



Quality measurement methods for video assisting refereeing systems

Kjell Brunnström^{1,3} · Anders Djupsjöbacka¹ · Oskars Ozolins^{1,4} · Johsan Billingham² · Katharina Wistel² · Nicolas Evans²

Accepted: 23 February 2023 / Published online: 10 March 2023
© The Author(s) 2023

Abstract

Changes in the footballing world's approach to technology and innovation, along with major advancements in broadcasting contributed to the decision by the International Football Association Board to introduce Video Assistant Referees in 2018. The change meant that under strict protocols referees could use video replays to review decisions in the event of a “clear and obvious error” or a “serious missed incident”. At the time of writing 48-Member Associations have introduced the Video Assistant Referees protocol in at least one of their tournaments and there are many technology providers who work with organisers to implement the Video Assistant Referees systems. To ensure that the use of Video Assistant Referees has a positive effect on the game, Fédération Internationale de Football Association collaborated with the RISE Research Institutes of Sweden to develop objective test methods that could be used to ensure that a system can provide an adequate solution. Each provider must be able to pass requirements that ensure that they can deal with the challenges of processing, coding, decoding, synchronising, and re-formatting of the broadcast feeds. This article will describe the development of the test methods and illustrate some initial results from a test event on Video Assistant Referees system candidates. The methods have shown to be robust and appropriate for their intended purpose and will be developed over the years to ensure the quality of Video Assistant Referees. The developed measurement methods are general and can be applied to other broadcast and video systems as well as to other sports.

Keywords Video assistant referee (VAR) · Football · Video quality · Latency · Synchronicity · Measurements · Broadcast

1 Introduction

Video officiating technology is used in a number of different sports e.g., Rugby, American Football, Volleyball, Handball, Hockey and Ice Hockey. For all sports, the video quality is equally important as in the game of football to

make the correct decisions while consulting the video referee. Advancements in technology over the past decade have afforded a significant number of innovations within the area of broadcasting. In the 2018 FIFA World Cup Russia™ for the first time in history, Video Assistant Referee (VAR) was implemented in every game in the tournament. This meant that fans were not the only ones to benefit from the broadcasting advancements, but referees would be able to use the technology to prevent “clear and obvious errors” or “serious missed incidents” [1, 2]. In addition to the four on-field referees, four additional referees were positioned in the Video Operation Room (VOR) to check potential incidents of four game-changing decisions—goals, penalties, red cards, and mistaken identity. With this implementation came a large host of technology providers offering VAR services to leagues.

To ensure that technology providers were able to provide the VAR services to an acceptable level, Fédération Internationale de Football Association (FIFA) collaborated with the RISE Research Institute of Sweden to develop objective test

This article is a part of Topical Collection in Sports Engineering on Football Research, edited by Dr. Marcus Dunn, Mr. Johsan Billingham, Prof. Paul Fleming, Prof. John Eric Goff and Prof. Sam Robertson.

✉ Kjell Brunnström
kjell.brunnstrom@ri.se

¹ RISE Research Institutes of Sweden AB, Stockholm, Sweden

² Football Technology Innovation Subdivision, Fédération Internationale de Football Association (FIFA), Zurich, Switzerland

³ Mid Sweden University, Sundsvall, Sweden

⁴ KTH Royal Institute of Technology, Stockholm, Sweden

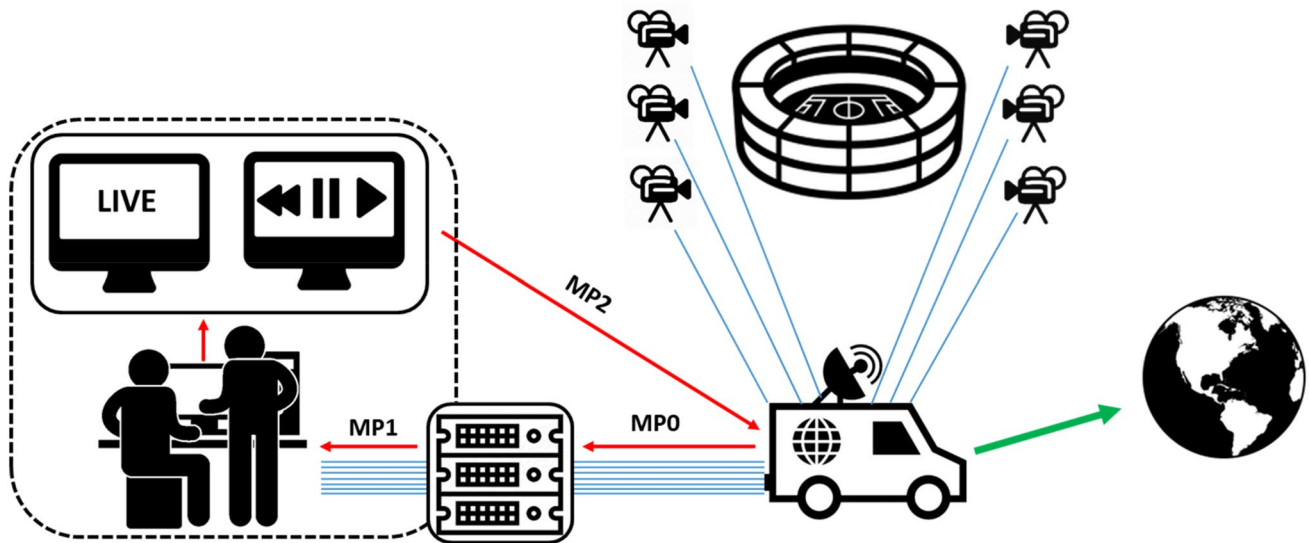


Fig. 1 Schematic overview of the VAR setup. Three measurement points (MP) are indicated MP0, MP1, and MP2, for the evaluation of latency, synchronisation, and video quality

methods that could be used to approve suitable technology providers for use. Figure 1 illustrates three key measurement points in the VAR workflow that would be used to conduct a variety of different tests on the quality of the VAR system. MP0 occurs before the broadcast signal reaches the VAR system. MP1 is what is visible on the VAR monitors. MP2 is the signal that is sent back to the broadcaster for broadcasting across the world.

The main concerns were:

- Measurement of synchronicity between a variety of different cameras at MP1.
- Measurement of latency in the VOR at MP1.
- Measurement of latency for the 3 s delayed feed at MP1.
- The video quality of the resulting output at MP1.

We present four tests aimed at quantifying the capability of VAR providers with regards to the main concerns outlined above. Two tests have been developed to measure the latency of feeds, one synchronicity test, aimed at assessing the ability of the provider to supply synchronised camera feeds from the broadcast. Finally, the video quality test aimed to quantify any degradation of video quality caused by the VAR system. Although, the methods were developed for football, the measurement methods are general and can be applied to other broadcast and video systems as well as to other sports.

2 Background

2.1 Synchronisation and latency

Many sports broadcasting situations are multicamera in nature and that is true for football as well. For example, the cameras must be synchronised when applying the use of VAR to the offside rule. The key factors that determine whether a player is offside are the position of the 2nd last defender and active leading attacker at the moment the ball is kicked. If the kick point and offside player are in two different camera angles and the cameras are not synchronised, then the position used to determine whether the player is offside will be incorrect.

Another time aspect is the delay (latency) of the video signal. In the VOR, a live screen and a delayed screen with about 3 s latency are used in the setup; this allows according to the VAR protocol the referees to shift their gaze from live-action to a replay of the live-action without requiring any interference with the system. Here, the latency measurement accuracy is not needed to be on field/frame level, a measurement accuracy of 0.1 s is sufficient, to accurately measure according to the requirements by FIFA [3]. In this case, traditional glass-to-glass (G2G) measurement will do, i.e., from the lens of the camera to the glass surface of the display [4]. This type of latency measurement is important not only in football, but is relevant in a number of other fields and applications, e.g., remote operation of machines [5] and digital rear-view mirrors.

2.2 Objective video quality algorithm

Objective video quality models are mathematical models that approximate results from the subjective quality assessment, in which human observers are asked to rate the quality of a video. Various models and approaches have been suggested in the literature [6].

To find a suitable algorithm, a user study with 25 invited video experts was conducted at RISE and several measurement methods were evaluated [7]. Although the top-performing model in this investigation was Video Quality Metric for Variable Frame Delay (VQM_VFD) [8], we decided to use the model that came second was Video Multimethod Assessment Fusion (VMAF) [9]. VMAF has very good performance and the software is more flexible and runs on more platforms, which will be important for VAR providers to be able to test their systems in preparation for an evaluation. The method of determining video quality is not depending on that the system is used for VAR but could be for any video or broadcast application. Published studies tends to only focus on the quality at the end users, but not at the contribution and production side as here.

3 Method

3.1 Latency measurements in video streams

A common method to measure latency for one video stream is to impose time information into it at the transmit side [10–13]. In the references, above the ‘reference mark’ relies on either colour coded pictures [10, 13] or on a European Article Number (EAN) code [11, 12] into the encoded video. There are also examples in literature based on a common clock and on a Quick Response (QR) code. By comparing the ‘reference mark’ with a running clock, the latency at the received side can then be measured.

All these methods suffer from the issue that the latency calculation relies on an image processing algorithm which can be difficult to implement into an unknown system. Published studies tends to deal with measuring delay for one video stream [4, 10–15], not synchronicity between multiple video streams. When synchronicity is mentioned, it is often referred to as the synchronicity between video and audio in one video stream [16].

The task, in this case, was measuring with a non-intrusive method the synchronicity between multiple video streams which is not generally studied in the literature. However, some papers give a hint of how such a measurement can be performed. In Jacobs [17] describes a non-intrusive system that measures glass-to-glass (G2G). In references [4] and [16], Jacobs’s method is further developed, but only for the measurement of the latency of one video stream.

3.1.1 Synchronous vs. asynchronous measurements

Most latency measurements presented in the literature are asynchronous measurements. Since we are going to measure latencies down to frame-level we prefer to perform synchronous measurements. This means that we need to have a clock that is frequency synchronised to the frame rate of the video stream. The easiest way to do this is to tap the Serial Digital Interface (SDI) video stream at some point, e.g., at the camera, at the input to the VAR-room’s SDI-server, at the output from the VAR-room’s SDI-server. Since the loop time for the video signal in the measurement setup is much longer than one frame, see Fig. 2, and since we want to have only one pulse in the loop at a time, the frame sync signal needs to be divided. In Fig. 2, we divide the extracted clock 64 or 256 times, depending on if we are measuring a delay less than 1 s or less than 4 s. We also like to keep the light flash within just one frame, and it is, therefore, important that we can adjust the delay (Δd) and the pulse width (Δw) at the stroboscope [18]. If a frame sync cannot be extracted, we need to apply asynchronous measurements which means that the light flash might be distributed (sometimes) over two frames which will lead to an added uncertainty in the measurement of one frame.

3.1.2 Latency and synchronicity measurement set-ups

The test equipment will be used in the field and often under unknown conditions, and hence we prefer a non-intrusive test method as well as both camera and display agnostic. Further, it shall not rely on any specialised hard- or software. We also prefer a test method that can measure both latency and synchronicity with essentially the same set of test equipment. We also anticipate that the captured image originates from one point in time, as for an image taken with a global shutter and not an image taken with a rolling shutter.

The external event introduced in the system that acts as a ‘time mark’ is here chosen to be a white flash. Since the flash can be 100 m away from the camera, we have chosen a Light Emitting Diode (LED)-based stroboscope with light flashes that are visible in the video streams at that distance (Check-line Europe QBS-LED’ equipped with 118 LEDs. Peak brightness of 10,000 lx). The in-frame delay of the flash, as well as its time duration, shall preferably be chosen so that it fits into just one field/frame of the video stream [10]. To measure the time delay between the ‘reference mark’ and the ‘detected mark’, we have chosen to use a counter, another way to do this is with an oscilloscope. In this case, we have here chosen a solution based on a counter which will display the time delay directly or can be read by a computer or both, see Fig. 2.

A more direct method is to measure the SDI signals directly which we do in the synchronicity measurement. In

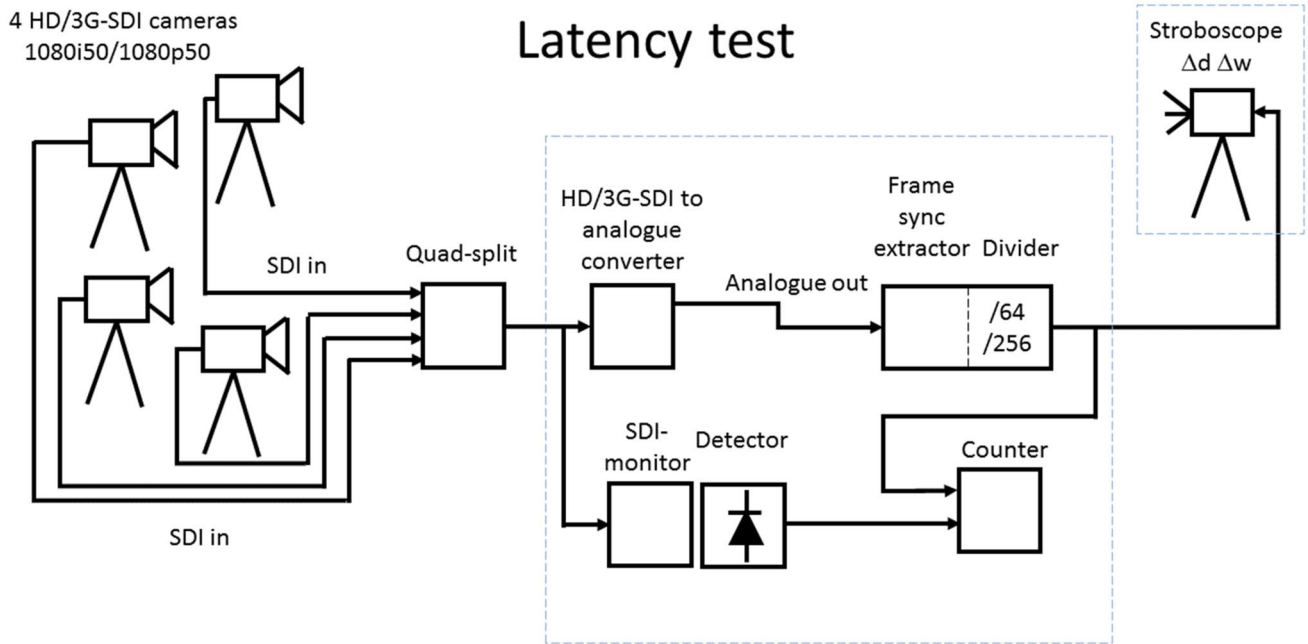


Fig. 2 The measurement setup for absolute latency measurements. A relative latency measurement is done with two monitors and two detectors. The dashed blue frames are the parts of the measurement setup. Stroboscope: check Line QBS-LED. SDI-to-analog converter:

Black Magic Design SDI to Analog 4 K. Frame sync & Divider: home-built frame-sync extractor and binary divider 1/64 & 1/256. Detector: Thorlabs PDA36A2. Counter: Rhode & Schwartz HM8132

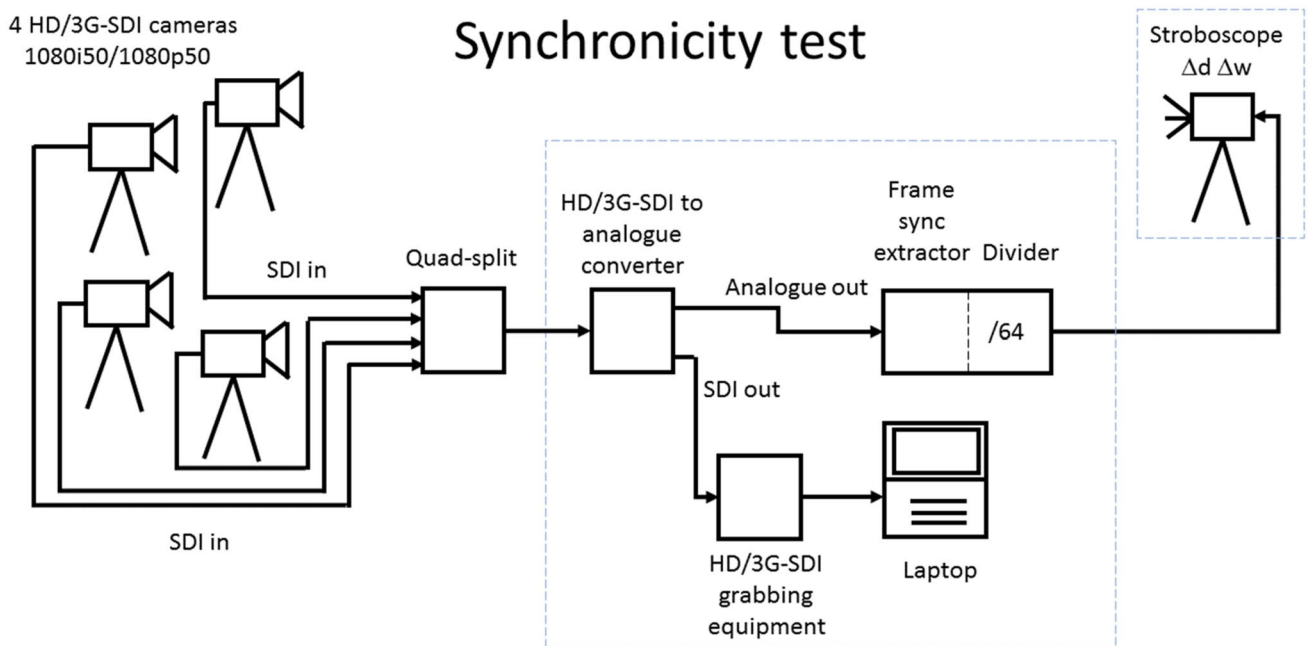


Fig. 3 The measurement setup for latency measurements. The dashed blue frames are the parts of the measurement setup. Grabbing equipment: Black Magic Design Ultra Studio HD Mini. Laptop: HP Zbook 15 G6

our case, we tap a quad-split signal into a computer and analyse the synchronicity of four channels in parallel, see Fig. 3. The software, in our case an off-line analysis written

in MATLAB, automatically detects the video format, i.e., progressive, interlaced, or progressive segmented frame, and then performs the synchronicity analysis. The only

input it needs is how many frames/fields there are between the flashes, which in our case are 64 field/frames for most measurements, see Fig. 3. If a plurality of cameras needs to be measured, several synchronicity measurements must be performed. If the number of channels that had to be analysed is m and the number of channels that can be analysed in one measurement is n (i.e., 4 for a quad-split and 16 for a 16-split), the total number of measurements required is then, see Eq. (1):

$$N = \text{ceil} \left(\frac{m-1}{n-1} \right), \quad (1)$$

where $\text{ceil}()$ is the ceiling function, m is the total number of channels/cameras, and n is the number of channels/cameras that can be analysed at the time.

3.2 Video quality

The video quality testing is based on ingesting a known uncompressed video into the VAR system via SDI and then replaying this video from the VAR system and recording it when it is sent back via SDI or High-Definition Multimedia Interface (HDMI), see Fig. 5. The videos of known quality are available in both 1080i50 (1920 × 1080 interlaced with 50 fields/s) and 1080p50 (1920 × 1080 progressive with 50 frames/s) and should both be tested. It will compare the video quality degradation inflicted by the VAR system on video that has been processed by the VAR system to be stored and played back but has not changed in size, format, scaling, or resolution.

The following equipment has been used for assessment video quality:

- A laptop with similar or higher specifications to the ASUS GX501 (with Intel i7-7700HQ@2.8 GHz, 16 GB RAM), with a Thunderbolt 3 connection.
- Software that is compatible with HD/3G SDI generator and sampling equipment for ingesting (video player) and recording (video recorder) uncompressed video. The video player and recorder used here was AJA Io 4 K Plus.

The following procedure is used:

- Ingesting
 - Connect HD/3G SDI generator's SDI output to SDI input from VAR system.
 - Send test video to the VAR system using the video player.
- Playback and recording

- Connect HD/3G SDI sampling equipment's SDI/HDMI input to the SDI/HDMI output from the VAR system.
- Play test video in VAR system.

- Record the output with the video recorder.
- The output video is then assessed using a video quality model.

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

The suggested requirements are to obtain:

- Average Mean Opinion Score (MOS_i)¹ ≥ 4 for each set of test videos.
- $\text{Min}(\text{MOS}_i) \geq 3$ for each set of test videos.
- The above two requirements should be fulfilled on two or more of the three sets.
- The final measurement value will be given from the highest scoring set.

The MOS requirements will be verified by the following VMAF scores on the individual test videos [3, 7]. Interlaced grabbed video will be deinterlaced before calculating the VMAF scores, using FFMPEG's yadif (Yet Another Deinterlacing Filter) mcdeint (motion compensating deinterlacing) "slow processing option" [3, 7]

- 1080p: mean (VMAF_i) ≥ 92 and $\text{Min}(\text{VMAF}_i) \geq 85$.
- 1080i: mean (VMAF_i) ≥ 85 and $\text{Min}(\text{VMAF}_i) \geq 75$.
- The above two requirements should be fulfilled on two or more of the three sets.
- The final measurement value will be given from the highest scoring set.

where ($i = 1, 2, \dots, 7$) in all cases.

4 Results

4.1 Latency measurements

In the field measurement, we used SDI cameras (1080p50) and a quad-split generator in the loop. The latency was the measure for one of the cameras. The latency result can be seen in Fig. 4A, with a mean value of 202 ms and a standard

¹ Actually DMOS i.e., difference MOS, which is the difference in MOS between the reference and the degraded video.

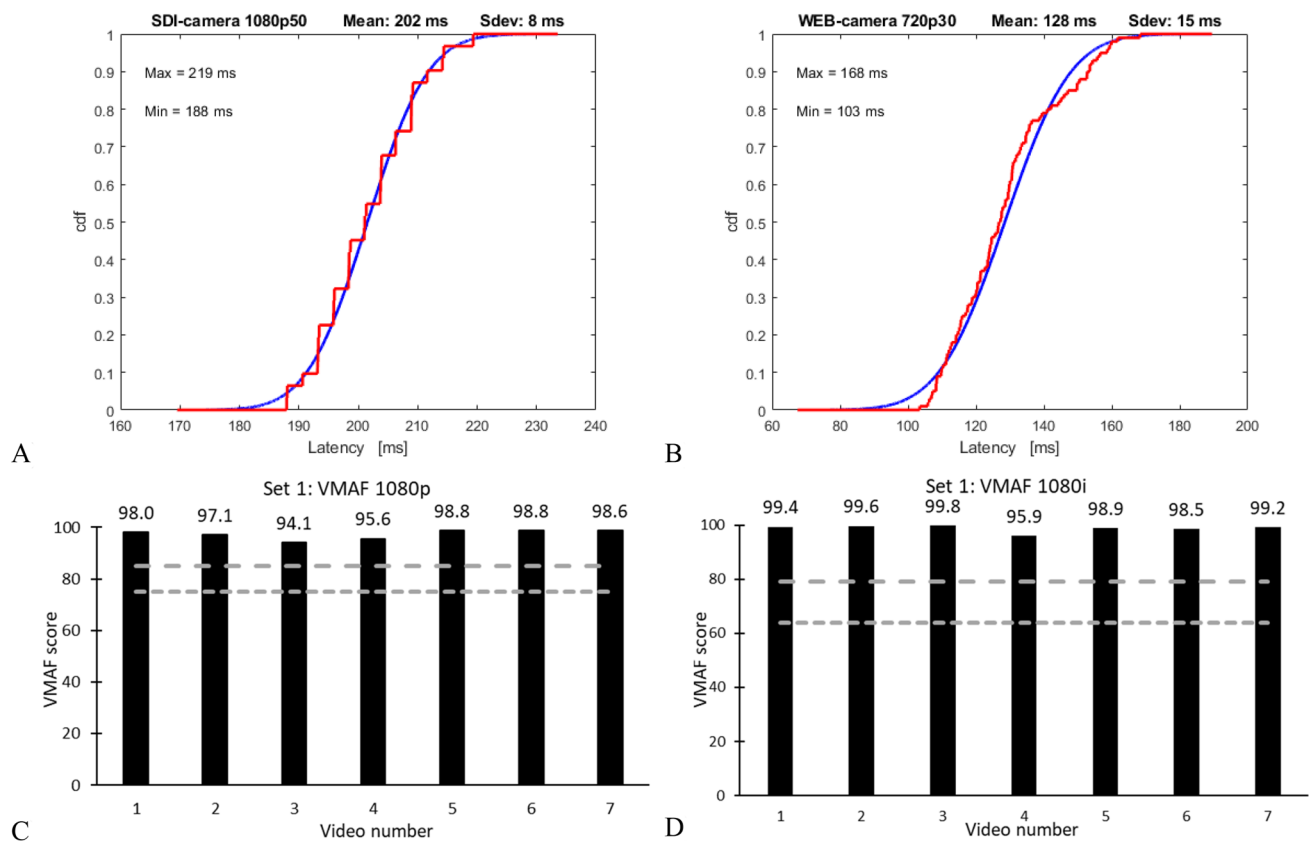


Fig. 4 **A** Latency measurement (synchronous) for an SDI camera in a full field setup. **B** Latency measurement (asynchronous) for a web camera in the lab directly fed to a monitor. **C** The vertical bars

indicate VMAF scores for 1080p and **D** 1080i for each of the seven ingested video clips. The horizontal long dash line and short, dashed line indicates the performance levels system is evaluated against

deviation of 8 ms. In the lab test, we used a web camera (720p30), and the stroboscope was running on an external pulse generator running freely on its frequency. In this case, the video signal goes directly from the web camera to the monitor which means much fewer processing steps and thereby an expected lower latency, which was also measured. Since the measurement is asynchronous, the standard deviation is higher. For the web camera, the mean value was 128 ms and the standard deviation 15 ms, see Fig. 4B. The measurement results are in line with what can be expected from these two types of camera-to-monitor systems.

4.2 Synchronicity measurement

Synchronicity tests are always performed in synchronous modes since the aim is that all cameras in the quad-split shall be (ideally) synchronous down to the last frame. The software needs to know how many frames there are in the flash cycle, but that is all. It automatically detects video format, and an interlaced signal gets deinterlaced before analysis. For time processing reasons, the colour signal is also converted to a black-and-white signal. The test in Fig. 5 is a field

test according to the setup depicted in Fig. 3 using four SDI-cameras (1080i50). The quad-split signal is first stored in a computer, and then the analysis begins. The software identifies in which frame the flashes occur for the four channels. If the cameras are out of sync the flashes appear in different frames, see Fig. 5. The pictures in Fig. 5 are mosaics of the quad-split signal putting together four pieces of the quad-split video with their specific frame number. If there are 64 frames in the flash cycle the counter goes from 1 to 64 and then starts over again. Please note, the frame number does not represent a latency value, it is just a counter for the flash cycle. In Fig. 5, one camera (the lower left) is one frame too early, and that flash occurs in frame 39.

4.3 Video quality

The measurement method was tested in the field at an unofficial VAR test event in the Netherlands [19]. Videos were ingested into one of the tested systems of the event, according to the method described above. The videos were stored on the system and then played back and grabbed. VMAF scores were then calculated for comparison with the required

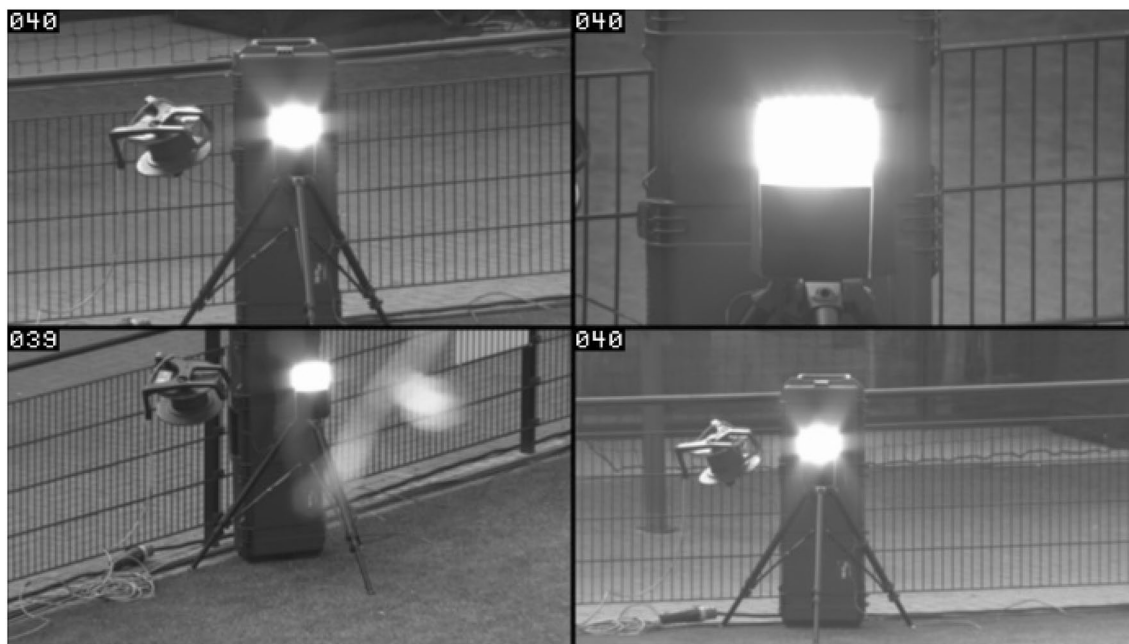


Fig. 5 Synchronicity measurement with four SDI-cameras (1080i50) in a full field setup, see Fig. 3 Here, one camera is out of sync

scores, see Fig. 4C, D. The vertical bars indicate the performance of the tested system for a particular test video. The horizontal long dash line indicates the level of mean performance that the tested system is evaluated against (VMAF = 97.3 > 92 for 1080p and VMAF = 98.8 > 85 for 1080i) and the short-dashed line the level of minimum performance (VMAF = 94.1 > 85 for 1080p and VMAF = 95.9 > 75 1080i).

5 Discussion

The methods for measuring latency and synchronicity presented in this article are based on time marks (stroboscope flashes, see Figs. 2 and 3) that are synchronous with the video streams. Published studies tends to be based on asynchronous time marks. The synchronicity measurements presented here can also be adapted for multicamera 3D systems.

Synchronicity is often measured towards a real event, whereas latency is measured as the time it takes for a video frame to travel from one place to another. Latency can thus also be measured with a video containing timecodes. A time-coded video makes the measurement easier to perform but the resolution will be reduced to one frame. This is most likely how the methods will be developed as it requires substantially fewer equipment on-site and are easier to operate.

The video quality measurement method VMAF [9] is considered currently to be the state-of-art method. The development done in this work is to make a practical method for testing systems in operation and not just on video files.

To avoid tuning the system to particular video clips, these needs to be updated on a regular basis. With the introduction of new video formats, the method will have to be updated, either with another algorithm or retrained. Monitoring live with methods not requiring a reference, as in contrast to VMAF, the quality degradations could be discovered immediately as it happens. This should be the target method in the long run. However, this is expected to be a few years in the future before it is mature enough [20].

6 Conclusions

In this work, we have presented a working method to evaluate the quality of VAR systems based on methods performance on latency, synchronicity between video feeds, and video quality. The methods have been tested in the field by using commercial broadcast equipment and on VAR system candidates. The methods have shown to be robust and have the expected accuracy. The measurement methods are today designed for capturing raw data on-site and for processing and evaluation off-site. The latency measurement will catch true latency including the latency induced by the broadcast cameras, but the video quality has the camera quality as baseline and cannot capture poor camera quality. The development of these methodologies offers tournament and competition organisers across a variety of different sports the possibility to explore both high- and low-end solutions for video refereeing. This work can be leveraged to further

improve how video refereeing technologies are best implemented and optimised to ensure a more transparent decision-making process and a fairer game.

Acknowledgements We acknowledge the invaluable help and assistance from Börje Andrén, Pär Johansson, Andreas Langell, Benny Norling, and Anders Gavler for the successful completion of the project. This research has been partly supported by the Sweden's Innovation Agency (VINNOVA, dnr. 2021-02107) through the Celtic-Next project IMMINENCE (C2020/2-2) and RISE internal funding.

Author contributions KB: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, software, validation, visualisation and writing—original draft, writing—review and editing. AD: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, software, validation, visualisation and writing—original draft, writing—review and editing. OO: data curation, validation, investigation, visualisation and writing—review and editing. JB: data curation, validation, investigation, visualisation and writing—review and editing. KW: data curation, validation, investigation, visualisation and writing—review and editing. NE: conceptualization, funding acquisition, project administration, writing—review and editing.

Funding Open access funding provided by RISE Research Institutes of Sweden. This project was funded as a development project together with FIFA during 2018–2020. RISE has received grants from FIFA during the development project. The research was also partly supported by internal funding of RISE.

Data and materials The data measured, analysed and presented in this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest This project was funded as a development project together with FIFA during 2018–2020. RISE has received grants from FIFA during the development project. The research was also partly supported by internal funding of RISE. RISE has business contracts with FIFA and is the organisation to perform certification measurement for Quality of VAR on behalf of FIFA. One co-author serves as a Guest Editor for the Topical Collection for Football Research in Sports Engineering, and another serves on the Editorial Board of Sports Engineering. Neither of them was involved in the blind peer review process of this paper.

Ethical approval Not applicable.

Consent to participants Not applicable.

Consent for publications All authors have given their consent for publication.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will

need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. FIFA (2017) FIFA at the forefront of broadcast technology. Available from: <https://www.fifa.com/about-fifa/commercial/fifa-tv/news/fifa-at-the-forefront-of-broadcast-technology-2916003>, Accessed 7 Mar 2022
2. IFAB (2021) Video assistant referee (VAR) Protocol. Available from: <https://www.theifab.com/laws/latest/video-assistant-referee-var-protocol/>. Accessed 7 Mar 2022
3. FIFA (2020) Handbook of test methods for video assistant referee systems <https://digitalhub.fifa.com/m/2ad4cd98fdc246aa/original/VAR-Test-Manual-June-2021-V-1-2.pdf>. Accessed 7 Mar 2023
4. Bachhuber C and Steinbach E (2016) A system for high precision glass-to-glass delay measurements in video communication. In 2016 IEEE international conference on image processing (ICIP)
5. Brunnström K, Dima E, Qureshi T, Johanson M, Andersson M, Sjöström M (2020) Latency impact on quality of experience in a virtual reality simulator for remote control of machines. *Signal Process* 89:116005. <https://doi.org/10.1016/j.image.2020.116005>
6. Liu T-J, Lin Y-C, Lin W, Kuo C-CJ (2013) Visual quality assessment: recent developments, coding applications and future trends. *APSIPA Trans Signal Inform Process* 2:e4. <https://doi.org/10.1017/ATSIP.2013.5>
7. Brunnström K, Djupsjöbacka A, Andrén B (2021). Video quality based on a user study with video professionals for Video Assisted Refereeing (VAR) Systems (RISE report 2021:29). RISE Research Institutes of Sweden, Stockholm, Sweden, doi: <https://doi.org/10.23699/79ja-gj68>.
8. Wolf S, Pinson M (2011) Video quality model for variable frame delay (VQM_VFD) (NTIA Technical Memorandum TM-11-482). National Telecommunications and Information Administration (NTIA), Boulder, CO, USA
9. Li Z, Aaron A, Katsavounidis I, Moorthy AK, Manohara M (2016). Toward a practical perceptual video quality metric. *netflix technology blog*. Available from: <https://medium.com/netflix-techblog/toward-a-practical-perceptual-video-quality-metric-653f208b9652>. Accessed 23 Oct 2018
10. Sielhorst T, Sa W, Khamene A, Sauer F, and Navab N (2007). Measurement of absolute latency for video see through augmented reality. In 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality. p 215–220, doi: <https://doi.org/10.1109/ISMAR.2007.4538850>.
11. Boyaci O, Forte A, Baset SA, Schulzrinne H (2009). vDelay: a tool to measure capture-to-display latency and frame rate. In 2009 11th IEEE International Symposium on Multimedia.
12. Boyaci O, Forte A, Baset SA, Schulzrinne H (2009). Demonstration of vDelay: a tool to measure capture-to-display latency and frame rate. In Proceedings of the 2009 11th IEEE International Symposium on Multimedia. IEEE Computer Society. p 444–445, doi: <https://doi.org/10.1109/ism.2009.129>.
13. Hill R, Madden C, Hengel AVD, Detmold H, Dick A (2009). Measuring latency for video surveillance systems. In 2009 IEEE Digital Image Computing: Techniques and Applications. Melbourne, VIC, Australia, p 89–95, doi: <https://doi.org/10.1109/DICTA.2009.23>.
14. CAST (2013). White Paper: understanding—and reducing—latency in video compression systems. White Paper Tech Note. Available from Accessed 29 Nov 2018
15. Eberlein P (2015) Understanding video latency—what is video latency and why do we care about it? Sensoray Company, Inc.

16. Science Mosaic (2013). Video and audio latency measuring system (DLY002 Version: 4). Teadusmosaiik OÜ, http://www.sciencemosaic.com/downloads/ScienceMosaic_latency_system_manual_v4.pdf.
17. Jacobs MC, Livingston MA, State A (1997). Managing latency in complex augmented reality systems. In Proceedings of the 1997 symposium on Interactive 3D graphics. Providence, Rhode Island, USA: ACM. p 49-ff, doi: <https://doi.org/10.1145/253284.253306>
18. Wei C, Chen H, Song M, Sun M, Lau K (2013). A capture-to-display delay measurement system for visual communication applications. In 2013 Asia-Pacific signal and information processing association annual summit and conference.
19. FIFA (2020). Latest FIFA VAR technology tests conducted successfully on KNVB campus. Available from: <https://www.fifa.com/news/latest-fifa-var-technology-tests-conducted-successfully-at-the-knvb-campus-in-th>. Accessed 8 Mar 2023
20. Pinson MH (2022) Why no reference metrics for image and video quality lack accuracy and reproducibility. IEEE Trans Broadcasting. <https://doi.org/10.1109/TBC.2022.3191059>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.