Old School: An 8K Multicamera Shoot to Create a Dataset for Computational Cinematography

GRAEME PHILLIPSON, STEPHEN JOLLY, ALIA SHEIKH, PETER MILLS, REBECCA STAGG, and MICHAEL EVANS, BBC Research & Development, United Kingdom

In our paper, we describe the rationale for, and the process and outcomes of, a shoot that BBC R&D carried out in 2021 to create the Old School dataset, to further research into AI automated editing systems. We introduce our way of thinking about these systems, and explain why we believe a widely available dataset is necessary if research into these systems is to reach its potential. We then explain why a new shoot was necessary to generate such material.

We proceed to give an overview of the creative process, explaining aspects of the shoot that differed from conventional TV production in order to produce a dataset with the desired technical and editorial characteristics. We describe the planning and pre-production, the shoot itself, and post-shoot work to process the material and generate metadata.

We finish with a summary of the material and metadata that have been included in the asset bundle for the dataset, and an overview of how we decided what to include in this bundle.

This paper was published as a BBC R&D White Paper in July 2022 [3].

Additional Key Words and Phrases: Artificial Intelligence, Machine Learning, ML, 8K, multi-camera, rushes

ACM Reference Format:
Graeme Phillipson, Stephen Jolly, Alia Sheikh, Peter Mills, Rebecca Stagg, and Michael Evans. 2023. Old School: An 8K Multicamera Shoot to Create a Dataset for Computational Cinematography. 1, 1 (May 2023), 2 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 OVERVIEW

Our paper describes a project to create a publicly available BBC R&D production dataset to support academic communities that would benefit from access to the raw material from which television is made. set during the filming of a fictional game show called “Old School”, it includes comedy fiction (“sitcom”), drama and gameshow elements.

BBC R&D work on Intelligent Cinematography has focused on “multi camera” television production, in which multiple cameras record the action simultaneously. We have avoided approaches wherein an AI director drives robotic cameras, as these lock in framing decisions at the point of capture. Instead, we use high-resolution locked-off cameras to frame the entire scene from different viewpoints. This allows footage to be processed in post-production by many framing algorithms. Real-world datasets for training and validating AI systems in this way are hard to obtain: conventional film and television production does not use this method, and rights and commercial issues limit the ability of media companies to share production footage with third parties.

BBC R&D tries to maximise the value of its work by supporting academic researchers in areas of mutual interest. In the last decade we have released datasets (e.g. [2]) to support several relevant AI and ML technologies. One fruitful approach has been to provide datasets in support of specific academic challenges (e.g. [1]).

After considering the limitations of the data available to support internal work on automated production, and after informal discussions with academic partners, we decided to commission the production of a television programme with several unique features:

Authors’ address: Graeme Phillipson, stephen.jolly@bbc.co.uk; Stephen Jolly, graeme.phillipson@bbc.co.uk; Alia Sheikh; Peter Mills; Rebecca Stagg; Michael Evans, michael.evans@bbc.co.uk, BBC Research & Development, Salford, United Kingdom.

2023. Manuscript submitted to ACM

- Simultaneous multi-camera recording using locked-off high-resolution (8k) cameras, each capturing the whole scene
- Varying levels of editing difficulty scripted into the performance, from segments with seated actors and little movement, to fast-paced action with actors occluding one another, speaking simultaneously, throwing objects etc
- Full rights over all aspects of the production, including the right to share the content for research purposes

While footage from one production is inadequate for training general-purpose deep learning models, we hope that it will be useful for algorithm development with less data-intensive methods, for validation, and as test material for academic challenges.

REFERENCES

[1] George Awad, Wessel Kraaij, Paul Over, and Shin'ichi Satoh. 2017. Instance Search Retrospective with Focus on TRECVID. *International Journal of Multimedia Information Retrieval* (2017-04-05 2017). https://doi.org/10.1007/s13735-017-0121-3

[2] J. S. Chung and A. Zisserman. 2016. Lip Reading in the Wild. In *Asian Conference on Computer Vision*.

[3] Graeme Phillipson, Stephen Jolly, Alia Sheikh, Peter Mills, Rebecca Stagg, and Michael Evans. 2022. *'Old School': An 8K Multicamera Shoot to Create a Dataset for Computational Cinematography.* Technical Report WHP 402. BBC R&D. https://bbc.in/3yCY4Eh