Applying Automation to High-Quality/Low-Cost Orchestra Live Streaming

By David R. Chalmers

Abstract

Many publicly funded orchestras are looking to engage new audiences via live streaming but this approach has always been expensive and complicated, and limited budgets prevent it from being a regular occurrence. The use of robotic pan-tilt-zoom (PTZ) cameras can reduce the cost. These cameras, however, require a large number of different shots to avoid becoming too static and boring for the viewer. This then requires more PTZ operators, which then starts to negate the original savings. This article outlines some experiments that we have conducted with the British Broadcasting Corp. (BBC) Scottish Symphony Orchestra (SSO) to repurpose and adapt an existing multicamera automation tool designed for capturing pop music performances to a different genre and style of music. We will consider if this approach can help bring costs down while maintaining the visual interest and complexity of a high-shot-count orchestral capture.

Keywords

Automation, orchestra, pan-tilt-zoom (PTZ), software, streaming

Introduction

British Broadcasting Corp. (BBC) Scotland is seeking to improve audience outreach and increase the number of people able to watch live concerts performed by its in-house orchestra, the Scottish Symphony Orchestra (SSO), at its permanent home at City Halls in Glasgow. It would like to increase the number of performances that are available live to a wider audience through live streaming to the BBC website and subsequent video-on-demand (VOD) viewing via the BBC iPlayer.

Unfortunately, the process of covering a live symphony orchestra performance via live streaming using traditional outside broadcasting techniques is prohibitively expensive.

To address this, many orchestras and concert halls have already explored the use of robotic pan-tilt-zoom (PTZ) cameras. PTZ cameras cost less, allow one operator to control multiple cameras, and are small and unobtrusive, and so are less distracting for both performers and audiences. They can also to be stored easily and recall preset shots.

Unfortunately, the process of covering a live symphony orchestra performance via live streaming using traditional outside broadcasting techniques is prohibitively expensive. To address this, many orchestras and concert halls have already explored the use of robotic pan-tilt-zoom (PTZ) cameras.

There are significant limitations, however, as the cameras are generally fixed to a single vantage point on or near the stage and cannot easily be moved around like a jib or a tracked camera without an additional kit.

Although there have been significant improvements in the motor control that allow some "in vision" movement, it is still a lot harder to do dynamic movement or to respond quickly to unplanned requirements.

However, by having many more framing and focusing snapshots saved as presets (up to 40 per camera is common)—it is possible to avoid the shots becoming too repetitive. However, as you increase the number of presets saved per camera, you also increase the workload for the remote camera operators, who are recalling them manually in time for the next cue.

This can become very stressful for the operators, especially on long, complex classical pieces with hundreds of shot changes during the orchestral pieces.

An Opportunity for Automation

We considered that a system for automated PTZ preset recall could potentially help in this context.

In some ways, classical music played by an orchestra really lends itself to an automated approach because most of the performers remain relatively static (with some exceptions!); improvisation is fairly unusual, and a description of the composer's original intent already exists—the score.

Digital Object Identifier 10.5594/JMI.2020.2969762
Date of publication: 15 April 2020

There remain, however, some significant challenges, specifically:

- Someone still needs to decide what visuals best represent the music at any given point.
- Matching the timing of any automation with the real-life timing of the orchestra performance can vary considerably, even from one run through to another! There is no easy solution for "beat-matching."
- Orchestral pieces are often quite long—up to an hour and may only be played in full *once* during rehearsals which does not give much chance to learn or adapt things.

So, what is required is a PTZ automation system that enables the capture from an orchestral score of the key information, allows the straightforward preparation of camera preset cues in advance, permits easy adjustment of timing to match the actual orchestra performance, and gives sufficient flexibility to allow for unforeseen changes.

To the best of our knowledge, such a system does not actually exist, so we are faced with the necessity of either building one ourselves or finding something similar and adapting to it.

CuePilot—Automated Camera Cue Cards

We had a serendipitous encounter with a company at the IBC trade show in Amsterdam in 2018. CuePilot[1] is a company that already produces a software-based camera cue card system that is well established in the market for high-end TV production of live music events such as the Eurovision Song Contest. It relies on a click track or playback system generating time code to keep the system in sync and can show the camera operators what shot they should be lining up next and how long it is needed.

An iPad or tablet app is used in lieu of the paper camera cards normally used, and it is synchronized to a server running in the gallery. It can also directly control the vision switcher, so that the right camera shot is cut up at exactly the right moment without requiring an operator to intervene.

Our first thought on seeing a demonstration of this system was that it was extremely close to what we needed and could just require the addition of PTZ recall to make it usable in this application.

Fortunately, the developers understood our requirements and offered to implement the PTZ recall function. They added support for the Sony VISCA over the IP (a common protocol used by a number of PTZ camera manufacturers) and subsequently have also added support for Panasonic PTZ cameras.

The software now allowed the user to specify whether a camera was manual or PTZ and to set up a list of preset shots for each PTZ camera, which would send the correct VISCA commands to recall that preset when triggered.

Since the software could "look ahead" and see what the next preset for each camera was going to be, and since it was controlling the vision switcher (so it knew which camera was currently "on air"), it could recall the next preset as soon as the previous one was no longer needed.

Testing

Once the developers had implemented the PTZ recall feature, we arranged to run an initial test with a full orchestra. Following the success of the initial test, we used the system in production for six further live streams of SSO concerts. These were live streamed to the BBC website as well as other platforms and made available afterward on the BBC iPlayer VOD service.

The setup for the initial test is shown in **Fig. 1**. We had four broadcast-quality PTZ cameras remotely controlled by a single control panel and two locked off cameras—a wide shot of the stage and a reverse shot of the conductor. Then we had a single electronic field production (EFP)-style shoulder-mount camcorder on a tripod at the rear of the hall with a long zoom lens (48×) controlled by a dedicated camera operator.

Seven camera sources in total fed into a vision switcher, which was controlled by the CuePilot server. We also used iPads running the CuePilot CueApp to show the upcoming shots to the camera operator and to the remote PTZ operator.

For the subsequent concerts, we increased the complexity of the system by dropping one static camera and adding two additional PTZ cameras, taking the total camera count to eight, with seven under CuePilot control. This allowed us to position some of the cameras among the performers onstage for more close-up shots, which add to the visual experience.

We also used a larger Blackmagic ATEM vision switcher with two mix/effect busses. This allows the CuePilot system to drive one M/E buss and present the output of that buss to the other buss, which serves as the main program output. This allows an "emergency manual cut" option if, for some reason, the CuePilot software failed or got out of sync during a performance.

The following sections outline some of our findings.

Preparation

Prior to the performance, it is necessary to prepare the run-order in the CuePilot tool. The director had already marked up the score with a choice of shot for the entire piece, specifying a shot number, a camera number, and a preset number in each case, see **Fig. 2**.

This produced a list of preset shots per camera, and these were then entered into the CuePilot system to produce a timeline of cues, as shown in **Fig. 3**.

The main challenge came when we started trying to map the shot timings against a commercial recording of the same piece imported into the tool.

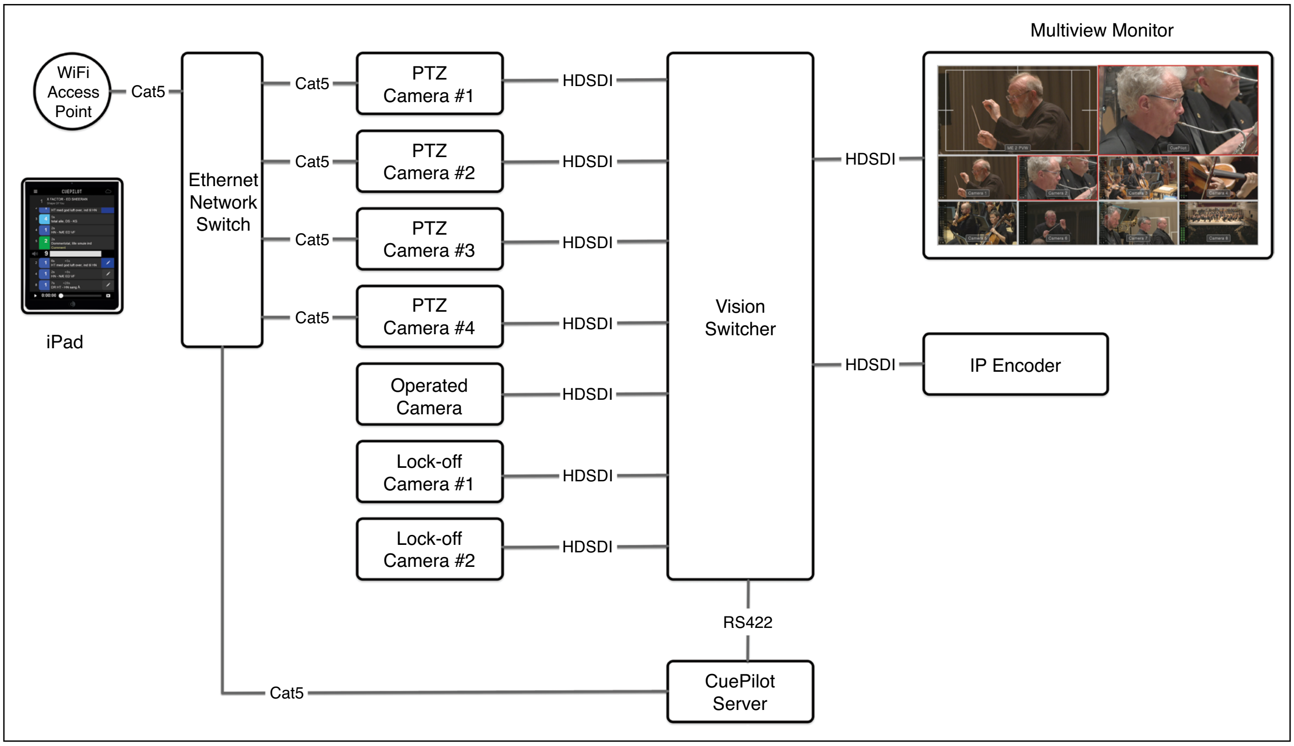

FIGURE 1. CuePilot test system setup.

It became clear that the method normally used in CuePilot of setting a master tempo for the entire piece and then using beats/bars markers to navigate through the piece was not really usable. The tempo variations within the piece required multiple time signature and tempo changes, which, while possible within the tool, was a very time-consuming process.

At this point, we decided to abandon any attempt to match the exact timings of the recording and instead to try matching the bar numbering on the score.

For the later production, we went even further and did not even try to set the correct timing—all cues were added with a default 4-sec gap. This worked well for the directors, who were getting their timing info directly from the score, but less well for the camera operators who had lost any sense of how long it was until the next cue.

Following this feedback, Cuepilot has now implemented support for dropping in cues "on the fly" from a pre-prepared list while playing back a recording.

FIGURE 2. Example of marked up score.

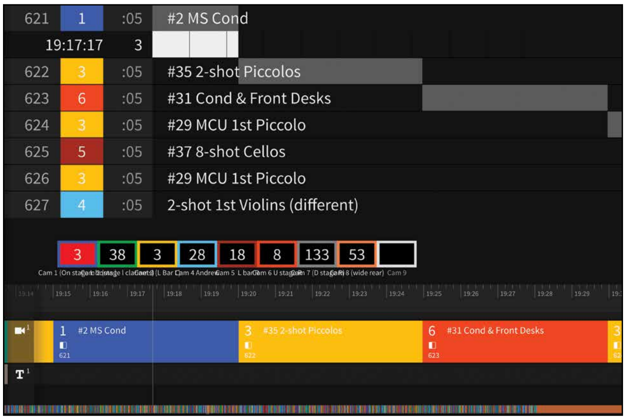

FIGURE 3. Corresponding CuePilot timeline.

This has significantly reduced the amount of time required to add cues during the preparation phase.

Synchronization

When we started trying to replay the cues based on the prepared timing, we immediately realized that it was not going to work. The operator was spending all their time worrying about keeping in sync and no time at all thinking about the actual shots.

We halted the testing, and after some discussion, the developers retired to their hotel and quickly implemented a new mode, which they called *step mode*.

In this mode, the timeline still played, but if it reached a cue point and the operator had not pressed the "next cue" button, it would pause.

Similiarly, if the operator pressed the next cue button early, the timeline would jump ahead to the next cue. It is a very intuitive approach as you are still taking the shots when you want them, but all the other parts of the directing and vision switching process are handled for you.

During testing, this proved to be very effective. The operator was able to keep in sync—as long as they knew where they were in the score.

Manual Cut-Always

Another feature added during testing was the ability to override the currently cued camera shot manually by pressing a number key matching the camera number. So, the operator could decide not to take the next cue but to tell CuePilot to cut up one of the currently available camera shots instead. This was useful if a particular shot had become unusable for some reason, or if a camera failed or lost connection.

Later software updates now make cutting shots manually very simple, while repositioning in the timeline, ready to rejoin the automation at the right point. This is very useful for ad-lib "as directed" sections or unrehearsed encores, which happen sometimes with this type of performance.

Crewing

We settled on a crewing arrangement where a score reader just followed the score and called out shot numbers, while the director operated the CuePilot software, stepping through the shots when the right time had been reached. A third vision engineer would be tweaking the recalled preset shots and making fine adjustments on focus, exposure, and framing.

In the later concerts, we found that while the director concentrated on the cue timings, it was difficult for them to go through the different camera options and make decisions about skipping unusable shots. The director was still inclined to look at the score, even though there was also a score reader shouting out cue numbers.

An alternative model would be to have the score reader responsible for cue timing, triggering the cues directly via a second USB keypad, while the director focuses more on the shot selection and overall "look." This is something we would like to experiment with further.

Further Potential Use-Cases

It could be argued that what we demonstrated was just a more complex way of achieving something that could have, and is, already done manually now. While that is possibly true in this specific test case given that we had hired in an experienced multicamera director and we "only" had 4–6 PTZ cameras—we believe there are significant opportunities to scale this approach in a number of interesting directions.

We identified the following potential use cases:

- Increasing the visual complexity of a performance capture without any corresponding increase in cost. So, adding more cameras to give more shots and angles without having to add more operators and increase cost. Our current system is constrained to eight cameras due to software limitation of the vision switcher.
- Reducing the technical skill level is required to achieve a productive output with a small team. An example of this would be radio visualization of a live music performance, something that is beyond the scope of a typical PTZ-type setup in radio studios.
- Splitting out the previsualization and preparation tasks from the operation tasks. This would maximize the value from the more specialist staff to choose shots and angles, without requiring them to be onsite during the actual recording. Someone else could "play back" their timeline without needing the same specialist knowledge. Thus, combined with the cloud-syncing capability of the CuePilot an interesting decoupled workflow could be possible.
- Exploring nontraditional ways of presenting the performance to audiences that would not otherwise be possible. One example of this would be to create multiple "mixer out" feeds at the same time that are all under CuePilot control that together make a quad split "montage" of angles all of which the audience see simultaneously. In this model, each "tile" would

need to rely on at least two dedicated cameras—one live while the other is moving.

Conclusion

Having adapted the CuePilot tool to this rather different use case and having been able to use this experimental workflow for a whole season of live streaming performances, we now believe that there is significant potential in using automation for orchestra live streaming. This approach allows you to add more cameras/angles/cuts to give a more engaging viewing experience without increasing the cost. The system can also be retro-fitted to existing PTZ camera installs.

We would be very much interested in finding a way to capture timing directly from the score without needing to enter it manually, and this is certainly an area for future development.

We believe this approach helps to achieve the goal of being able to live stream more performances in a more visually interesting way in the same budget.

Acknowledgments

The author would like to thank Johnny, Alan, Jenn, and other colleagues at BBC Scotland for their contributions to this work. He would also like to thank Kriss Russman and Morag MacIntosh, along with Per and Jakob from CuePilot.

Reference

1. "Cuepilot." [Online]. Available: https://www.cuepilot.com

About the Author

David R. Chalmers is the technology development manager for BBC Scotland, Glasgow, U.K. Along with his small team, he is responsible for evaluating and developing new technology solutions for TV and audio production. His background is in electronics and camera sensor design. Recently, he has specialized in innovation around IP contribution for outside broadcasts and in low-cost production.