A Framework for Classifying Creative Technologies

Andrew Chitty, Graham Hitchen, Conor Roche

March 2021

This Creative Technology Framework was developed in partnership with the Audience of the Future and Creative Industries Clusters team at UKRI to facilitate new collaborations between technology researchers and creative practitioners, and to help shape thinking and future investments in the Creative Technology space.

CRAIC is working with UKRI and with Conor Roche at BOP Consulting, to develop an interactive version of the Framework – creating a search function, and enabling researchers to cross-reference research and other projects being undertaken which are testing or developing the different practices and processes described in the document.

Intro

This document presents a framework for the purposes of identifying and classifying creative technologies. This includes the underlying technologies, technology-enabled applications, related areas of research and disciplines that are currently applied in creative practice across the creative industries. This work focuses on emerging technological trends and advanced technologies, not technologies that already have high adoption in the creative sectors.

For the purposes of this exercise, we have attempted to capture some of the most relevant and significant technologies relating to the use of advanced technologies in creative practice. **However, this framework does not contain nor attempt to contain an exhaustive list of all technologies, the intention for the framework is to provide a means to identify and classify technologies in areas of creative practice.**

We have classified the technologies, applications and disciplines for areas of creative practice under the following five headings:

1. **Content Creation and Production:** The creation and co-creation of new content, including technology assisted performance-based production activity
2. **Content Enhancement and Production Workflows:** Improving and up-sampling of existing content and media, and the use of technologies in production workflow and project management
3. **Media Storage, Extraction and Analysis:** The storage, extraction, use and analysis of content and media for creative applications
4. **Connectivity:** The technology enabled connection of creators and users, usually at distance - for the purposes of the creation and consumption of creative works
5. **Presentation and Consumption:** The technology enabled presentation and consumption of creative content for audiences

For the purposes of this work, we use the following informal working definitions:

- **Creative Practice:** a catch all term related to the creation or consumption of creative works
- **Discipline:** an area of research, a field of academic knowledge or business category
- **Application:** A programme or tool, typically software based, that enables the user of the application to carry-out specific tasks. The inner workings of the application/tool i.e., the machine/the code – is typically hidden from the user
- **Technology:** the practical manifestation of science for example as a machine, software or hardware based, or a material
- **Use case:** a specific situation in which an application, a technology or collection of technologies is used within an area of creative practice

This document includes an appendix that illustrates how this framework could be applied to the area of augmented reality.

1. Content Creation and Production

Discipline	Applications	Technologies	Use cases
Computer Vision (CV): a field of study and technology development that seeks to develop techniques to help computers "see" and understand the content of digital images such as photographs and videos. CV involves the use of specialised machine learning methods and makes use of general learning algorithms.	**Capture and Motion Prediction:** ML is particularly well suited to learning models of motion from captured real motion sequences. These motion characteristics can be learnt using deep learning based approaches.	**Machine Learning (ML):** The key ML technologies applied in creative practice include: - **Convolutional Neural Networks (CNNs):** a class of deep feed-forward artificial neural networks that are designed to take advantage of 2D structures, such as found in images - **Generative adversarial networks (GANs):** a recent algorithmic innovation that employ two neural networks: generative and discriminative. The GAN pits one against the other in order to generate new, synthetic instances of data/media that can pass for real data/media. - **Recurrent neural networks (RNNs):** widely employed to perform sequential recognition tasks; this makes them amenable to tasks, such as speech recognition, handwriting recognition, and music generation. - **Deep Reinforcement Learning (DRL):** a machine learning algorithm trained to make a sequence of decisions based in a virtual environment. Useful in creative applications where there may not be a predefined way to perform a given task, but where there are rules that the model has to follow. Current applications typically include gaming, story-telling and caption-from-image generation	**Rendering Synthetic Avatars, Objects and Scenes:** This includes the synthesis of 3D views from motion capture or from monocular cameras, shading and dynamic texture synthesis. Creating realistic lighting in animation and visual effects has also benefited by combining traditional geometrical computer vision with enhanced ML approaches and multiple depth sensors. For example, Google research has created software for pose animation that turns a human pose into a cartoon animation in real time. This is based on PoseNet and FaceMesh. Adobe has also created Character Animator software offering lip synchronisation, eye tracking and gesture control through webcam and microphone inputs in real-time. This has been adopted by Hollywood studios and other online content creators.

Natural Language Processing (NLP): analyses natural language data and trains machines to perceive and to generate human language directly. This is a machine learning process and NLP machine learning algorithms frequently involve speech recognition, natural language understanding, and natural language generation.		*ML*	**Script Writing:** The script of a fictional short film, Sunspring (2016), was entirely written by an AI machine, known as Benjamin, created by New York University. The model was trained using science fiction screenplays as input, and the script was generated with random seeds from a sci-fi filmmaking contest. In the sequel, It's No Game (2017), Benjamin was then used only in selected areas and in collaboration with humans, producing a more fluid and natural plot. Script Book introduced a story-awareness concept for AI-based storytelling. The generative model focus on three aspects: awareness of characters and their traits, awareness of a scripts style and theme, and awareness of a scripts structure, so the resulting script is more natural.
Synthetic Media: The artificial creation, production, manipulation, and modification of content and media, typically by automated means, through the use of artificial intelligence algorithms	**Synthetic gaming environments:** An AI generated synthetic environment that represent activities, places and environments at a high level of realism, e.g. Vid2Vid which uses a deep neural network, trained on real videos of cityscapes, to generate a synthetic 3D gaming environment.	*ML*	**Generative interactivity and interactive narrative:** - MADE (a Massive Artificial Drama Engine for non-player characters) generates procedural content in games, supporting design, decision-making and interactivity for games. - AI Dungeon is a web-based game that is capable of generating a storyline in real time, interacting with player input. Procedural generation has been used to automatically randomise content so that a game does not present content in the same order every time.
Generative Content: Content or art that in whole or in part has been created with the use of an autonomous system / AI	**Music co-creation:** Applications of AI in this domain include searching through large databases to find the most appropriate match for sound design and audio production purposes. There are several AI assisted music composition systems that support music creation. The process generally involves using ML algorithms to analyse data to find and suggest musical patterns.	*ML*	**Music co-creation:** Flow Machines by Sony, Jukebox by OpenAI19 and NSynth by Google (developed using Raspberry Pi)

	Image Generation and Augmentation: create new digital imagery or art-forms automatically, based on selected training datasets e.g. cartoon characters, celebrity headshots.	*ML*	**Image Generation:** pix2pix has been used to create a renaissance portrait from a real portrait photo **Image augmentation:** DeepArt transforms an input image into the style of the selected artist by combining feature maps from different convolutional layers. A stroke-based drawing method trains machines to draw and generalise abstract concepts in a manner similar to humans. **Neural Rendering:** Neural rendering is a new class of deep image and video generation approaches that enable explicit or implicit control of scene properties such as illumination, camera parameters, pose, geometry, appearance, and semantic structure. It combines generative machine learning techniques with physical knowledge from computer graphics to obtain controllable and photo-realistic outputs.
Computer engineering: The integration of electronic engineering with computer science.		**Graphics Processing Unit (GPU):** A specialized electronic circuit designed to accelerate the processing and creation of digital images. GPUs are used to process the kind of calculations a machine learning neural network needs. **Micro-controllers and Single Board Computers:** Micro high-powered computers often used with a variety of motion-sensor in location based immersive experiences and live performances, typically for the purposes of interacting with and controlling live media and content	**Micro-controllers**: Raspberry PI, Arduino - used to develop custom performance and production interfaces and robots.

Cognitive computing: Simply described as assisting humans to make decisions typically through AI and potentially using human brain interfaces, interactive cybernetic systems and emotion sensing technology.	**Emotion recognition:** AI methods have been used to learn, interpret and respond to human emotion, via speech (e.g. tone, loudness, and tempo), face detection (e.g. eyebrows, the tip of nose, the corners of mouth), and both audio and video. **Adaptive game-play experiences** respond to the player and their current psychological state **Creation of artificial agents** based on human behavioural data, and psychological cues **Informing and assisting creators** in the creative and production process.	*ML*	
Virtual environments: The development of applications that allows a user to interact with a computing, often immersive, environment and the work of other users	**Virtual 3D Interfaces for composition and performance:** The design, development and use of virtual interfaces for immersive environments, typically AR or VR, that enable the creation and production of creative works in that virtual environment.		**Virtual 3D Interfaces for composition and performance:** - Soundstage VR - Google TiltBrush

2. Content Enhancement and Production Workflows

Discipline	Applications	Technologies	Use cases
Distributed computing and systems: a field of computer science that studies systems whose components are located on different networked computers, which communicate and coordinate their actions by passing messages to one another from any system	**Cloud based production workflow platforms including:** - **Cloud based Creative AI services:** Cloud based services and programming modules configured specifically for creative applications These services enable the remote storage and analysis of media and content. - **Online Digital Audio Workstations (DAWs):** Cloud and browser based music and audio production platforms enabling real time cloud based co-creation and co-production of music and audio. - **Cloud rendering:** Cloud services specifically designed to enabling high-end video and animation based rendering, integrates directly with the most popular animation and production tools	**GPU**	**Cloud based Creative AI services:** - Google Cloud and Tensor Flow, - AWS and Augmented AI **DAWs:** - soundation.com - Soundtrap (Spotify) **DAWs:** - soundation.com - Soundtrap (Spotify)
	Game Engines and Frameworks for non-Game Production Workflows: A software-development environment originally designed to build video games, and have now been adopted to work in the immersive, film, architecture and performing arts sectors. These frameworks also integrate with lighting and audio production platforms.		**Game Engines:** - Unreal is being used by the RSC to capture and render in real-time a virtual performance environment for the immersive production entitled Dream
	AI assisted Visual Special Effects (VFX): The use of ML-based AI in VFX has increased rapidly in recent years. These both use a combination of physics models with data driven results from AI algorithms to create high fidelity and photorealistic 3D animations, simulations and renderings. With ML-based AI, a single image can be turned into a photorealistic and fully clothed production-level 3D avatar in real-time.	*ML*	

Audio, Image and Video Processing: addresses the problems of acquisition, storage, retrieval, processing and production of audio, images, videos, and high dimensional signals for extraction, analysis and enhancement for human users, robots, and autonomous systems.	**Contrast Enhancement:** Contrast is an important factor in any subjective evaluation of image quality. Low contrast images exhibit a narrow range of tones and can therefore appear flat or dull. Contrast enhancement uses ML to synthetically enhance the contrast of an image or other media **Colorization:** The process that adds or restores colour in visual media. **Super resolution imaging:** a technique that enhances and increases the resolution of an image and video. Enabling the up sampling of images and video spatially or temporally. **Restoration:** Divided into five classes - deburring, demonising, dehazing, mitigating atmospheric turbulence and in-painting **Ray Tracing:** is a rendering technique for generating an image by tracing the path of light as pixels in an image plane and simulating the effects of its encounters with virtual objects. **Spatial Audio:** is audio which gives you a sense of space beyond conventional stereo, allowing the user to pinpoint where sound is coming from, whether this is above, below, or a full 360 degrees	*ML*	

3. Media Storage, Extraction and Analysis

Discipline	Applications	Technologies	Use cases
Artificial Intelligence	**AI Training Data Generation and Management:** Services that enable AI developers to model and generate specific training sets for their machine learning tools.	*ML*	**AI Training Data Generation and Management:** - Amazon currently offers a cloud management tool, SageMaker, that uses ML to determine which data in a dataset needs to be labelled by humans, and consequently sends this data to human annotators through its Mechanical Turk system or via third party vendors. This can reduce the resources needed by developers during the key data preparation process. - Mostly.AI: Creates realistic, privacy-safe synthetic datasets that are compliant with the data protection laws.
CV	**Content Retrieval and Semantic Representation:** Modern image retrieval methods often employ deep learning techniques, enabling image to image searching by extracting low-level features and then combining these to form semantic representations of the reference image that can be used as the basis of a search.	*ML*	
	Segmentation and Recognition: Segmentation methods are widely employed to partition a signal (typically an image or video) into a form that is more meaningful and easier to analyse or track. The resulting segmentation map indicates the locations and boundaries of semantic objects or regions	*ML*	**Object detection:** YOLO and its variants represent the current state-of-the-art in real-time object detection and tracking.

Audio, Image and Video Processing	**Image and video compression:** The demand for increased qualities and quantities of visual content is particularly driven by the creative media sector, with increased numbers of users expecting increased quality and new experiences. Deep neural networks have gained popularity for image and video compression in recent years and can achieve consistently greater coding gain than conventional approaches. Deep compression methods are also now starting to be considered as components in mainstream video coding standards.		

4. Connectivity

Discipline	Applications	Technologies	Use cases
Distributed Computing and Systems **Next Generation Connectivity:** refers to anticipated technological advances which will offer greater capacity and coverage for internet connectivity, both fixed and mobile.	**Fog/Edge Computing** are concerned with leveraging the computing capabilities within a local network to carry out computation tasks that would ordinarily have been carried out in the cloud and at distance. This enables high speed computing tasks to happen locally, potentially providing low-latency connectivity for mixed reality experiences.	5G mobile networks, full-fibre broadband and satellite internet access *GPU*	**Ericsson Private networks:** a solution that integrates 5G and edge computing technologies, and other advanced connectivity and rendering technologies to provide AR and VR streaming and connectivity for businesses and users **NVIDIA CloudXR:** is NVIDIA's solution for streaming virtual reality (VR), augmented reality (AR), and mixed reality (MR) content from any OpenVR XR application on a remote server
Audio, Image and video Processing	**Streaming Media and Technologies:** media and technologies that allow for the constant consumption and presentation to an end-user while being delivered by a provider.	*GPU*	

5. Consumption and Presentation

Discipline	Applications	Technologies	Use cases
Interactive Design: the practice of designing interactive digital products, environments, systems, and services.	**Interactive and non-linear film, "TV" and video:** Interactive video and non-linear story-telling is not new, but with Netflix increasingly experimenting with the format for kids shows and mainstream series like Black Mirror, it has become more common place.		**Interactive film:** requires a new form of video production workflow, for this process Netflix develop their own development product called "Branch Manager".
Distributed Computing and Systems	**WebXR:** An API that provides access to input (pose information from headset and controllers) and output (hardware display) capabilities commonly associated with Virtual Reality (VR) and Augmented Reality (AR) devices. It allows the development and presentation of VR and AR experiences on the web through a browser		XRSwim is a curated list of WebXR experiences
User Interface Design and User Interface Engineering is the design of user interfaces for machines and software, and other electronic devices, with the focus on maximizing usability and the user experience.	**VR/AR Body Control:** body-based navigation technology that puts the body in direct control of movement through virtual space. Using sensors for first-person view navigation free of hand controls. Interactions are designed to respond to natural bearing and balancing instincts avoiding motion sickness.		**VR/AR Body Control:** BodyNav is a developer kit that enables developers to integrate body controls into XR experiences
	Wearables and trackers: The integration of tracking devices and sensors into jewellery, clothes and other everyday items		**Wearables and trackers:** Leaf Chakra, a tracking pendant and necklace you wear, it tracks your activity, sleep, meditation and general health.
	Haptic technology: refers to any technology that can create an experience of touch by applying forces, vibrations, or motions to the user		**Tesla suit:** full body haptic suit and training solution for physical VR experiences
Electronic Engineering		**MicroLEDs** allow for next-generation ultra-thin and low power AR displays and glasses using an array of LED that act as a micro image projector	**MicroLEDs:** Vuzix MicroLED smartglasses
	In-car entertainment for autonomous vehicles: The emergence of autonomous vehicles enables a new "stage" for audiences to access and experience entertainment		**In-car entertainment:** Audi teamed up with Disney to produce an in-car VR experience for backseat passengers, called "Marvel's Avengers: Rocket's Rescue Run"

Distributed Computing and Systems	**Distributed Ledgers (DLT):** a consensus of replicated, shared, and synchronized digital data, geographically spread across multiple sites, countries, or institutions - there is no central administrator	**Blockchain:** a form of DLT, that is a growing list of records called blocks, that are linked together using cryptography	**CryptoArt:** digital artworks, linked to unique tokens that exist on a blockchain, the tokens are in the form of a non-fungible token (NFT), which makes the ownership, transfer, and sale of an artwork possible in a cryptographically secure and verifiable manner
	Content and Media Authentication, Rights Management and Provenance: Technologies that enable a creator to be securely and robustly attributed to the content they create, often using DLTs	*ML*	**Content and Media Provenance and Authenticity:** Coalition for Content Provenance and Authenticity (C2PA) an initiative founded by Adobe, Microsoft, Truepic, Arm, Intel and the BBC is an organization that will advance the work of open specification development toward broad adoption of provenance standards.

Appendix: Augmented Reality

This section applies the framework outlined above to the area of Augmented Reality. The purpose of this section is to highlight how the framework might be applied to an area of creative practice so as to identify the ecosystem of interrelating research disciplines and technologies involved. This exercise doesn't attempt to exhaustively list all of the technologies involved in Augmented Reality, it intends to illustrate how the framework can be applied.

Content Creation and Production			
Discipline	Applications	Technologies	Use cases
Computer Vision	Capture and Motion Prediction	*Machine Learning*	FXMirror - AR Fitting Rooms
Cognitive Computing	Adaptive game-play experience	*Machine Learning*	

Content Enhancement and Production Workflows			
Discipline	Applications	Technologies	Use cases
Distributed computing and systems	Cloud based Creative AI services Cloud Rendering	*GPU*	Google Cloud and Tensor Flow AWS and Augmented AI Google Cloud and Tensor Flow
Audio, Image and Video Processing and Computer Graphics	Ray Tracing Spatial Audio	*ML*	 Fields - a spatial sound creation tool that uses augmented reality to turn your surrounding space into a sonic canvas.

Media Storage, Extraction and Analysis			
Discipline	Applications	Technologies	Use cases
Computer Vision	Content Retrieval and Semantic Representation Segmentation and Recognition	*ML* *ML*	
Audio, Image and Video Processing	Ray Tracing	*ML*	JPEG XS

Connectivity			
Discipline	Applications	Technologies	Use cases
Distributed Computing and Systems Next Generation Connectivity	Fog/Edge Computing	5G mobile networks, full-fibre broadband and satellite internet access GPU	Ericsson Private networks NVIDIA CloudXR
Audio, Image and video Processing	Streaming Media and Technologies	*GPU*	

Presentation and Consumption			
Discipline	Applications	Technologies	Use cases
Distributed computing and systems	WebXR		
User Interface Design and User Interface Engineering	VR/AR Body Control Haptic Technology		BodyNav Tesla Suit
Electronic Engineering		MicroLEDs	Vuzix MicroLED smartglasses