required to transfer stream data during poor network situations. Once the network is improved the background stream to the remote party is re-enabled. For transmission, each object in the frame gets a rating. That rating is decided based on gaze detection.

The system may use the following method.

1. Two parties (let's name them A and B) initiate a telepresence session using an open protocol or a proprietary protocol.
    a. This requires multiple systems like signaling server, relay servers, etc. However, those details are not important for this invention.

2. Once the telepresence session is established, systems of both A and B validates enough processing power and required algorithms are available in the system.

3. If both are available, the system jumps to the next step below. If the algorithms are not available, the system downloads them in the background. If the required processing power is not available, the system skips this feature entirely.

    a. Once the algorithms are available, systems of both parties start tracking the objects human observe the most using gaze detection. Take for example telepresence in a meeting room. Sometimes the remote party may be interested in looking at the speaker and at other time they might be more interested in looking at a board in the room. So, the importance of each object differs. All this information is tracked and stored locally.

4. The system keeps monitoring the network for quality.
    a. This may include multiple parameters like latency, bandwidth, packets received, etc.

5. If the system notices network performance degradation, the system turns on optimization mode. Or the user may manually turn on this feature to save on bandwidth.
    a. First, each party request historical gaze detection data from the remote party. If there are multiple remote parties involved, all data may get aggregated and then averaged.
    b. In this mode, the system tries to find the outline of every object. Next, it uses the historical data gathered above by the remote party to give priority for each object (for example, out of 3 people in a room, one of them is the most important because the remote party has been looking at them most of the time). This historical data is used to jump start the raking system. However, the system may also take into account real-time gaze detection of the remote party. Real-time gaze detection may get priority over the historical data.
    c. Next, the system just sends the pixels inside an outline of the most important objects. And ignores all the background pixels. The system may rely on machine learning vision algorithms to find an outline of the person or it may use a depth sensing device like Intel RealSense. Next, the system tries to put a bounding box around the person (x, y, height,

width). The system sends both the pixels inside the outline and bounding box to the remote system. The remote system uses bounding box information to place the person in the frame. This helps in the placement of the remote person.

6. Once the network improves to a preset threshold value, the optimization mode is turned off.

The system of outlining a person is used in other telepresence systems before. However, this invention provides further system and method around it to make it successful for poor network telepresence.

Disclosed by Arjun Angur Patel, HP Inc.