FIFA

Handbook of Test Methods for Video Assistant Referee Systems

January 2020

Contents

1. Introduction

Video Assistant Referees (VAR) was approved by The IFAB to be used, in accordance with the Laws of the Game, to support referees in case of a clear and obvious error. Video Assistant Referees (VAR) who have access to all broadcast feeds may assist the referee team in the following cases:

- Goals
- Penalties
- Red card offenses
- Mistaken identity

When it comes to implementing a VAR system three key components must be considered: the system set-up, the basic performance requirements of the VAR system, and the training of VAR operators. The FIFA Quality Programme has developed a testing and certification scheme which focusses specifically on the basic performance requirements of a VAR system. This manual is solely designed to provide technical information on the underlying VAR technology and should be understood as one part of the comprehensive IAAP processes.

2. Background and Approach

In order to establish technical guidelines for Video Assistant Refereeing (VAR) systems that wish to be approved for use a research project was undertaken by RISE Research Institutes of Sweden AB on behalf of FIFA. The project aimed to identify key areas of technical importance for VAR systems, establishing objective test methods to quantify these factors and set limits that ensure a tested and approved system is fit for use.

Initial discussions identified various challenges linked to coding, decoding, synchronizing, re-formatting broadcast feeds and the processing of images. Three measurement points (MP) were defined, see Figure *1*. MP 0 is where the camera signals enter the Video Operating Room (VOR), MP 1 is located after the video server in the VOR (i.e. the VAR system) and MP 2 is where the video is sent back to the Outside Broadcast (OB)-van or to the broadcast provider.

Various challenges were identified along this workflow:

- Measurement of time synchronicity of broadcast images between different cameras (immensely important for offside decisions) at MP1
- Conversion and integration of different formats (1080, ultra-motion cameras, varying frequencies & formats) and image sources into a single system and the quality of the resulting output at MP1
- Measurement of absolute latency of processed images vs. "live" feed at MP1
- Measurement of the output video quality from a VAR system back to the broadcaster for transmission on air at MP2

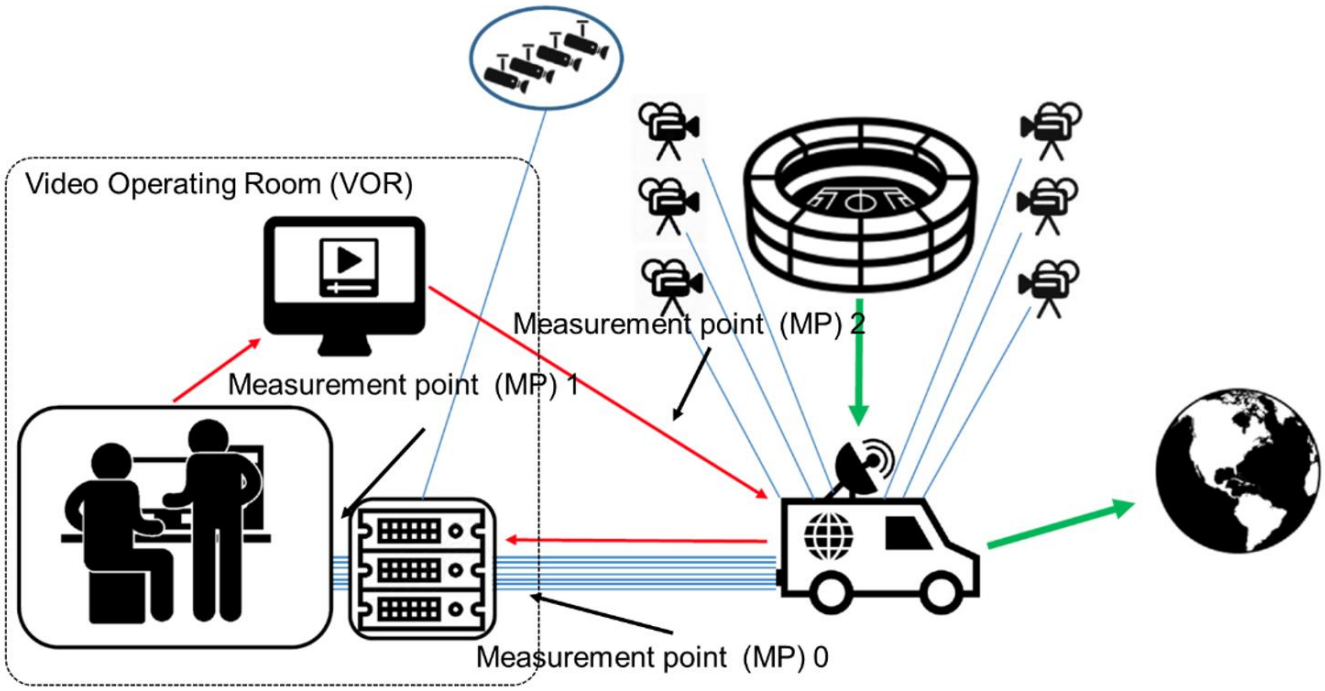

Figure 1: Schematic overview of the VAR setup. Three measurement points (MP) are indicated MP0, MP1 and MP2, for the evaluation of latency, synchronization and video quality.

As a result the basic performance requirements are based on the following:

- Synchronicity of video feeds.
- Latency of VAR system video feeds vs. broadcasting video feeds.
- Objective video quality assessment of the output of VAR-systems.

Therefore, the test methods described in this handbook include:

- Measurement of synchronicity and latency of feeds by analysis of the content in the video feeds.
- Measurement of synchronicity and latency of feeds by using optical detectors on monitors.
- Measurement of video-quality of VAR system feeds using sample video of known and varying quality

3. Test Protocol

With regards to the Latency tests, 2 tests are being conducted in the testing with a decision on the final test to be made after the collection of sufficient empirical evidence.

3.1. Synchronicity Test

3.1.1. Test setup

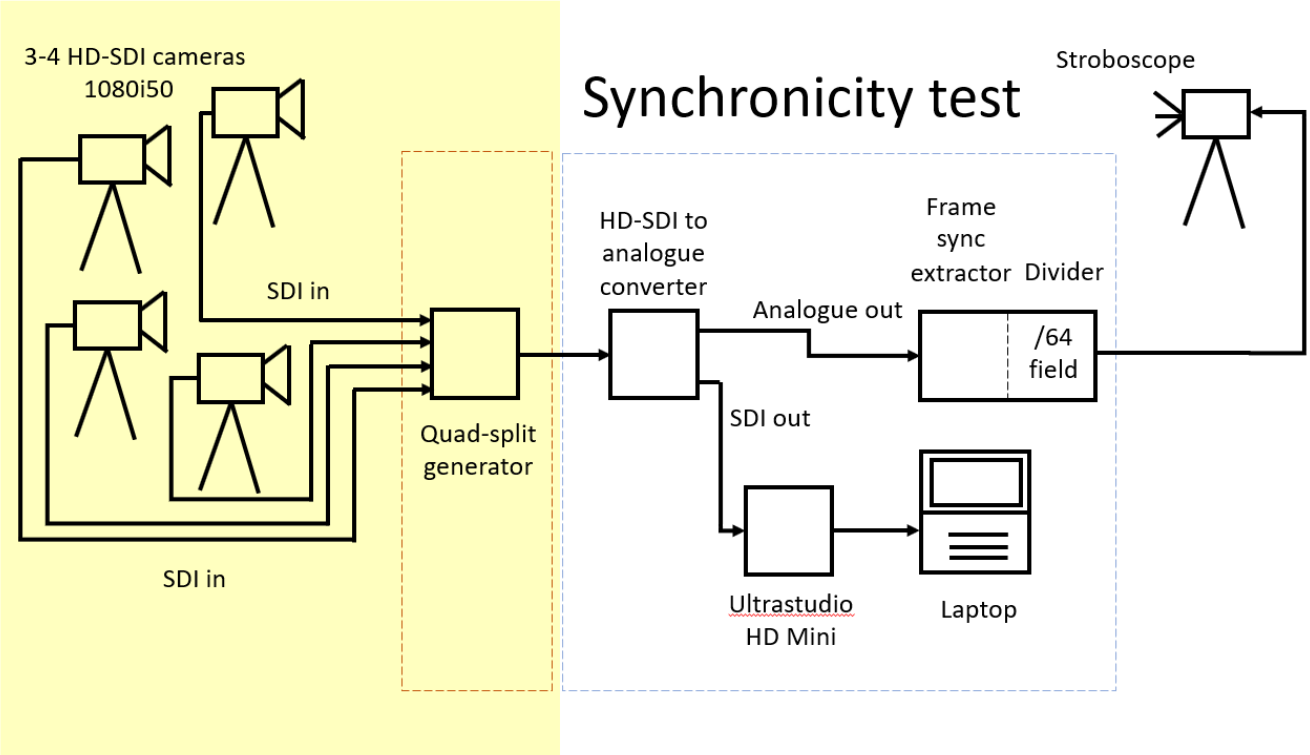

Figure 2: Synchronicity test between channels in a quad-split.

The following equipment shall be used for assessment of synchronicity between video feeds:

- 4 HD-SDI cameras capable of filming in both 1080i and 1080p
- A quad-split generator.
- An HD-SDI to analogue video converter, e.g. Blackmagic 'Mini Converter SDI to Analog'.
- An analogue frame synch extractor.
- An SDI video grabber, e.g. Blackmagic 'UltraStudio HD Mini'.
- A lap-top computer, optional.
- A stroboscope, e.g. CheckLine 'QBS-LED' with an optional tripod.
    - The stroboscope needs to be adjustable in intensity and in delay between the trig signal to the stroboscope and the flash.

3.2. Methodology

The test set-up (Fig. 2) is set up in a pitch side location where cameras have a clear line of sight to the stroboscope.

A flash rate of ~1 flash per second is selected which means that a division rate of 64 of the analog field/frame sync signal works for most application, however the test method is not strictly dependent on the flash rate. The test method can also be used with a free running stroboscope, although a synced stroboscope to the video feeds is recommended.

The stroboscope is placed where all the cameras in the test can see it, note that the suggested stroboscope only has a visibility range of 90° which means that not all cameras will be able to focus at one time. It is also recommended to ensure the stroboscope is set in a position with clear sight to all cameras as the analysis algorithm can get confused by sudden changes in the feed.

Once in place the flash intensity and delay is then adjusted for a proper flash strength in the video feeds. When complete a video of the quad-split feed is grabbed with the video grabber. A 1 minute-long sequence is normally sufficient for analysis, however, to ensure sufficient usable footage is collected a 5 minute-long sequence will be recorded.

The video analysis algorithm reads the captured files and presents in which field/frame the flashes occurs for the four video feeds in the quad split signal.

Latency Test 1

3.2.1. Test setup

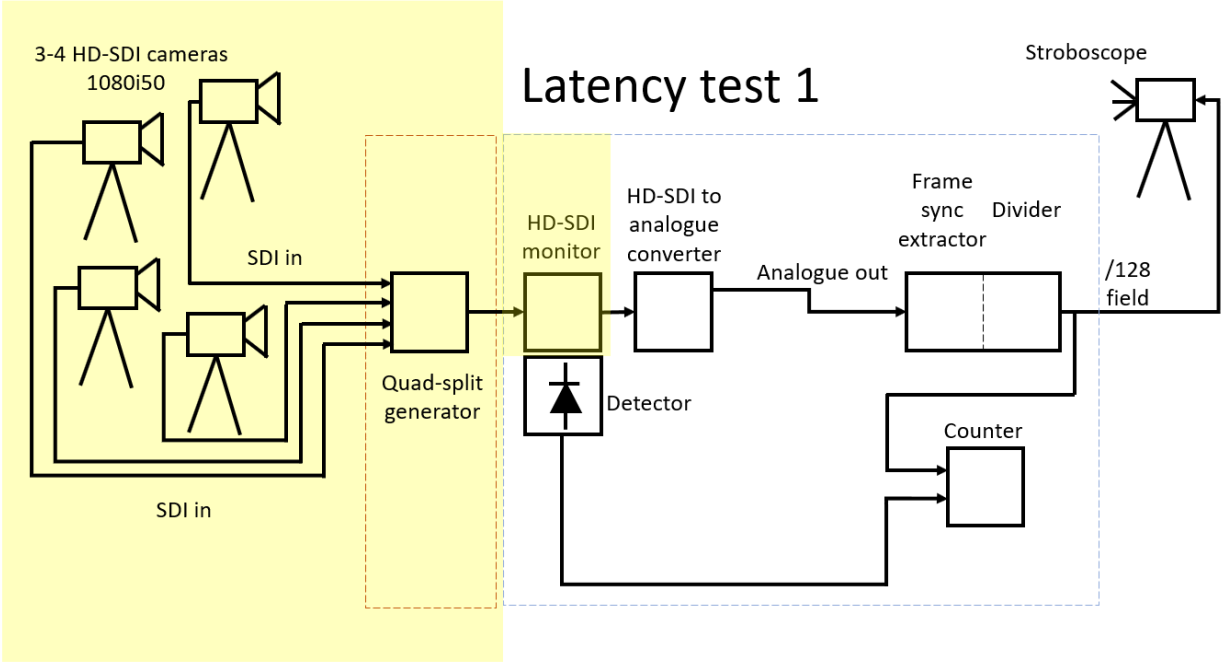

Figure 3: Latency test between trig signal and stroboscope flash in a monitor.

The following equipment shall be used for assessment of latency between trig signal and monitor response:

- 4 HD-SDI cameras capable of filming in both 1080i and 1080p.
- A quad-split generator.
- An HD-SDI to analogue video converter, e.g. Blackmagic 'Mini Converter SDI to Analog'.
- An analogue frame synch extractor.
- An SDI video grabber, e.g. Blackmagic 'UltraStudio HD Mini'.
- A lap-top computer, optional.
- A stroboscope, e.g. CheckLine 'QBS-LED' with an optional tripod.
  - The stroboscope needs to be adjustable in intensity and in delay between the trig signal to the stroboscope and the flash.
- An optical detector, e.g. ThorLabs 'PDA36A2 Si Switchable Gain detector'.
- A counter, e.g. Rhode & Schwartz 'HM8123'.

3.2.2. Methodology

The test set-up in Fig. 3 is used. A flash rate of ~0.5 flash per second is set on a stroboscope which is synched to the video feeds. The stroboscope is placed were all the cameras in the test can see it. Flash intensity and delay is adjusted for a proper flash strength in the video feeds. The trig signal is connected to input 'A' on the counter. The optical detector is attached to the monitor/video feed under test and the output signal from the detector is adjusted to ~1 V. This signal is then connected to input 'B' on the counter. The Rhode & Schwartz' HM8123 is used to measure the time interval between input 'A' and input 'B' and the time difference is recorded from the counter.

3.3. Latency Test 2

3.3.1. Test setup

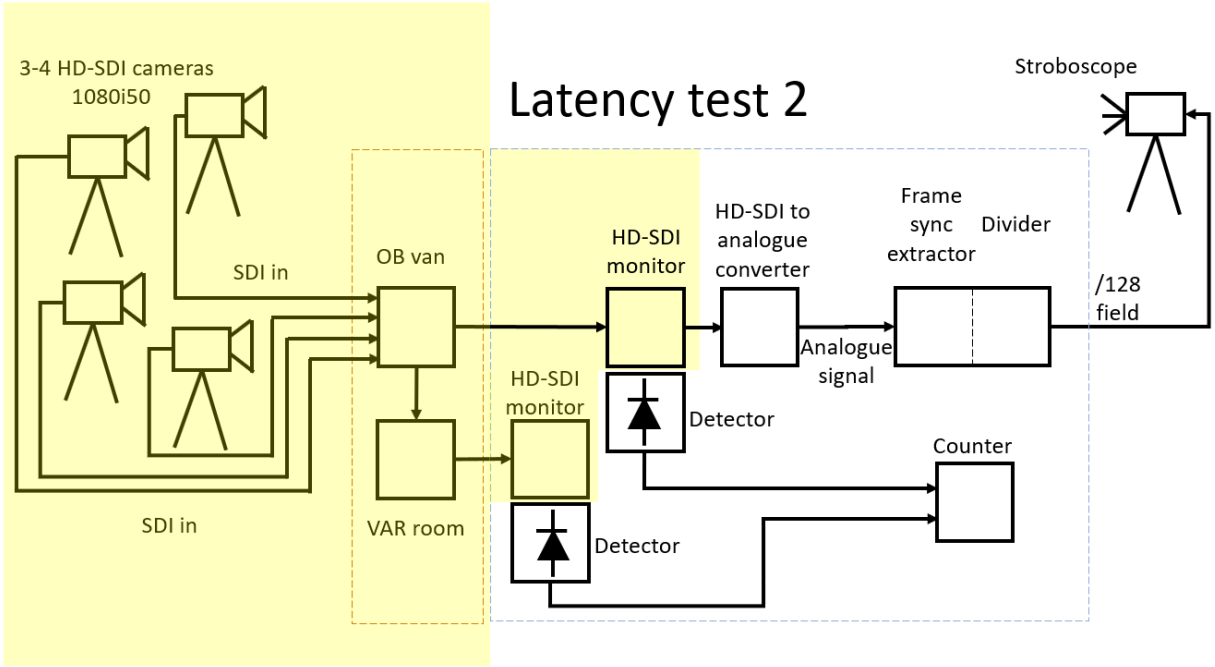

Figure 4: Latency test between stroboscope flashes between two monitors.

The following equipment shall be used for assessment of latency between two monitors:

- 4 HD-SDI cameras
- A quad-split generator.
- Two HD-SDI monitors.
- An HD-SDI to analogue video converter, e.g. Blackmagic 'Mini Converter SDI to Analog'.
- An analogue frame synch extractor
- A stroboscope, e.g. CheckLine 'QBS-LED' with an optional tripod.
    - The stroboscope needs to be adjustable in intensity and in delay between the trig signal to the stroboscope and the flash.
- Two optical detectors, e.g. ThorLabs 'PDA36A2 Si Switchable Gain detector'.
- A counter, e.g. Rhode & Schwartz 'HM8123'.

3.3.2. Methodology

The test set-up in Fig. 4 is used. A flash rate of ~0.5 flash per second is set on a stroboscope which is synched to the video feeds. The stroboscope needs to be adjustable in intensity as well as having an adjustable delay between the trig signal to the stroboscope and the flash. The stroboscope is placed were all the cameras in the test can see it. Flash intensity and delay is adjusted for a proper flash strength in the video feeds. The optical detectors are attached to the monitor/video feed under test and the output signal from the detector is adjusted to ~1 V. This signal is then connected to input 'B' on the counter. The Rhode & Schwartz' HM8123 is used to measure the time interval between input 'A' and input 'B' and the time difference is recorded from the counter.

3.4. Video Quality

3.4.1. Test setup

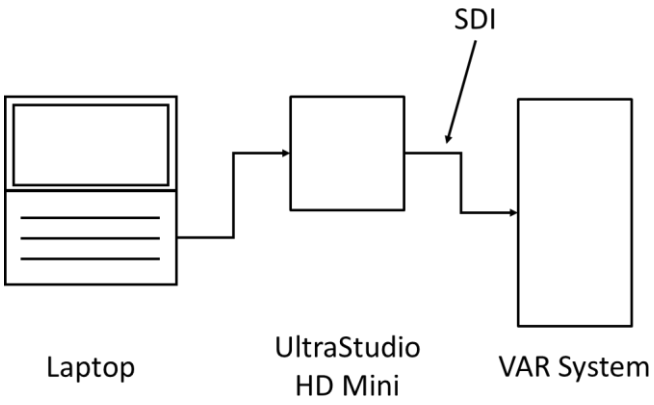

Figure 5: Set-up for ingesting test video into a VAR system

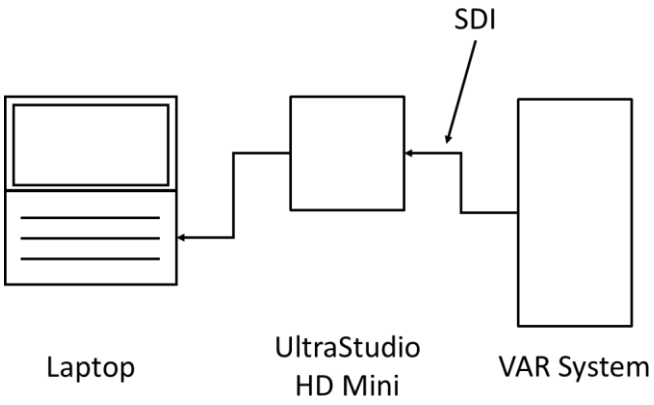

Figure 6: Set-up for recording test video from a VAR system

The following equipment shall be used for assessment video quality

- A laptop with similar specifications to the ASUS GX501 (with Intel i7-7700HQ@2,8GHz, 16 GB RAM), with Thunderbolt.
- Playback software e.g. BlackMagic MediaExpress
  - Ingesting: BlackMagic MediaExpress (playback mode)
  - Recording: BlackMagic MediaExpress (Log and Capture mode)
- SDI generator e.g. BlackMagic UltraStudio HD Mini
  - SDI output from box connected to VAR-system at ingesting
  - SDI input on box when recording.

3.4.2. Methodology

The video quality testing is based on ingesting a known uncompressed video into the VAR system via SDI and then replay this video from VAR system and record it when it is sent out via SDI, see Fig. 5.

- Ingesting
  - Connect UltraStudio HD Mini SDI output to SDI-input from VAR-system
  - Play test video in VAR-system BlackMagic MediaExpress (playback mode)
- Playback and recording
  - Connect UltraStudio HD Mini SDI input to SDI-output from VAR-system
  - Play test video in VAR-system
  - Record the output with BlackMagic MediaExpress (Log and Capture mode)
- The output video is then assessed using the Video Multi-method Assessment Fusion (VMAF) a perceptual video quality assessment algorithm developed by Netflix.

4. Results & Evaluation

4.1. Synchronicity & Latency Evaluation

For a synchronicity test, the maximum value for the absolute difference in field/frame number between any of the cameras (i.e. measured for the same flash) during the test period shall counted as the synchronicity value. An example table is included below which illustrates in which fields/frames the flash occurs, the synchronicity value between the four cameras is set to 3 fields/frames. The term field is used for interlaced videos. Interlaced technique is used to reduce the bandwidth while keeping the same perceived frame rate. In an interlaced video each frame contains two fields, captured at different times. Normally, the odd lines contain the first field and the even lines contain the second fields. An i50 video thus contain 25 frames per second which is interpreted as 50 fields per second. A p50 video contain 50 full frames per second.

Analysis of the synchronicity and latency test will be assessed by measurement values, either by running a software analysing the captured quad-split vide, or by reading the values presented on the counter an example is displayed in Table 1. The values represent the maximum difference of these frames during the quad split video. A value of +1 means that the flash appears one field/frame later than the flash in the reference camera, -1 means that the flash appear 1 filed/frame earlier than the flash in the reference camera.

For synchronicity, all values must be <3 fields/frames for FIFA Quality Pro.

No. filed/frame distance relative to reference camera	Camera 1 (ref.)	Camera 2	Camera 3	Camera 4
Filed/Frame	0	+1	+1	-1
Filed/Frame	0	+1	-1	-2
Filed/Frame	0	+1	-1	-1
…	…	…	…	…
Filed/Frame	0	+1	+1	-1

Table 1: Example of measurement result for a synchronicity measurement

For the latency test, the maximum value read during the test period shall be considered as the measurement value. Latency measurements are generally a delay measurement between a mark event, e.g. trig signal is sent to the stroboscope or that a stroboscope flash detected at one screen, and the time that the mark appears on the screen chosen for the latency measurement.

For latency the limit is <0.5 s for FIFA Quality Pro.

4.2. Video Quality Evaluation

After a detailed exploration of possible analysis methods the Video Multi-method Assessment Fusion (VMAF) was selected as the base of the video quality test. As mentioned earlier, the VMAF is a perceptual video quality assessment algorithm developed by Netflix. VMAF follows a machine-learning based approach to first extract a number of quality-relevant features from a distorted video and its reference full-quality video, followed by fusing them into a final quality score using a non-linear regressor.

The video quality will be evaluated on seven 14 s test videos. The evaluation will be done by comparing the quality of each individual 14s video, before and after the ingestion. In order to avoid temporary glitches affecting the results, the ingestion and grabbing will be repeated three times.

The requirements are to obtain:

- Average Mean Opinion Score (MOS_i) ≥ 4 on all test videos
- Min (MOS_i) ≥ 3 on all test videos
- The above two requirements should be fulfilled on two or more of the three set of ingested videos.

The MOS requirements will be verified by the following VMAF scores on the individual test videos.

- **1080p:** Mean ($VMAF_i$) ≥ 92 and Min ($VMAF_i$) ≥ 85
- **1080i:** Mean ($VMAF_i$) ≥ 85 and Min ($VMAF_i$) ≥ 75