**Source: Huawei**

**Title: On Network Based Video Super Resolution**

**Document for: Discussion and Agreement**

**Agenda Item: 9.7**

1. **Introduction**

This contribution discusses the benefit of AI base Super Resolution solutions in enhancing video quality, and proposes to consider such techniques in FS\_AI4Media as a use case.

1. **Discussion**

With the tremendous success of mobile internet services over the past decade, user demand on wireless data is ever increasing. Popular applications like Youtube, online music, online news, audio/video conferencing, media sharing, etc., are driving this demand up. However, market surveys show that this demand is by far asymmetrical, users need more downlink (DL) data capacity than uplink (UL) data capacity, even though there are applications like on-line meeting and video calls that depend on symmetrical network capacities.

Due to this UL/DL imbalance in demand on network capacity from users, operators provision their network build out to meet this demand profile in order to minimize deployment cost. Consequently, within the foreseeable future, this imbalance in network planning and availability of capacity will remain unchanged. There is roughly a factor of 4 or more between DL and UL capacities.

Table 1: Example of Deployed Uplink and Downlink Resources

|  |  |  |
| --- | --- | --- |
|  | 4G | 5G |
| TDD | FDD | TDD | FDD |
| Uplink Data Rate at the edge of Single Carrier | 256Kbps | 1Mbps | 3~5Mbps | --- |
| Downlink Data Rate at the edge of Single Carrier | 1Mbps | 4Mbps | 100Mbps | --- |
| Peak Uplink Network Speed | 568Mbps | 2Gbps | --- |
| Peak Downlink Network Speed | 1.2Gbps | 5Gbps | --- |

Note1: The above data from the telecom operator basically reflect the characteristics of wireless resources of base stations.

Note2: The corresponding data in FDD of 5G has not yet been released.

While this imbalanced deployment profile fits well the current wireless data traffic characteristics, future HD video enabled conversational applications like HD video calls or AR/MR conversations will drive the need for a more balanced symmetrical network capacity. When those new services are deployed and broadly used, the network demand will tend to follow a balanced profile. This miss-match between operator deployed imbalanced network capacities vs. a balanced user demand profile will contribute to more and more network congestions especially for the UL direction.

Dynamic rate adaptation can be an effective remedy to control network congestion, but that is typically achieved by reducing video resolutions, therefore sacrificing user perceived service quality.

How to preserve the service quality while dropping the video resolution? The answer could well be by introducing advanced techniques like AI based super resolutions. Below table shows test results that support this line of thoughts, i.e., instead of dropping 9dB PSNR from 1080p to 540p, with AI Super Resolution applied to the 540p to “rebuild” 1080p, the reduction in quality is only around 4dB PSNR that seems to be more acceptable.

Table 2: Video Quality Comparison

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Anchor:1080p without compression | 1080p@2mbps(H.265) | 540p@500kbps(H.265) )=>1080p (Linear interpolation) | 540p (H.265)=>1080p(Bicubic Super-Res) | 540p (H.265)=>1080p(AI Super-Res) |
| PSNR | ~36dB | ~27dB | ~29dB | ~32dB |

Note: PSNRs are calculated based on output from Super Resolution, without re-encoding. Additional artifacts/degradation in quality are typical when transcoding is applied. Also, transcoding latency increase can be another quality issue that needs to be considered in overall performance assessment.

Test results also show that although video quality can be improved even with simple algorithms like Bicubic, AI based Super Resolution solutions are more capable. Another potential benefit to deploy Super Resolution solutions is to improve video quality when the transmitting UE is not capable of sending out higher resolution contents. In such a case, it is a net gain by using Super Resolution solutions.

1. **Deployment challenges**

Unfortunately, AI based solutions need in general much higher computational complexities, making them hardly implementable on handheld/head-mounted devices due to cost, power consumption and heat dissipation. It is therefore necessary to investigate how to apply Super Resolution techniques in the network while minimizing the impact.

The following areas might need further investigations:

1. Investigate and document the value of Super Resolution in achieving better congestion control, image/video quality enhancement, and/or preservation of video quality for Split-Rendered scene re-encoding;
2. Investigate the best algorithms available for network based Super Resolution applications;
3. Determine the constraints and impact to the network architecture when deploying Super Resolution solutions in the network;
4. Investigate more appropriate video quality assessment methodology and investigate achievable minimum performance requirement to justify the deployment of Super Resolution solutions;
5. Investigate and recommend suitable media negotiation processes to support integration of Super Resolution solutions.
6. **Proposal**

It is proposed to include the network based Super Resolution as a use case in the FS\_AI4Media study and work out recommendations on how to minimize its deployment impact to current network.

The following section is proposed to be included in the FS\_AI4Media PD:

4.5 Network based video AI super resolution

Due to the current wireless network build-out, there is an imbalance in the UL/DL capacity provision: smaller capacity for UL, and larger capacity for DL. As a result, network congestion can happen more easily in the UL direction, and dropping the resolution for video stream is usually the remedy to ease network congestion, therefore sacrificing user experience. Also, for some legacy UEs that have limited resolution capabilities, the resulting video based services also lack up to date qualities. Deploying network based video AI super resolution techniques can effectively address these issues. Figure 4.5 shows a scenario for network based video AI super resolution.



Figure 4.5 Network based video AI super resolution

Table 4.5 shows an example of test results for AI based video super resolution; instead of dropping 9dB PSNR from 1080p to 540p, with AI Super Resolution applied to the 540p to “rebuild” 1080p, the reduction in quality is only around 4dB PSNR which seems to be more acceptable.

Table 4.5: Video Quality Comparison

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Anchor:1080p without compression | 1080p@2mbps(H.265) | 540p@500kbps(H.265) )=>1080p (Linear interpolation) | 540p (H.265)=>1080p(Bicubic Super-Res) | 540p (H.265)=>1080p(AI Super-Res) |
| PSNR | ~36dB | ~27dB | ~29dB | ~32dB |

Note: PSNRs are calculated based on output from Super Resolution, without re-encoding. Additional artifacts/degradation in quality are typical when transcoding is applied. Also, transcoding latency increase can be another quality issue that needs to be considered in overall performance assessment.

Test results also show that although video quality can be improved even with simple algorithms like Bicubic, AI based super resolution solutions are more capable. Another potential benefit to deploy super resolution solutions is to improve video quality when the transmitting UE is not capable of sending out higher resolution contents. In such a case, it is a net gain by using super resolution solutions.

For this use case, the network aspects which may need to be considered include investigating the constraints and impacts to the network architecture, and the media negotiation procedures to support the integration of AI super resolutijon solutions in the network.-