

XR Stories and SIGN Podcast

Series 2, Episode 6 - How to integrate VP with AI?

Nina [00:00:02] Hello. I'm Dr. Nina Willment and you're listening to the XR Stories and Sign podcast. In this series, we delve into all things virtual production, meeting the experts, academics and creatives working on every aspect of virtual production, or VP, you'll hear that term a lot in this podcast. In this episode, we're looking at artificial intelligence. We've seen a real explosion of AI tools over the last few years, creeping into our everyday lives. For example, you might have used ChatGPT to write an email response, a CV or even an academic essay, although I'm not encouraging that one. And we're seeing this flow through to VP with tools like DALL-E and Midjourney, two AI software platforms that enable creatives to generate visual images from text prompts, but how far have we got to go before we successfully integrate AI with VP? And what does this all mean for creators, film crews and the audiences who experience what is made? To help me answer that question, I'm joined by Florian Block, Professor of Digital Creativity at the University of York, R&D Lead for AI & Immersive at Dock 10, and Co-Founder of innovation consultancy, Fast Forward. I started by asking Florian to tell me all about his VP facing work in both academic research and industry R&D.

Florian [00:01:24] So in my industry, in my industry role, I work very closely with the team at Dock 10. So we are the UK's largest television broadcast facility, over in Media City in Manchester, Salford. There I'm very much focused on multi-camera virtual production, so it's a distinction I'll draw on quite a few times from in this podcast, when we shoot a film we call this Single Camera because all shots are planned and recorded and then after we record all the different shots, we then piece them together. For anything that's live or recorded as if it was live, that doesn't work, so you, essentially what you have is multiple cameras in the studio from different perspectives, some maybe close ups, some maybe wide shots, and then the director in the television gallery mixes the different perspectives sort of in real time. And that's an important distinction because it has quite important consequences for virtual production, for the technologies used, but also for what you can and cannot do in post-production. So really, for what I do at Dock 10, I very much look at the liveness of virtual production. So anything from sports shows, to maybe live game shows, to political reporting, anything that has to sort of- that happens and needs to go out to broadcast right away. So that's my virtual production context there. And in my academic work, I pretty much look at creative AI or AI in creative practice. So I look at data driven storytelling, I've looked at that in the past where, for instance, for live sports we analyse performance of what's going on and you'll see in tennis you often have augmented reality overlays about where the shots landed on the court. So it's that kind of data driven narrative which AI is used essentially as a real time analyst, that's sort of one example. And I'm just really, really interested in the convergence between immersive technology and live events, performing arts and film and video games, we see a convergence between all those genres that were traditionally separate, that are now becoming almost the same, you know. Movies become and TV becomes a little bit more like games, music is now streamed into game environments, so I'm also really, really interested at that sort of level of media convergence.

Nina [00:03:59] That's amazing, thanks Florian. Just to bring it back maybe, how would you define AI? People might have seen it in action but maybe they don't know exactly how to define it or what exactly it is. How would you define what artificial intelligence is?

Florian [00:04:13] Yeah, I think that's a very, very frequently asked yet very, very hard, hard question. The key distinction between what AI is to what we had before which is sort of smart algorithm, is that while you architect as a human a general framework for what the computer needs to do, the actual details of how to achieve that are learned by the artificial intelligence by means of training on existing data that essentially shows both the inputs, the material you want to- the AI should process and the output, which is the material that the AI should produce. And so by that means you don't necessarily need to tell the computer beforehand exactly how to do it, you just need to show it what to do. And then it learns that, which is very exciting but it also comes with lots of issues. You lose some degree of control. We can problematise also the issues of representation and sort of what it produces based on what it learns, and if what it learns is problematic it'll reproduce those problems. That's a different story, but it ought to be brought up every time we talk about this. But in my view, in the creative sense, AI is something that- it's an algorithm that is not predetermined that figures things out based on observing other material.

Nina [00:05:44] So just picking up on some of those challenges I suppose around AI, we're hearing a lot about the threat from AI in the mainstream media, it's going to steal jobs, it's going to take over our lives, we can't escape- all of that dystopian future sort of thing. What are your thoughts on that rhetoric?

Florian [00:05:58] I think one of the biggest challenges of AI in general currently is truth and how ---, reliability of what you see outside of your physical environment, so anything that's mediated digitally that you see on social media or on TV or that is produced by people's cameras, can now very easily be forged, generated completely from scratch using these emerging technologies. So I think that is obviously something that scares a lot of people and rightfully it should worry us and we should take measures to address that. I think there is also some really viable and interesting theories about AI essentially socially engineering humans, you know, I mean, ChatGPT already is a very convincing conversational partner, so I think as the technology matures over the next years and then combine that with voice synthesis and video synthesis, now these techniques start to work in real time. So it's possible at some point you get a call, a video call, a zoom call or a WhatsApp call by someone that you think is your relative, sounds like your relative, looks like your relative, responds to your questions, but it's potentially completely artificial and potentially not with the best intentions. So I think we need to be really- what in the context of entertainment and storytelling creates this excitement because these technologies allow us to generate and explore beyond current limits, in the context of authenticity and truth poses a huge challenge. So I think that needs to be addressed and I think the first way of doing that is, as you do, you start dialogues about it and you try to reach as many people as possible to think about this.

Nina [00:07:57] Fantastic. So you mentioned obviously storytelling there and we're seeing the writers strike in Hollywood and things like that going on at the moment obviously about pay, but also about some of the fears around AI and how it will be used. What potentials and challenges do you think there are for AI in storytelling?

Florian [00:08:16] As with all the AI debates, there's always the question about is it going to replace us? Is it going to eliminate lots of jobs? And we're starting to see that certain things like detecting skin cancer on tissue samples or analysing stock markets for insurance brokers or other data analysts, these roles are already being replaced by AI. I think in the creative sector there's a valid concern by everyone who, you know, feels that their contribution is valued, that their livelihood depends on their voice as a voice actor, or on their ability to write as a writer. So I think there's absolutely necessity to recognise that

and also to really think about regulation. I think there has to be something done by policymakers. This is not something we can overnight figure out ourselves, this needs to really rapidly come from policymakers. I had a chat with a good friend of mine, Sol Rogers, global head of innovation at Magnopus, and I take his provocation on how music works in France. There's a certain percentage of music that has to be a French language, you know, and it's regulated by law. We could potentially face a similar situation, particularly in the transitional period where we don't yet fully understand the technology, that content can have a maximum of X percent of AI generated content. So I think this needs to be regulated. On the other hand, I will say there's something really interesting happening. You know, a lot of people that are very prominent are talking about, 'oh, AI will never be able to do as well as we do', 'can't write stories like we do' and that may be true but one important aspect that we shouldn't forget is that a lot of the recent AI advances have been really democratic from the start. So from ChatGPT to Stable Diffusion, Midjourney, all of these new capabilities were given into the hands of the people and even big companies like Adobe only now released AI products. So I think that's a little bit also of a threat to the sort of elitist, creative community, that all of a sudden they are saying, 'well, now everyone can create something great, where does that leave us?' but, you know, we must not forget that creativity and innovation in the space has traditionally been in the hands of a very limited demographic, and what we define as what good content is, and what creativity is, and what innovation is, was again defined by a very limited demographic, usually white males, right? And then the other thing that is really important to consider in that context, I think, is that audience standards are shifting. And I'm not saying this as a value statement, but we look at user generated content, we look at a shift from, you know, the traditional sense of production value to personal effects, personal relationship to personalities, light, short form content. You know, again, a lot of voices saying, well, yeah, maybe not for most people, but why don't they realise, you know, these stories aren't as good as our stories that we write? Yes, that may be true, but we've got to look at what's actually happening and how it empowers broader set of people to contribute to the creative endeavour and how that also really challenges us about how we think about creativity, storytelling and content production in general. I think that's really provocative number of things that we need to ask ourselves, and I'm talking about myself here because I've always considered myself part of the creative community. It's easy to forget that it's not been open to a lot of people traditionally.

Nina [00:12:11] Thanks so much, there's some really fantastic critical reflections there that's really, really encouraging to hear. You mentioned obviously a couple of things there, Stable Diffusion and Midjourney. For people that aren't familiar with those, how are they used in virtual production?

Florian [00:12:25] Okay. Basically, Stable Diffusion and Midjourney are both examples, different brands so to speak, of image generator- image generators. So these are machine learning models that can convert something that you say in text. So you say a picture of a horse on a cliff overlooking a valley in the sunset, and then these algorithms have learned to generate something on that. The current practice in virtual production for generative AI I think is mostly, by my knowledge and how my sort of involved in practice goes, I think it will very quickly disrupt concept art. I'm saying this from a commercial side instead of contracting thousands of pounds worth in a concept artist, and usually a concept- an artist is an intermediary between a client who has a vision and a technical team that does the delivery, there's really some interesting things here in making that sort of interaction more concise, right? And why not give the ability to generate some drafts and sketches into the hands of the people who are the creative decision makers? Now I'm realising saying that, that that creates also a problematic thing about are we replacing artists? Again, in the

same vein of the things that I've said, you know, it's also really important to say that using these new AI generators is not as easy as one thinks. The example I gave about the simple text prompt will not give you a good result to start with. There's more complex things there and I think as time progresses in the next few years, it'll become more obvious that it will become a dialogue of these new techniques. So think about it as you have an overarching idea, it generates something and you say ahh, actually, I don't like the way the horse faces. So then I draw a horse very crudely in a different pose, and then that will be redigested by the AI and it will take my post but then generate a photorealistic horse on top of that in the same pose, right. So I think I'm just trying to illustrate that these methods are not as simple as one thinks. But again, I think for concept art in virtual production, that will be really powerful. I also think that one of the key problems in virtual production is to mix, essentially to really teleport the physical presenter into the world. And a huge challenge there is that the lighting in the virtual scene should ideally reflect on the physical actor. So if it's a sunset scene, the actor needs to be drawn in a sort of slightly orange tone, right? The angle of lighting should throw different shadows in the face and so on. And likewise, if that actor is standing on a pier on a lake, that actor should reflect on the water. It should affect the environment around them, right? And generative AI, I've got some brilliant colleagues here at University of York who are in computer vision, they look at AI that understands light fields and scenes and composing that- using that to compose virtual productions into one believable scene, that's another. And the third one is Style Transfer, for instance, and I'm just saying examples here, there's many more but Style Transfer would be, you know, if you wanted to create like an 'aha' take on me style movie and you wanted to create the virtual you know, it's all looks like someone, a sketch artist has drawn the movie, you know you could use Style Transfer. That's another technique which essentially takes an input image, which can be a normal video, and then you train it based on a certain style that you want to achieve, and then in real time it can transfer that into that style. So that again will create quite a few interesting creative possibilities to make something look like an anime show or a Disney movie or a show from the seventies, a sitcom from the seventies, right. So again, so this is where I see some of the core uses of AI come in fairly quickly.

Nina [00:16:48] So you've covered quite a lot of maybe some of the different ways that we might think about AI with virtual production. For me, I've heard a lot of really exciting things about AI being used in motion capture, but without people having to wear motion capture suits. Motion capture, for those that maybe have not come across it, is where you basically can record the movement of objects or people in film and TV and more broadly. So yeah, it'd be great to hear where you've seen some really exciting stuff with AI in virtual production.

Florian [00:17:15] Yeah, Yeah, absolutely. I mean, look, there's, there's a ton of really exciting things happening. Because you mentioned motion capture I think, what we often consider it slightly different. Virtual production where someone is standing in a greenscreen studio and then we make it look like they're in the virtual world, but just as valid an approach as to take the movement of a human puppeteer, so we take all the movement they do, the fingers, the facial expressions, we have technology to scan all of this and suits, spandex suits that have sensors in them are tracked with cameras and, you know, there are some virtual production workflows now that don't involve filming a real person anymore, right. So, because you can then create an avatar, a virtual copy of the human that moves like the physical human, looks like the physical human, has the same facial expression but now that character's in a virtual world and can maybe even interact with the environment. Can, you know, hold their hands under a waterfall and it starts flashing and you can have all this, this lighting. So to your question on AI in this sense,

yes, I'm very excited about the markerless motion capture. So just so that people understand, motion capture has traditionally involved quite expensive- you can almost think about it as 20, 40 little cameras around the room and then you have little dots on your shoulders, on your elbows, all around your body and then people run around and all these different cameras look at these constellations of dots and then essentially in real time recreate the position of these dots because they could see it from different angles, you'd triangulate it, brilliant technology but obviously there is something about that technology that makes it quite expensive. It's not obtainable for most people and markerless tracking, AI based tracking essentially looks at sometimes just one, but usually multiple phone cameras. So again, they can be friends phone and your phone synchronised, and they look at just the two dimensional video of a scene and they can extract the different- the skeletal information where the people's arms and legs and head movement and so on. So that's very exciting because already it just eliminates 95% of the expensive infrastructure that traditional motion capture relies on. But, you know, I think again, the exciting thing is- and you have companies like MOVE AI which have a brilliant product out, there are some Open Source academic led projects like the CMU, OpenPose, the Vision Group there have created something called OpenPose that everyone can download for free. These are all examples for AI understanding human motion and then being able to in real time, transfer that into a virtual environment and then basically use that to puppeteer anything you want. It could be a humanoid character, it could be a fantasy, could be an animal. It gives a lot of creative possibilities but the role of AI here in recent years has just accelerated the computational power and the ability because again, there were some approaches before but for the first time the AI by training, by teaching it on a handful of data, it can then take that knowledge and without us telling it how to, it can recognise the different body parts and track them at a remarkable procession. So this is definitely a very, very exciting development.

Nina [00:21:01] So you mentioned there some university led projects, and you've been involved in one as R&D lead at Weaver, a £5.1 million project funded by the UK government to explore how Esports audiences engage in immersive experiences. Could you chat to me a bit about that and how you're seeing similar workflows integrated as standard VP techniques?

Florian [00:21:19] Absolutely. So this project kind of connects to, a little bit to what I said earlier about my academic background. So I've been really spending a lot of time in the last five years thinking about data driven analysis and storytelling in the context of video games, competitor video games, so also called Esports and traditional sports. And so basically, again, when you have an artificial intelligence could watch a million sports matches, say in football, and then it can watch a live match that's going on right now and it can tell you all the things that are extraordinary. It can tell you that that player for the first time has run more than 25 miles in the first half of the match, for instance. Or it can say that the serve in tennis was the fastest. Or it can even connect different stories such as, oh, that player started off really poorly compared to others like them, but in the second half they made a huge return, a huge comeback, right. So I think that's the kind of stories I got really, really interested in and in automating that process of finding these stories. And the motivation here, again, was not to replace human storytellers. You know, you have someone like Gary Lineker or the tennis commentators who create wonderful narratives, right, so again that's brilliant but it is one story for all. The intention behind the AI driven narratives was to be able to personalise the commentary to every viewer. So if you are someone new to tennis, the AI could explain to you a very different level of play and explain to you the basics. If you are a very seasoned viewer, but you may even play tennis yourself, it can give you very, very deep insights and specifically maybe even compare

what the professional tennis player does to what you do in your private court, because you have your camera set up and you're tracking your shots. So that is not possible in sports just yet but the project was in the context of video games, a competitor video games which are recorded at high fidelity, even for amateur matches. So we looked at this and these new means of --- driven storytelling as a means of delivering tailored stories to everyone and not to isolate people, but in fact to, you know, even though everyone might have their own story, the highlights bring people together and they're just the same highlights I explained in different ways so that someone who is new can say, wow, this is amazing, there's something cool happening, which they wouldn't otherwise be doing. And then a professional or more advanced viewer could say, yeah, yeah, this was incredible and you create, all of a sudden you create a connection. So I'm really interested in these AI driven narratives to create connection and sort of individualised content. And the way we've then used it in the virtual studio, you know, the idea originally of virtual production was that you had, you know, that you save that set up time, you could literally just flick a switch and you're on the new set. You can put new commentators in and produce a new show. Now that has then expanded to, oh wait, if we have a virtual environment we can also all of a sudden do some really funky stuff. We can show off a stadium in the background, even though we are in a five metre by five metre room. And so you create some interesting opportunities there to create more space and so on. What we wanted to do with Weaver really was to think more about how can we fully leverage the dynamic nature of this virtual studio so that the pundits could bring up statistics that float in midair, can bring up three dimensional maps of the pitch and points, right? Immerse themselves almost in the world of the sport and, you know, look at different angles when you have like a free kick, you can almost teleport yourself onto the pitch and see what the person taking the shot would see. It is more like this evolution of truly facilitating the capabilities of these virtual environments that we're now sort of getting into and that's the exciting bit. It's making more than just replicating the physical environment.

Nina [00:25:36] How do you think that will flow through or should flow through to grassroots studios, indies, more grassroots creatives?

Florian [00:25:44] So this is the fascinating thing, right? That I think at the top level AI will have an impact but, you know, it's almost like everything I just said makes almost more sense because, you know, a professional, like someone- like the BBC doing Olympic coverage, they have analysts, they can afford a huge crew of spotters and people looking at different stats and statisticians, but the sort of things that we built for Weaver really work for every match automatically. So you can particularly think about grassroots productions achieving similar results, even though they've got a crew size of three people maybe. You have a camera operator, director and maybe someone recording sound. And so I think again, the democratisation of AI is again one of the opportunities that excites me. Similarly, in virtual production, MOVE AI, OpenPose, is a fraction of the cost of other motion capture systems. Really a power of ten, where you have 200,000, you have maybe 2000. It's like several powers of ten reduction in costs. And likewise, I think all the virtual studio systems that are currently very expensive would become much cheaper. And things like removing the green screen, doing post-production tasks that big, big triple-A movie productions do now will soon be obtainable by many, many more people. So that is a really exciting point you're raising, Nina.

Dr. Nina [00:27:19] Fab, thank you. So you've seen obviously virtual production change traditional production pipeline, traditional production workflows in different ways, how do you think AI will then shift virtual production workflows? Other particular roles that will change, any particular skills that will change? Be great to have your insights about that.

Florian [00:27:38] Yeah, that's a really interesting and multifaceted question, I think. In the process we refer to as post-production, so after the actual action is shot, we already see a lot of AI power tools that help with the post editing. So normally you have things like colour correction or you have to understand the depth of a scene in order to replace the background or to change, tweak certain parameters. There's lots of very detailed ones that I think it's maybe not worth just going into detail on every one of them here, but there's lots of things that, in the post-production that can help. And then lastly, I will say when we go down that strand, we need to think about the way that audiences consume content, right? So traditionally, I mean, virtual production is still used pretty much predominantly in linear video. So by linear I mean it plays from start to finish. You have no way of influencing it other than maybe pause it quickly to nip to the loo and then you come back but, you know, whatever's captured happens. So some really exciting things happen when you think about creating interactive version of content where all of a sudden, you can interact with the actors, right. And so there are some really interesting things here with, again, that are controversial in nature but imagine you can pause the main script of a movie and just walk up to the actor and ask them to act out a different possibility. Say, oh, why don't you go down that path? Go into that door and not the other, and then things happen. Or let's say you watch a sitcom or like Hollyoaks and you are in a pub and things happen, the main actors, but on the side there are characters that are in the pub that you can talk to almost like, you know, they can tell you some back stories about the action that's taking place, right? So again, AI, in the context of real time generating live narrative, creates some really interesting opportunities to bring people closer to the action and make them more active. It's the same what I talked about, the data driven narratives that sort of tailor to people. Likewise, in virtual production, you can think about shooting some raw material and yes, people can watch it in a video format and that's great, but they can also put their VR headset on and dive into the world and then walk up to the actor and then all of a sudden an AI trained on the actors and on the character's background takes over and generates a character that can talk to you that you can interact with. I know it's sort of, it has controversial and some eerie aspects to it, but it's at the same time also, has the potential to really transform the way we consume content.

Nina [00:30:40] It's really exciting things to think about for the future, particularly Hollyoaks *Florian laughs*. So I suppose my final question before we move to some exciting quickfire questions at the end, there are lots of possibilities there, any areas that you think should never be integrated with AI in virtual production?

Florian [00:31:00] I think that we need to really stay clear of everything that presents itself as factual, that uses these techniques. I think there's a- again, it goes to this corrosion of truth. I think if it's clear that it's fiction or maybe even satire and it's sci fi or anything like this, I think is okay. But I think we really need to stay clear of appearing factual. And then the other big thing is consent, you know, I never want to see a deep faked or synthetic version of an actor doing anything that that actor doesn't want to do. And same goes for witness statements or for people that are not celebrities but are featured on camera. So I think that's the, that's the kind of thing that's obviously a big red flag, and there's probably others but that's certainly a big one.

Nina [00:31:57] Fantastic, thank you. You've given me and hopefully everyone listening loads of food for thought there. And it's really, really great to hear some really critical reflections on both the potentials and also some of the challenges we face. So just a round off, as a tradition on the podcast I'm going to throw some quickfire questions at you now.

Florian [00:32:15] Oh dear *laughs*.

Nina [00:32:15] So as difficult as it may be, if you could try and answer these as quickly and concisely as you can, that would be amazing. So the first one is, why should we care about virtual production?

Florian [00:32:26] New means of storytelling and and... Just new means of storytelling *laughs*.

Nina [00:32:30] Amazing, thank you. What's your favourite TV program or film that you use as virtual production?

Florian [00:32:36] Ooo, probably Mandalorian. It's a really boring answer, but it is brilliant.

Nina [00:32:42] What companies are on your radar in the virtual production space at the moment?

Florian [00:32:45] I need to say Dock 10 of course. MOVE AI, Zero Density, Pixotope and Epic, needs to be Epic.

Nina [00:32:54] If you had a magic wand, the laws of physics and the laws of finance do not apply, what virtual production problem would you solve?

Florian [00:33:03] Photorealistic volumetric capture.

Nina [00:33:07] For people that might not know what that is, if you could explain?

Florian [00:33:08] So it's essentially, it's a little bit like motion capture, but it's essentially a camera system that can capture 3D volumetric high detailed scans of the environment in real time. I would like that.

Nina [00:33:25] And finally, my favourite question. What TV show from your childhood would you want to remake in virtual production?

Florian [00:33:32] Ohhh God! I've grown up in Germany, so you won't know. But I don't know, this is so silly. I mean, as a German, I have to say David Hasselhoff's Knight Rider, It needs to be done, right? *Nina laughs* It needs to be done.

Nina [00:33:47] Incredible. So, Florian, thank you so much for joining me today. It's been such an interesting, really informative conversation. Thank you so much again for joining the XR Stories and Sign Podcast.

Florian [00:33:58] It was a pleasure. Thanks for having me.

Nina [00:34:03] Thanks for listening to the XR Stories and Sign podcast, exploring all things virtual production. Next time, we're looking at the skills gaps in virtual production. If you enjoyed this episode, you can subscribe, rate and leave a review. I'd love for you to suggest other topics you'd like to hear on the series. You can find more information about XR Stories and Sign through our websites, xrstories.co.uk and screen-network.org.uk. This episode of the XR Stories and Sign podcast was produced by the wonderful Olivia Swift and it's a Reform Radio production.