June 20, 2018

WIRELESS AUGMENTED REALITY (AR) / VIRTUAL REALITY (VR) CHANNEL PREDICTION

Ardalan Alizadeh

Amir Kamalizad

Matt Silverman

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

WIRELESS AUGMENTED REALITY (AR) / VIRTUAL REALITY (VR) CHANNEL PREDICTION

AUTHORS:
Ardalan Alizadeh
Amir Kamalizad
Matt Silverman

ABSTRACT

A wireless AR/VR headset may constantly report the exact coordination and orientation of itself and/or a user to a server. This information can be initially used at an access point (AP) to generate a look-up table (LUT). The LUT may include estimated channels from the AP to all spatial points that the headset may potentially sweep (e.g., in a certain application). During regular usage of AR/VR, the reported location of the headset is extracted in the AP to predict the channel at each point from the LUT and to update a multi-user MIMO steering matrix accordingly. A sounding process is then carried out only once (e.g., at the beginning when generating the LUT) and spatial beamforming is updated with no sounding process. This significantly reduces the latency of the network.

DETAILED DESCRIPTION

Popularity in AR/VR is trending upwards. For example, immersive forms of video entertainment, such as immersive gaming, as well as live viewing events, are areas of AR/VR that are becoming more popular. Certain applications, such as 360-degree videos, may be extremely sensitive to delay and require very low end-to-end latency. For example, inconsistencies between a user's perception of movement and the display of a corresponding image may result in the user experiencing digital motion sickness or cyber-sickness. As such, reducing the delay between the perception of movement and the display of the corresponding image may be beneficial. However, processing-intensive operations associated with AR/VR applications, systems, devices, etc. add to such latency.

An orientation and position of a headset rapidly changes in interactive applications. As a result, the wireless channel changes frequently. In the design of the next generation of wireless standards, the process of channel estimation plays an important role since access points (APs) / E-UTRAN Node Bs (or Evolved Node B) (eNBs) have a large number

1

5638X

of antennas (e.g., Massive MIMO). Further, the network is usually dense and, as a result, a large portion of transmission is generally consumed by pilot transmissions that are only for channel estimation. Moreover, downlink beamforming relies on accurate channel estimation and outdated channel state information (CSI) can cause large packet error-rate (PER). This may result in higher latency.

Accordingly, a need exists for an effective channel estimation technique that does not impose extra delay on wireless VR/AR devices. Such a need may exist for many different wireless technologies, such as, IEEE802.11ax and 5G NR.

In interactive VR/AR applications such as multiplayer VR/AR gaming applications, the position and movements of the headset may be constantly transmitted to one or more server(s). This information may be required at the server(s) to, for example, update a field of view and a location of one or more players in the game with very low latency to provide a close-to-real-world experience.

Fast movements of a VR headset that includes embedded antennas changes the wireless channel significantly in an indoor environment with multiple non-line-of-sight signal receptions. When the access point is working under downlink multi-user multi-input multi-output (DL-MU-MIMO), which is now part of 802.11ac/ax, the accuracy of channel estimation is critical. This is because any error in channel state information (CSI) results in degradation on the gain of beamforming. Such high packet-error rate (PER) may cause re-transmission of data, and more frequent updating of the channel sounding will accordingly increase latency.

Techniques are presented herein to "predict" a wireless channel based on the accurate coordination of the headset information transmitted to the server, instead of estimating by sounding. More precisely, when a user changes the position of a headset, the built-in sensor information is sent to the server (e.g., an gyroscope, accelerometer, and/or magnetometer may be employed in the headset to provide accurate three-axis measurements). The sensor measurements may be extracted in an access point, which has at least the delta (difference) coordination of the headset in each reporting time instance.

In certain examples, the position information transmitted by the sensors is very accurate. However, current commercial wireless location (positioning) products do not provide such accuracy of users' location in an indoor environment(s). Therefore, the

channel prediction/tracking mechanisms provided herein may not be feasible with the information gathered only from a location estimation process of an access point. This high accuracy of the location (or changes in coordination), for example, provides a unique scenario for wireless VR/AR headset to make the proposed algorithms feasible.

In examples described herein, the AP may collect the exact location of a headset, which has very high accuracy in short time intervals (every ~1mSec). The AP may then map the coordination to the index of look-up table that stored the channel vector or steering matrix corresponding to that point of space.

Figure 1 illustrates an example scenario in which a single VR user is present in a room with one multi-antenna AP and several fixed wireless clients. The AP is required to update/adjust its beamforming vectors to be able to send independent streams to each client with minimum interference and high throughput. As such, depending on the time coherency of the channel, the steering matrix should be updated every 5-25 mSec.
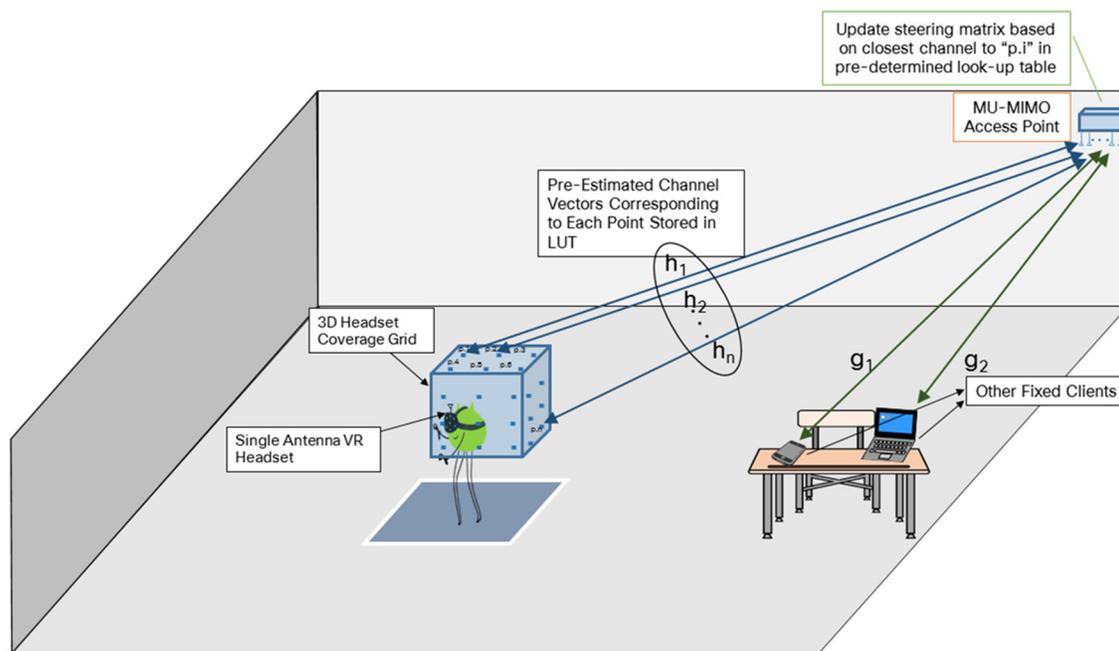

Figure 1

Techniques described herein minimize the need for sounding processes during the active session of VR interactive applications by splitting the beamforming update process into three different modes, namely:

- **Mode 1 - Generating the Look-Up Table**: When a VR user decides to start an interactive application, a short training session is considered with 3D visualization to guide the user to stop on spatial points that likely will be covered during the active session. With a short pause at each point, the sounding process is performed to estimate channels for a set of points that likely cover the head movements of the user. These channels are stored in a table at the AP. The i-th value of the table is $LUT(i) = h\_i$. In this equation, the term $i = L(x(t), y(t), z(t))$ and it is an index of a point at location $(x(t), y(t), z(t))$ in time t . The term h_i is the channel vector size Nrx x 1. Nrx is the number of antennas at the AP. Instead of storing channel vectors, the access point can pre-process channel vectors and can store the steering matrices. This may be done to ignore the overhead of re-processing.

- **Mode 2 - Sounding-less Transmission**: In this mode, the steering matrix is updated based on the channel extracted from the LUT of Mode 1. Some current commercial VR products measure and send the coordination of VR every 1mSec. This coordination is extracted in the AP and the index is calculated based on the closest point to this coordination in the LUT. If the coordination of the user is $(x(t'), y(t'), z(t'))$, the AP will first find the index $i = L(x\_i; i = min(x(t')-x\_i), y\_i; i = min(y(t')-y\_i), z\_i; i = min(z(t')-z\_i))$. The AP then finds the channel vector corresponding to this index. The steering matrix can be calculated based on the predicted channel or, if the steering matrix has been already pre-processed and stored in LUT, it can be directly used.

- **Mode 3 - Update Look-Up Table**: Depending on the stability of objects in the environment, the channel vectors may be updated after a certain amount of time. If the performance is degraded, the AP may determine to adjust the LUT based on implicit sounding. This will not be as frequent as the updating mechanism in conventional MU-MIMO (which is every 5-25mSec). Instead, it would generally only occur if a significant change in the environment occurred.

By employing these mechanisms, the beamformer can be updated based on the predicted CSI with no need to transmit pilot packets (sounding) at least for a short time period. Accordingly, the packet-error rate (PER) will be reduced and there will be lower

5638X

overhead to increase the data rate. These improvements can significantly reduce the latency of transmission.

Figure 2, below, illustrates a comparison of an example of conventional beamforming and an example of a predictive channel estimation method. The channel sounding is performed at the beginning of generation of the LUT and updating will be much less frequent than in conventional arrangements.
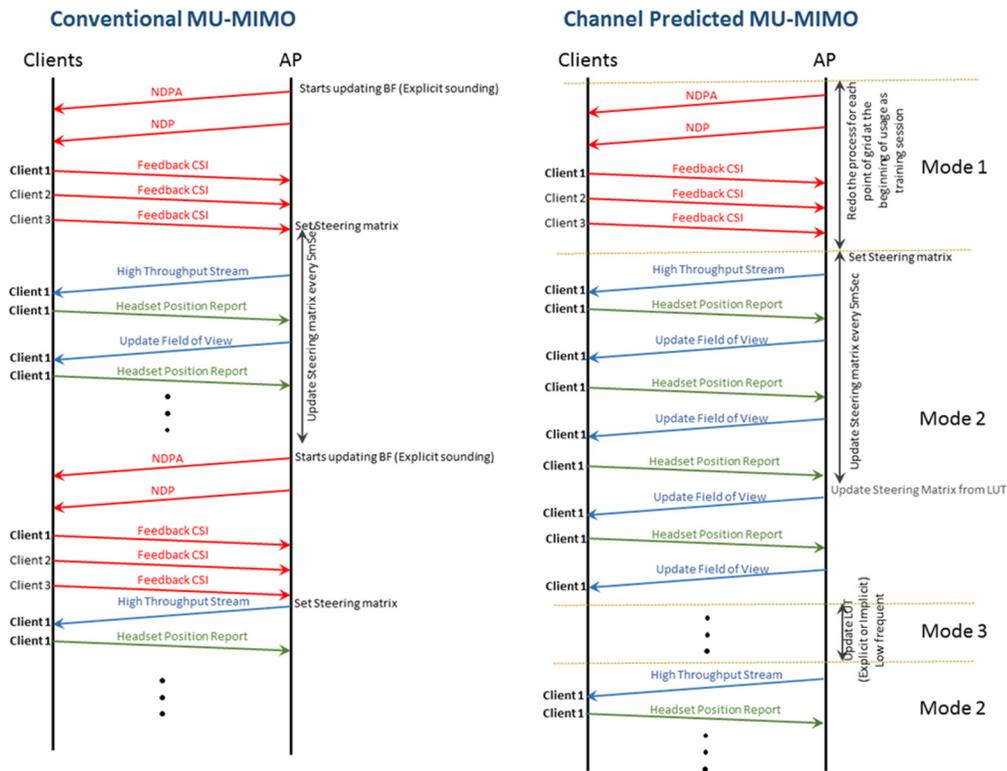

Figure 2

5

5638X