Do It Yourself - Immersive Performance

A Low-Cost XR & Live Media Hacks Toolkit

By Nathaniel Mason

Contents:

0. Introduction

Common/Wealth has been working with Creative Technologist Nathaniel Mason to explore and integrate low-cost XR technologies and immersive audio into site-specific performance. We see the power immersive technology has in shaping our futures - how stories are told and by who. Like theatre, we think this should belong to everyone.

This toolkit is designed for theatre makers, artists, and performers who want to experiment with technology to augment reality through immersive media in live performance, all without needing a huge budget or a dedicated tech team. It focuses on DIY methods, accessible tools, and practical, experimental workflows that can be adapted to various creative contexts.

Rather than offering a step-by-step manual, this is a collection of approaches, techniques, and insights that can be mixed and matched depending on the needs, but it is by no means exhaustive. It covers everything from using projection as a dynamic light source to integrating real-time AR elements and live-streaming visuals into a performance space. The emphasis is on flexibility, working with what you have, making quick prototypes, and testing ideas in the room rather than getting stuck in technical setups.

Theatre has always been an act of transformation, using simple means to create expansive worlds. This toolkit is an extension of that tradition, offering ways to layer digital media into performance while keeping the focus on liveness, presence, and storytelling.

1. Live streaming techniques for performance

Considerations to make live streaming more immersive for the remote viewer

1.1 Mobile phone as a Streaming Camera

It’s an obvious choice to use mobile devices to live stream video. It's done all the time on social media. We can harness this in performance and add to it, to have a more cinematic feeling. First thing is to get a gimbal device. We used hohem iSteady M6 kit. Then there are a variety of cheap clip-on lenses that play a massive part in making the image more cinematic. You may read reviews of how terrible they are, but they are certainly dynamic! Look at a super wide option, a macro lens that can allow for extreme closeups of eyes and other details and a long lens 75mm or so, that isn’t macro. These lenses can easily be clipped on and off and kept on the gimbal stand whilst the camera operator is moving around allowing for extreme flexibility! Granted the chromatic aberration is increased, you get swirly background focus and the image is softer than without, so those are considerations, but these are far outweighed by the flexibility and dynamics.

You need to consider how to stream from the phone - you could do this directly to YouTube et al, but if you want to use the footage in the performance or add effects you need to stream to a PC. Get yourself a decent router, any router will do, check your granny’s cupboards! But if you can use Wifi 6, that will go a long way for stability and speed.

Then decide on the image protocol to use. The very simplest is to download Iriun on phone and laptop, make sure they are on the same network and then you're done. There are loads of apps out there, but in our experience - Iriun has worked flawlessly.

The detailed version - The two protocols to think about are SRT or WebRTC. SRT is what we believe (amongst others) the BBC use and it's super reliable but has a big buffer so there is a considerable delay- 0.5 or 2 seconds or more. This latency makes it really stable but not so good for real-time communications. An alternative is WebRTC - this is the protocol used by Zoom and other similar platforms where communication is important. It's a mad protocol! it can go halfway across the world and back in a 0.2 of a second! But as you know with Zoom, it can be low resolution and not stable. The final, most innovative and most expensive is Dolby.io and their WebRTC platform. This platform enables one to stream multiple WebRTC feeds directly from the phone into their online platform and then streamed on to the viewer. The feed from there can be embedded in a video game environment (more on that later) or enable the viewer to choose which camera they would like to look through. But at £500 per month it's not all that cheap.

Haivision do a streaming app - it's quite complicated to get up and running as you will need some understanding of networking - but their team in Ireland are extremely helpful on the phone. Our favourite is Iriun as mentioned before, to go between device and PC, but only one device can be selected at a time. You can use multiple apps as a hacky DIY approach to allow for multiple mobile cams. Haivision app streaming to vmix is a great free video mixing software that can accept multiple streams from the Haivision app. And also the Haivision app has the capability to stream using SRT or WebRTC. But it's a fair bit more development heavy.

1.2 360 Camera for Live Streaming

A few select Insta360 models, like the Insta360 X and the newer Insta360 X4, can live stream directly to YouTube. We used a Pilikoll camera mount adapter to attach an Insta360 X to GoPro mounts, allowing us to secure the camera to body mounts, head mounts, and other rigging options for dynamic live streaming. One key limitation of the Insta360 X is that it must be wired to a phone for live streaming, meaning it cannot be charged at the same time, giving you only about 30 minutes of battery life. The newer Insta360 X3/4 solves this issue by supporting wireless live streaming, but it requires a 4G connection, so make sure you have enough data allowance! One of the biggest advantages of a 360 camera in live streaming is its discreet and flexible placement. Unlike traditional cameras, which require careful framing and often stay static, the Insta360 allows you to capture the entire scene without worrying about shot angles. This is particularly useful for immersive performances, where getting close to the action usually demands high-end lenses and bulky equipment. With a 360 camera, you can be inside the performance space without disrupting the experience.

1.3 Empowering Remote Audiences: Creating Interactive Experiences

There are tons of interesting ways to enable a more immersive experience for a remote audience. The challenge is how to give a remote audience agency, a question that underpins

Common/Wealth's approach to audiences in real life. For this toolkit, we want to talk about one specific way that brings together elements of live streaming and utilises Unreal Engine, a software traditionally used to make 3D video games.

The short version

The remote audience clicks a link and enters a virtual world, where they can move around freely. Within that world, the live video streams from the performance are streamed into the digital world in real time. The worlds are connected so the remote audience can see each other in the space by means of a digital avatar whether semi-lifelike or a mysterious floating blob - anything. Once inside they could interact with the world and that interaction data could be sent to the live performance and affect it in some way. Whether that's through sending messages, changing light colours, making decisions about action…. the list can go on.

The slightly longer one

Unreal Engine, Dolby.io, Pixel Streaming
Using the WebRTC protocol, Dolby.io platform and Unreal Engine it's possible to send the video feeds directly into the 3D world of the audience. The video is sent from the mobile devices directly into Dolby's platform. The Unreal Engine game environment embeds the Dolby plugin which then receives the stream from the Dolby site.

You can choose for the game to be either downloaded by the audience, or host it on a cloud server such as AWS which streams the game to the audience. This is particularly good as any device can access it - they don't need a gaming PC (this is called Pixel Streaming). It is a complex and involved process to do from scratch but there are platforms that offer a plug and play system. We experimented with Arcane Mirage. In talking with them it was shown that they can stream a game to the audience for a multi-player experience.

The more common use case of Unreal in live performance, is to stream body-tracking data onto a digital character in a digital world. This method tends to confine the performers into a greenscreen box, covered in cameras and/or motion suits. What we like about streaming video into the digital world is that the live performance can be experienced by an audience simultaneously. It's basically a multi-camera live stream viewing in a 3D space environment, for the users to have agency to explore and watch. The surfaces used to project the video feed on can be anything that has a material. For example:

1) a simple floor with the video feed from one camera as its material
2) a floating sphere playing the video
3) the material of an object/ a scanned model
4) the walls, floor and ceiling of a room. The list goes on - any digital object that has a material assigned to it (which is basically everything)

Pixel streaming resources:
Dolby.io documentation
Unreal setup
Arcane Mirage

2. Augmenting Reality in Live Performance

2.1 Projection-Based AR: Expanding Your Space

Once you have the mobile video feed coming into the PC, then it's a case of sending it through a software package and to the projector(s). There are multiple platforms that do this, our favoured option is Touchdesigner. The free version offers all the features with just the resolution limited to 1280/720p - so not quite full HD. It is a bit of a learning curve but once you're in, you won't be able to live without it. It can be as simple and complex as you need, and there are plenty of tutorials available online, and a thriving DIY community who are always happy to help on facebook. Here is a LINK to get you started for routing video footage from the phone to the projectors over NDI. It is simple then to create other videos live, or pre rendered, and switch between the two. The alternative is Resolumne Arena. But at £800 a pop and nowhere near the same flexibility, you really have to know you need it.

As a side note, within Touchdesigner, it has a projection Mapping tool, which, although slightly strange and cumbersome - it is fantastic considering its free. If you are mapping something quite complex, maybe a 3D structure, then we would also use Madmapper just for that part.

I am obsessed with projectors in performance, they can be so much more than big TV's. They can be a really dynamic lightsource, when shone through haze and using high contrast images and patterns, you can get amazing laser light visuals LINK. To achieve this look the only thing you have to be careful of, is that the final image is out of the way of the audience's eyeline. Angle the projector any which way, except how it would normally be positioned (behind the audience facing the stage).

The other fun thing to do with projectors is to throw the image over the performers and/or audience. This can give you a really nice layering effect and also nice technique to bring the audience into the work and make them part of it. It is of course important to consider safe projection angles so as not to continually blind the audience.

Bring a projector into the space: place it everywhere it shouldn't go and see how it interacts with the space. Also have a little mirror to easily send the angle of the projector around the space without having to move the projector itself.

- Aim for a projector at least 3500 lumens or more if possible, multiple if you can. Although brightness is great, if you are shining it through/over an audience brighter is not always better.
- The standard lamps are DLP but we would recommend to look for a class 1 Laser or LED lamps as they allow you to mount the projector any which way and don't get so hot (quieter).
- A note on resolution, it works a bit differently to TV's, anything over FullHD (1920x1080) is not really visible unless viewed from very close up. So save your money and go for Lumens (brightness) over resolution.

Projector Resources:
Great place to get info on projectors before heading over to Ebay

2.2 Mobile AR: Layering Digital Realities

This is a big one, and there are lots of ways to go about it. We have been playing with two specific types of AR which we will focus on here; a) image tracking, and b) Geo-located content

a) This is where a mobile app recognises an image in the real world and anchors a digital 3d object to it when viewed through a phone.
b) Where a digital object is located with reference to its gps coordinates, the object appears in the phone's view when the user arrives at that location.

Image tracking

A challenge using AR in a performance is that the lighting and set is likely to change dramatically throughout the performance. We came to the idea of using image tracking so that the performers can have control of the AR by moving images (a book, photograph, postcard etc). It is also fairly simple to ensure there is a light source on that image.

It is fairly straightforward to create a basic image tracking app using Unity game engine and there are tons of tutorials to get you started and side-load this onto the phone. If you want to just test the functionality out we would recommend using ArtVive as it's super quick and easy to get a prototype up and running.

A major consideration when thinking about this tech in performance, is of course - do you want all the audience getting their phones out in the show? So that's where live streaming with mobile phones becomes quite nifty. These can be the phones that pick up the image tracking AR and then the audience sees the digital overlay when the live feed is sent to the projectors. To make this work you need to send a live screen capture over the network. We would recommend NDIScreenCapture app for this purpose. It does suffer from screen tare a bit, but we like that aesthetic.

Geo-location

There are a few companies offering different technologies to get you started with GPS located AR. The two we have explored are Google Cloud Anchors and Niantic's Lightship. Both are installed as a plugin in Unity.

Google's account access and UI overlay is complex - even just getting started navigating around the web portal and understanding costings was extremely confusing and complicated. And at the moment it is free but that is likely to change, so beware!

Niantics AR platform Lightship is the main alternative to Google. Most of you probably used or heard of Pokemon Go - this is the platform that developed that game. There are plenty of tutorials online and a community of developers using it - and the online documentation and sample content is great to get you started. It is not without its frustrations - technology moves so fast - so even making sure

pairing the right version of Unity to the right version of Lightship is a complicated matter. But overall it's technology that combines GPS location and the scanned version of the real environment, making the tracking extremely precise!

Google's tracking feels less precise, but once you have got your head around their website account faff, it's much easier to implement than Niantic. So for AR that does not require extremely precise interaction with the real world I would settle for Google. Also it looks like Google will be integrating AR into their Maps app soon. Which could be a super direct and interesting way of engaging an audience in your AR experience. There have been updates since we used this and I believe they have improved on the stability of AR tracking.

AR Resources:
We would recommend checking both THIS and THIS channel out to get started with these Niantic and Google Geospatial AR.

2.3 Mixed Reality (MR)

We just want to plant a seed here because this is one of those exciting areas it's worth keeping an eye on, as tech keeps evolving. To quickly define the difference between Augmented Reality (AR) and Mixed Reality (MR) in big tech terms - while these terms are often used interchangeably, they have distinct differences that are helpful to understand. AR layers digital elements over reality using a device, whereas MR adds an interactive dimension, allowing participants to engage with both the real and digital layers simultaneously.

The new headset from Meta (facebook) offers an affordable way to start dabbling in mixed reality experiences (it's old enough now to pick them up on Ebay). It uses a pair of cameras on the front to capture the real world and then blends that image with digital elements right in your headset view. This kind of tech opens up a playground for creative experimentation - imagine live visuals that interact with you in real time! But it does still require you to wear an awful headset so that's an obvious con!

3. Immersive Audio for Live Performance

3.1 What is it?

Immersive audio as a term is pretty meaningless really. What audio isn't immersive?! When you are in the middle of a pumping gig or a delicate baroque instrumental, you may feel fully immersed in the experience. Maybe the uncomfortable wooden chair or the pint that was just accidentally tipped down your back, is trying to pull you out of that deep immersion. But essentially music is immersive no matter how many speakers you have. So why is 'Immersive Audio' as a form interesting/useful in live performance? And how can we avoid spending £200,000 on a proprietary system from D&B audiotechnik? Dear readers, this is where things get really nerdy!

3.2 Understanding Ambisonics

So, to understand why Immersive Audio, we need to understand Ambisonics a little bit.

"Ambisonics is a full-sphere surround sound format or a means of representing the sound field at a point or in space." Rode

So we use multiple microphones pointing in all different directions to capture a 360 sound field. This can then be reproduced using the speakers in a sphere formation. The more mics and speakers you have, the more audio resolution you can achieve. This goes up in order of magnitude.

1st order Ambisonics = 4 Mics/Speakers
2nd order Ambisonics = 8 Mics/Speakers
3rd order Ambisonics = 16 Mics/Speakers
And so on…

For live events we would say 3rd order Ambisonics to Immersive Audio is a bit like FullHD is to video; really great resolution, and for anything above this the law of diminishing returns really starts to kick in.

3.3 Considering Speaker Layout

The point where Immersive Audio becomes more immersive than just audio is when you start thinking about speaker placement. Particularly, placing speakers above the audience as this is a very unfamiliar position for audio to come from and can give the audience a whole new feeling.

In a perfect world, the speakers would be placed in a precise sphere around the audience - think geodesic polyhedron. This obviously isn't practical in many situations. So we settle for rings around the audience. For 2nd Order Ambisonics we might go with 6 speakers in a circle and 2 in the ceiling. For 3rd order we might go with 8 at just above head level, 2 directly above pointing down, and the remaining 6 in a circle between the lower circle and the top speakers. But this might change if you feel that sound coming from feet level is important for your project. And of course if your audience is suspended on a perforated floor it would be wicked to have speakers below them pointing up!

3.4 Why do it?

Besides the obvious (eg if you are a 360 video Dome or an audio soundbath experience) - the reasons for using it in a theatrical production are intriguing and specific and depend largely on the staging setup. Here's a few ideas to get started:

For the classic end-on setup that creates a distinct boundary between performance and audience, throwing the sound out and around the auditorium can be a really effective way of bringing the audience into the world; essentially, it's what cinema has been doing for years. But as soon as you start to break this format down, and spread the performance out/around/amongst the audience, immersive audio soon becomes a really crucial tool. Not least to represent a prerecorded soundscape, but to have the ability to follow the performance live, in real time around the space.

Example 1:

A performer with a radio mic is traveling from one performance area to another, but they want to keep speaking their soliloquy throughout their journey. We want their amplified sound to come from where they are, as best as possible. You can see how an immersive speaker setup would be able to do this. Similar to a manual folly spot.

Example 2:

The voice of a god

Example 3:

The audience are inside someone's head, hearing their thoughts spiralling (literally) into a cacophony of noise.

These are just a few quick examples, but they give you an idea of its use case. Lastly, the sense of space is very unique to immersive audio. The audio can really feel like it is hanging between the speakers or outside of the bounds of the performance space. One experience I had in Plymouth's Dome Market Hall, that uses, I believe, 96 speakers! There was a moment the audio felt like it was being produced actually inside my ear.

3.5 How to do it

So, besides spending ££££££ on a fancy L-acoustics or D&B system, how can we get this to work? There are a few ways, but the way we have tested and has worked well for us so far is using a free plugin in Ableton Live Suite called Envelop for Live, run over a Dante network.

Envelop for Live

It is designed primarily for live immersive music production, but it can function equally well for live theatre, especially when musicians are involved. It can work with almost any speaker setup, and there are plenty of tutorials online. It's possible to either pre-animate the movement of sound, or move it live with a midi controller.

Here's a LINK to get you started and can't recommend Michael G Wagner enough for all things immersive audio

One key difference between this system and a D&B system is latency (delay). From the note being sung, to that sung note being sent out of a speaker. D&B use clever processing to achieve this super fast, just a couple of milliseconds - so let's call that zero latency. They use a Dante network to help achieve this, which essentially means sending audio using ethernet cables, and of course their main processing brain of the system is dead quick.

Now we can use Dante too…

Dante

The nuts and bolts of Dante are many and thus beyond this toolkit. But we will touch on the components you need for an absolute minimal setup.

1) On the PC running Ableton and Envelop 4 Live you need a Dante Virtual Sound card (£50). This does the job of an audio interface, but can achieve up to 128 inputs/outputs.

2) Then you plug in the ethernet cable (pref. Cat6 or above). Either send this to a Dante-enabled mixing desk or directly to a (preferably Dante certified) managed Network switch with QoS dealt with properly. THIS thread is a good starting point.

3) From this network switch you can send a separate cat6 cable to the speakers, and at each speaker you then need a Dante to analogue converter, something like this: LINK.

You can do all of this without needing to use Dante, but you just need a lot of available analogue inputs and outputs and the cable runs get expensive compared to Cat6. Just a word on Dante network that it will introduce extra latency per hop (between switches/devices) so be careful when adding in switches.

DIY project tangent… I have to share THIS incredible DIY project that achieves absolute zero-latency wireless audio using super affordable components. The core idea? It transmits optical audio signals by bouncing light off small mirrors to reach the speakers, where optical-to-analog converters handle the final output. The only catch is that it requires a clear line of sight. Even if you don't plan to build the full speaker units (though they're pretty cool too), the wireless audio method alone is worth checking out!

3.5 A final thought

I want to wrap up this section with something a little personal. I'm stone-cold deaf in my left ear, officially called a "Dead Ear", and have been for 11 years. One of the biggest impacts of this is that I've lost the ability to spatialise sound. I can hear whether something is loud or soft, but I can't tell where it's coming from. My brain has adapted through deduction, but in an immersive audio environment, my experience is fundamentally different from most people's.

On paper, spatial audio might seem pointless for someone with single-sided deafness (SSD). But I'd challenge that assumption. Having all sound processed through one ear is exhausting, and I often switch music to mono just to make sure I hear everything but that does cause listening fatigue. What I've found in immersive audio spaces is that they actually relieve this pressure. The sound world feels broader, more open, less like everything is being funnelled into one ear at once.

It's a bit of a tangent but I just thought this was something to consider when thinking about access and inclusion in immersive audio. Alongside this, it's always worth thinking about sub-bass levels, giving people the chance to literally feel the music through sound pressure and floor vibrations. Immersive experiences aren't just about what we hear, but how we sense sound in different ways.

4. Toolkit conclusion

This toolkit is meant as a hands-on guide to begin thinking about integrating immersive media into performance using accessible, low-cost tools. It's not about high-end production values or the latest tech for tech's sake, it's about finding practical, flexible solutions that serve the work. Whether you're live-streaming from a phone, projecting AR elements into a space, or experimenting with audience agency in a virtual environment, the focus is on making these technologies work in real-world performance settings. And this toolkit just scratches the surface of what's out there to experiment with.

None of this is plug-and-play; every setup will need testing, tweaking, and adapting to suit the demands of the project. But that's part of the process, learning by doing, pushing the limits of the tools, and figuring out what fits. The aim is not to replace live presence but to extend it, to bring audiences into the work in new ways, whether they're in the room or connecting remotely.

If there's one takeaway, it's this: don't wait for the 'right' equipment or a perfect setup. Start with what's at hand, experiment, and build from there.

And to plant a last idea - look into Touchdesigner and the MediaPipe Toolkit, alongside StreamDiffusion to begin experimenting with AI in live performance.

Glossary

360 Camera: A camera that captures video in all directions at once with two or more lenses, creating an immersive viewing experience. Useful for live-streaming performances where the audience can look around in real-time.

Ambisonics: A method of recording and playing back 3D sound, allowing audio to come from any direction. It's often used in immersive audio setups to create a more realistic sense of space.

Augmented Reality (AR): Derived from the Latin augmentare, meaning 'to increase' or 'to enhance.' Its helpful to consider augmentation in its broadest sense, as something greater in size, value, or effect. In technology, this often relates to Augmented Reality (AR), where digital elements are layered onto the real world viewed through a device. In music, an 'augmented' interval is one that uses notes outside of its original key to expand an interval, this tends to create a sense of tension or unease. This dual meaning, both enhancement and tension, offers an interesting lens through which we can consider augmentation in performance and immersive media.

Dante: A digital audio networking system that allows multiple audio sources to be connected over a network using ethernet cables instead of traditional analogue connections.

Dolby.io: A WebRTC-based platform that enables high-quality real-time streaming. It can be used for integrating live video feeds into digital or mixed-reality environments.

Envelop for Live: A free plugin for Ableton Live that enables immersive audio by allowing sound to be positioned and moved in 3D space.

Geo-Located AR: Augmented reality content that appears at a specific real-world location, often using GPS to determine where digital elements should be displayed.

Immersive Audio: Audio that surrounds the listener, often using multiple speakers to create a sense of space and movement. It can be used to place sounds in specific locations around the audience.

Image Tracking: A way for AR applications to recognise and anchor digital objects to specific real-world images (like posters, books, or objects) when viewed through a device.

Lightship (Niantic): An AR platform that enables developers to create precise, location-based augmented reality experiences. It's the same technology used in Pokémon Go.

Live Streaming: Broadcasting video in real time over the internet. In performance, this can be used to extend the audience experience beyond the physical space.

Mixed Reality (MR): A blend of real and digital worlds where virtual objects interact with the real environment in real-time, often viewed through headsets like the Quest 3.

NDI (Network Device Interface): A protocol that allows video to be transmitted over a local network with low latency. Useful for integrating multiple live video sources in performance.

Passthrough AR: A feature in mixed reality headsets that allows the wearer to see the real world using integrated cameras while also viewing digital elements overlaid on top.

Pixel Streaming: A method for running high-quality 3D applications remotely on a cloud server and streaming them to a user's device, reducing the need for powerful hardware.

Projection-Based AR: Using a projector to display digital images or effects onto real-world surfaces, transforming the space dynamically.

Remote Audience: A remote audience is a group of people who are reached and engaged using digital technologies.

Site-Specific performance: A site-specific performance is a performance that takes place in a non-traditional space, such as a park, warehouse, or house. The site's characteristics are used to enhance the story.

SRT (Secure Reliable Transport): A streaming protocol used for high-quality video transmission over the internet, often used in professional live streaming setups.

TouchDesigner: A visual programming tool from Derivative used for live visuals, projection mapping, and interactive media. Often used in performances for real-time video manipulation.

Unreal Engine: A powerful game engine that can be used for creating virtual environments, real-time rendering, and integrating live video into digital spaces.

WebRTC: A protocol that enables real-time video and audio streaming with low latency, commonly used for online meetings and live streaming applications.

Find out more about us and how we use this tech:

www.commonwealththeatre.co.uk

https://nmason.xyz/