The global data palette: massive databases and the reformation of content creation in film/video and music/sound art practice

Jefferson Goolsby and Dr Mei-ling Lee
University of Maine, Orono, USA

Introduction: the global data palette

For film/video and music/sound arts practice there is an emergent production reformation that represents a significant shift away from the intrinsic and fundamental restrictions of twentieth century models of time-based image and sound composition. After 100 years of incremental advancement, current moving image and sound production developments offer a radical shift not only in production technologies but also distribution and delivery systems, and what may prove most significant - aesthetic values and paradigms. Not only how we create, but how we acquire, own, store, and engage with visual and sonic materials is clearly changing. The global data palette proposes that time-based media artists will move away from the resource intensive aspects of content capture and redirect production emphasis to the post-production construction of projects by working primarily with existing content from massive databases. Such a shift places digital media creators in a substantially altered environment of creative practice, and offers greatly changed production, economics, and aesthetic engagements previously unavailable.

Several technological developments have initiated this reform, particularly access to means of production (cameras, computers, etc.) and access to means of distribution (the Internet). As access to means of production and distribution becomes ubiquitous, three emerging developments function as a critical subset in the transformation of production practices. These subset changes are: 1) the affordability of massive data storage; 2) the restructuring of data ownership; and 3) the malleability of data. Through these three developments in storage, ownership, and malleability, a fundamentally changed production model may be decisively enabled.

Affordability of massive data storage

Data palette production originates in the ability to affordably store, and quickly access, immense quantities of data on independent hard drives or RAID (Redundant Array of Inexpensive Disks) systems in one's studio. By creating massive libraries of time-based sound and image content, the artist/composer is able to create works without the necessity of capturing raw data materials.

The remarkable changes in storage capacities and costs over the last 30 years have created the necessary conditions for affordable massive data storage. In US dollars, $500 in 1989 would buy 20 megabytes of hard drive storage, $4,500 less than that amount of storage had cost eight years previous. Ten years later, in 1999, that $500 would buy 1000 times more storage: 20 gigabytes. Ten years after that, $500 in 2009 buys 250 times more storage than 1999: over 5000 gigabytes (five terabytes). The megabyte-per-dollar change in cost has been dramatic (Historical Notes 2008). For video and sound artists working back in 2000, the availability of eight gigabytes of storage felt remarkably empowering, while the notion of an affordable terabyte of storage seemed yet unimaginable. Today, that prized eight gigabytes that cost several hundred dollars in 2000 can be found on a $20 thumb drive that gets lost at the bottom of a handbag, while eight terabyte home RAID systems are readily affordable in near plug-and-play configurations.

Considering the rate of storage increase from 1999 to 2009 it can be loosely anticipated that in 2019, $500 might buy 1,250 terabytes (1.28 million gigabytes or 1.22 petabytes) of storage. While there is a slowing of the precipitous cost-per-megabyte decrease and a need for increases in performance to accompany capacity growth, today's storage costs and capacities strongly indicate that we will soon reach the point of being able to reasonably afford as much storage as needed. Through this affordable massive storage capacity the working artist can begin to assemble enough raw materials for years of work, minimising the need to go into the field or studio for content capture.

Though feasible to implement, such necessarily localised massive databases still have data-quantity limitations and represent a first wave conversion that is archaic on arrival. Instead, such a model simply moves storage towards a tremendously more expansive second wave database structure that connects a global network of individual databases. This second wave model of networked community databases

will assemble enough raw aesthetic data resources to provide filmmakers/composers with a lifetime of raw content to draw from, nearly eliminating the need for project-specific field or studio content capture. Such cloud computing, as Ramnath Chellappa has termed it (Chellappa 1997) would follow his paradigm wherein 'the boundaries of computing will be determined by economic rationale rather than technical limits.' This ever expanding, dynamic repository of real-time accessible uncompressed data will provide a global database of raw materials to be perpetually accessed, modified, culled, and reformed into distributable content.

Three readily apparent technological issues to be resolved in completing the data palette framework include: further increases in storage capacities and performance to accommodate fully uncompressed data; real-time data flow and exchange speeds for that data; and automated identification systems that allow for searching and sorting resources, with provisions for micro-payments to content generators. The exceptional logistics of cataloguing, tracking, identifying, and filtering such massive quantities of aesthetic data will require a semi-automated system that begins at the hardware level and is able to configure terms of ownership as well as content, and importantly, allow for access to all links in any chains of source files, whether raw, processed, or reconstituted.

Restructuring of ownership

To attain the quality and pervasiveness needed to be sufficient, these community databases require the support of redefined cultural and legal structures of ownership. Revised ownership paradigms need to allow for fair trade and use of raw and processed content materials based on community market valuations and authorial prerogative in an authentic free market environment. Despite some problematic limitations, licensing systems such as Creative Commons begin to move the culture and law of ownership and licensing in this needed direction. Particularly necessary and presently insufficient in the Creative Commons structure will be unimpeded access to the entire chain of source materials, whether raw or processed (Bell 2009). Access to such source files will enable micro-revenue distribution configurations for content uses that are technologically, legally, and culturally fair and reasonable. On-the-fly use and micro-payment revenue calculations can be balanced, allowing for multiple content generators to opt-in for participating in revenue sharing with creators of distributed works. For example, embedded ownership and licensing tags would allow a project to be constructed from several hundred or even thousand content

sources yet permit micro-payments based on market driven valuations. Commercial and public use can be defined under license or work for hire agreements that provide revenue distributions that are market driven and undistorted, forming a production-consumption exchange that exemplifies a functioning micro-mass market.

Malleability of data

Despite its material and aesthetic qualities, analogue technology suffered from a fixed rigidity of content - an 'embeddedness' and incompliancy of the content within the media. The exposing and developing of film or the laying of an analogue audio track resulted in a sequence of images or sounds with burdensome limits on post-production restructuring possibilities. Whether visual or sonic, these post-production limitations channelled creative emphasis towards the original 'take'. Once the take was completed, the primary evaluative consideration for audio was whether to keep or discard the take, and for film whether to use or not use the take, generally in its entirety. Thus a necessary pre-eminence was given to the capturing of raw materials with a corresponding de-emphasis on post-production interventions and processing. Conversely, working in digital media with uncompressed data, software and hardware tools enable a malleability of content - an 'unembedding' from the media - that dispatches with the need for idealised, overemphasised content capture, and instead moves principle production decidedly from the field to the post-production environment.

With improved processing tools and advanced artistic control over those tools, content capture remains foundational but loses its privilege. Data palette production allows content to be not only modified, but primarily constructed in post. The 'take' becomes informal, a simple process of gathering data, with the analogue ideal of a perfect take replaced by the digital ideal of a constructed take assembled from a variety of raw materials, allowing for a greater scope of possibilities in final content. Thus raw materials need only be adequate to function as a starting point of production, and of a high enough resolution to allow substantial processing and reconstruction. The capture of content no longer needs to be exact or literal. Through this increasing versatility and malleability of uncompressed data, a 'digital materiality' comes forth - a data physicality of sorts - allowing aesthetic data to fully function as a raw material, like iron-ore that can be alloyed and fabricated into myriad materials and shapes as directed. Just as a sculptor uses fundamental raw materials, so the

day nears when the video/sound artist simply requires only generally suitable raw (data) materials.

Such data-priori ordering has been inherent to the rap music genre since its inception, and commonly emerges in recent feature film productions where financial and human resources are concentrated in a late capitalist market-driven effort to advance tools, materials, and product. From motion capture to the body scanning of actors to the physical simulations of hair movement (Docter, Silverman 2001), tools and strategies of data-priori production then filter across to consumer-level applications.

Drawing from a global database of content, works can be constructed and processed using pre-set, pull-down, and configurable parameters to achieve desired content outcomes. This ability to extensively combine, redefine, and reshape raw data relocates the traditional roles of production, moving the task of content capture from its position of primacy to a subjugated role in the creative production process. Content capture then becomes something of a manual labour, a sort of 'field work' that can be outsourced to a globally market-driven labour force. This shift encourages ideas rather than production regimen to again move to the fore, and in a sense purifies artistic intent. Content previously moored in media - such as image focal point, depth of field, framing, and eventually even camera angle - becomes 'unfixed', so the director of a project will find it unnecessary, even constraining, to generate determinate raw imagery.

Aesthetic transformation

These developments enable an analogous aesthetic reformation in time-based arts for both creator and consumer. In all arts practices there exists an elemental relationship between the tools, materials, economics, and production processes that, interacting with culture, determines and directs the aesthetics of a form. The constricted - and by today's standards acutely attenuated - practices of twentieth century film and music composition are painfully manifest in the dull redundancy of form and narrative that has resulted in a 'malaise of sameness' for the commercial film and music industries. Similar redundancies, familiarities, and excesses of self-referential ironies have settled upon time-based fine arts practices: found materials works have been bound by a rigorous indebtedness to source materials; avant-garde works have often been stymied by formalism; and so forth for various genres.

At the same time, technologically driven shifts in time-based media consumer behaviours have been phenomenal and well-documented, from on-demand content to iPods to the 'YouTubing' of distribution to mobile content delivery and more. Along with these and other changes in consumer and delivery conventions, user acceptance - and demand - for altered constitutions of content has initiated a generous allowance for a reformed aesthetic of time-based media.

An additional source of advancement is found in the hybridisation of animation and live-action imagery in major film productions such as Pixar's *Wall-E* and Columbia Picture's *Spiderman* series (Stanton 2008, Raimi 2002). While economically bound to narrow forms of narrative, these endeavours contribute tremendously to the technological changes that will eventually enable seamless data palette constructions. By current trends, it can be expected within the next 20 years that computer-generated animated films will successfully cross the uncanny valley and become visually indistinguishable from live-action cinema. Correspondingly, this capability of creating 'live-action animation' will be distributed to mass market tools. With those tools in hand and access to a massive global database, post-production ascendancy will be firmly established.

While the construction of 'seamless' distributable content assembled from a disparity of source materials will become achievable, more questionable is whether the culturally constructed valuation of the seamless - the cinematic 'real' - will persist. The 'invisible' edit and the ideal of 'realism' as admired by André Bazin (Bazin 1967: 1, 23-40) may yet be another casualty - if not already - of post-production predominance, to be as outmoded as the 'perfect' take. A relaxing of consumer expectancy with a reinterpretation of how time-based content threads and suggests its depictions, narratives, and constructions seems a more appropriate response to advanced visual and sonic remix forms. While the twentieth century aesthetic yet casts a shadow across early twenty-first century works, inevitably a new aesthetic takes shape that more accurately reflects contemporary lived experience and relationships between individuals and cultures.

Conclusion

Contained within the macro-revolution taking place in means of production and distribution, this paper has looked at an important three-part subset evolution of data storage, ownership, and malleability. The arrival of these technologically enabled

conditions provides a reformed structure and approach to the creation of time-based works and frees film/video and music/sound artists from the need for raw content generation. The data palette allows for a global community of fair and reciprocal content exchanges along with micro-revenue distribution in an authentic free market economy. As importantly, these changes in production encourage an attendant shift in aesthetics and a potentially dramatic reform of creative practice and cultural engagement distinct from the atrophied practices of twentieth century time-based arts. By alleviating production limitations, whether from an orthodoxy of tools, rigidity of media, or constricted ownership and distribution options, global data palette production avails to all genres an accessible production methodology. From there, new voices are allowed to emerge, and time-based arts practices are provided an environment in which to thrive.

References

Bell, John Patrick. Personal interview. 22 April 2009.

Bazin, André. "What Is Cinema?" 1967. *Defining Cinema.* Ed. Peter Lehman. New Brunswick: Rutgers, 1997.

Chellappa, Ramnath. "Intermediaries in Cloud-Computing: A New Computing Paradigm." *INFORMS* Dallas, October 2007: 26-29

Historical Notes About the Cost of Hard Drive Storage Space. 21 January 2008. http://www.alts.net/ns1625/winchest.html (Accessed 5 May 2009.)

Monsters. Inc. Docter, Pete and Silverman, David. Perf. John Goodman, Billy Crystal. Pixar Animation Studios, 2001.

Spiderman. Raimi, Sam. Perf. Toby Maguire, Willem Dafoe. Columbia Pictures Corporation, 2002.

Wall-E. Stanton, Andrew. Perf. Ben Burtt. Pixar Animation Studios, 2008.