

Video Call Performance Testing

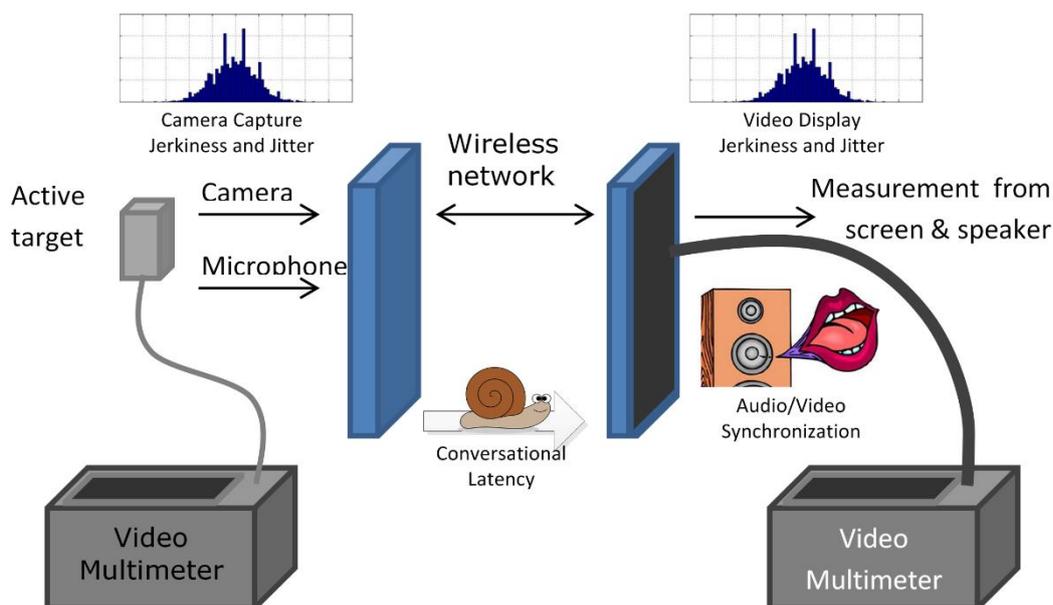
Video calling – What kind of impression do you give of yourself?

Select your video calling and sharing applications with care! We measured the performance of video call apps of the three major ecosystems (Apple® FaceTime®, Google® Hangouts™, Microsoft® Skype®), and one rival from China (Tencent WeChat).

After reading this article you will know a whole lot more about measuring the video playback performance, and how well the different ecosystems' video call apps perform both in a varying network environment and in different devices.

How did we carry out the measurements?

It is a hard task to control video call quality from end to end. In addition to spatial video quality, a group of temporal quality aspects has to be controlled: video jerkiness, jitter, conversational latency and lip synchronization, which together form the basic set of KPI's (Key Performance Indicator). Jerkiness and jitter are the cumulative result of both captured and displayed images, meaning that unevenly captured images in the sending party will cause a poor user experience, despite smooth video playback in the receiving party.



OptoFidelity has developed a novel measurement technology, Video Multimeter, which enables quick and accurate video call measurements, end-to-end. In our Video Multimeter solution we use controlled visual and audible targets, which are presented for the sending party. In the receiving party, Video Multimeter collects the visual and audio data non-intrusively.

OptoFidelity Video Multimeter outputs the results as a MOS (Mean Opinion Score) whose scale is from 1 to 5. A value of 1 means bad, unacceptable performance, and a value of 5 means excellent, flawless performance. The scaling of Video Multimeter MOS is based on well-known industry standards, suggestions and research studies.

Scores	Quality	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible, but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying, unwatchable

In the following sections we refer always to the MOS values achieved by tested applications. Here below are the results in a nutshell. The MOS values presented in the table are so-called 'composite' values, which represent the weighted average of the jerkiness, jitter, conversational latency and audio/video synchronization.

We have visibility to individual MOS values, frame jitter for example. From the MOS values it is possible to drill down to detailed measurement data because frame and audio timestamps are recorded and stored in microsecond resolution.

Video Call MOS (Wi-Fi Connectivity)					
	iPhone+iPhone	iPhone+Android	Android+Android	AVG Score	Rank
FaceTime	3,6			3,6	1
Skype	3,4	3,5	2,4	3,1	2
Hangouts	3,6	3,7	1,9	3,1	2
WeChat	3,6	2,0	2,7	2,8	4
Wi-Fi MOS AVG:				3,1	

Video Call MOS (Cellular Connectivity)					
	iPhone+iPhone	iPhone+Android	Android+Android	AVG Score	Rank
FaceTime	3,8			3,8	1
Skype	3,1	2,1	2,1	2,4	2
Hangouts	2,4	2,2	2,4	2,3	3
WeChat	2,4	2,5	1,5	2,1	4
Cellular MOS AVG:				2,7	

Terminology

Jerkiness = Average FPS (Frames Per Second) of the video playback

Jitter = Standard deviation of video playback frame intervals

Conversational latency = End to end audio and video latency, sometimes called 'mouth to ear' delay

Audio/video synchronization = The delay between the user-experienced video and audio stream, sometimes called 'lip sync'

Apple FaceTime is not king - Google Hangouts and WeChat get even in the IOS® environment

When measuring the video call performance in the IOS-environment, and specifically when using Wi-Fi connectivity, Google Hangouts and WeChat get the same performance score (3.6), meaning that the performance is somewhere between fair and good. When drilling down to the root cause the main issue with Apple FaceTime is the video jitter. Google Hangouts and WeChat have also some additional issues with the audio/video synchronization.

Microsoft Skype has a tight battle with Google Hangouts

When looking at the average MOS scores Microsoft Skype and Google Hangouts are very close to each other (3.1 in Wi-Fi). Quite interestingly Google Hangouts has a hard time achieving a decent performance in a pure Android™ platform environment (only 1.9 in Wi-Fi). One could easily think that Google could do better because it is the Android platform ecosystem owner.

Microsoft Skype ranks the #2 in overall performance. In Wi-Fi connectivity it shares the rank with Google Hangouts, and in cellular connectivity it beats Google Hangouts slightly.

Quality rival from China - WeChat surprisingly good in the IOS environment

WeChat is a very popular online chat and video calling application, especially in China. It is believed to have over 1.1 billion users, so it is the most popular video calling application in the whole world!

Although having the last place in the overall ranking WeChat performs very well in the IOS environment with Wi-Fi-connectivity. Performance in Wi-Fi is very important because video calls are very data consuming, and typically users seek a place to hook up to a Wi-Fi hotspot instead of using expensive cellular data.

As mentioned earlier WeChat has some issues with frame jitter and audio/video synchronization. Users may perceive relatively poor performance if WeChat is used in the Android platform environment with cellular connectivity.

What would you like to see measured in the future?

OptoFidelity Video Multimeter is able to measure almost any local or streaming video playback performance. New video services are arising - you probably have already heard, and possibly used as well, these: Periscope, Facebook® video sharing, Snapchat™ 2.0 etc. One should be very critical about the video performance and quality in these services because they affect the impression that people try to give when making personal videos. **No matter how good is the video content, a bad video performance can ruin the whole message.**

Test details

Mobile phones used in the testing:

- IOS: iPhone® 6, iPhone® 6S
- Android platform: LG® G3, Huawei® Y5
- All phones had the latest OTA (Over The Air) software updates

Cellular networks used in the testing:

- 4G from DNA Finland
- 3G from Sonera Finland
- H+ from Saunalahti Finland
- Reception quality was at least 4/5 in all cases
- Networks were all public

Wi-Fi network used in the testing:

- 802.11 a/b/g/n/ac, connected to 1 Gbps broadband
- Reception quality was 5/5 all times
- There was random traffic at all times (typical office traffic)

For each test case a 20 second video call was measured twice, and an average was taken from those results.

All trademarks belong to their respectful owners.