Composite Media; A new paradigm for online media

Ingar M. Arntzen, Njål T. Borch, Northern Research Institute & Motion Corporation

Abstract

Media has been shaped by inherent limitations of the available distribution mechanisms since the advent of broadcasting. We seek to break free of this heritage by fundamentally reconstructing media. We promote the concept of motion as a fundamental building block in all media. By structuring and executing media according to shared motion, a world of opportunity opens up. In particular, our recent invention of highly scalable, cross-Internet motion-synchronization implies that motion-based media is collaborative and multi-device by design. We envision a shift away from the current paradigm, where fixed pieces of content are produced, transmitted and consumed. Instead, media will be composed again and again from a continuously developing corpus of online content and online motion. In this world, viewing, navigation, interaction and authoring can all be collaborative, multi-device activities. We call this Composite Media.

Introduction

Over the last decades broadcast media has grown much in complexity. The origins were simple, at least conceptually; TV providers used a one-way broadcast network to entertain a largely passive audience. These days viewers are increasingly interacting using a wide variety of devices, and live TV shows are coupled with web offerings, SMS services, Twitter streams, Facebook integration, infographics, live analytics, user feedback, custom apps, and more. The life cycle of a single TV program has also become more complex, as it is made available through broadcast networks, web-based live streaming options, and catch-up services for archived content. In addition, programs are time-shifted using personal video recorders (PVRs) and DVB receivers.

This increasing complexity is challenging. In particular, there are issues relating to coupling of media, services, devices and users. For example, Twitter messages and web offerings associated with live TV shows are coupled only as they coincide in real-time. This coupling is weak, and breaks as TV shows are time-shifted. Or, as TV shows are migrated from broadcast networks to archives, the lack of coupling forces viewers to manually navigate services, viewing devices, media applications and media offsets. On the other extreme, time-sensitive infographics merged into a streamed TV show couples these media types too strongly. As the TV show is time-shifted, one might want the infographics to reflect this, for instance by including the effects of other interacting VOD viewers. Instead, overlayed infographics are frozen in time.

We argue that such issues with coupling are pervasive and illustrate fundamental limitations of the current paradigm. Broadcast media is still conceptualized in terms of its simple origins; fixed pieces of media content are produced, transmitted and consumed. In particular, the coupling of media types, services, devices and user navigation has no prominent place in this model. To address these issues, we propose Composite Media (CM) as a new paradigm for online media.

Composite Media

Composite Media (CM) is online and multi-device by design. A wide range of networked devices, including smart phones, tablets, laptops and smart TVs may connect to a single, shared instance of CM. As this media presentation is requested to *play* by one device, this affects all devices in synchrony. If another device requests the presentation to *pause*, or perhaps to *skip* or *fast-forward*, all devices behave accordingly. Conceptually, instances of CM execute online so that client devices are free to connect and disconnect at any time. Client devices host independent views into the presentation, suitable for their capabilities and the role they play in the presentation. A single user can use multiple devices covering different aspects of the presentation. Collaboration is supported as multiple users access the same presentation.

For example, a Tour-de-France CM production may define multiple roles, such as “video”, “infographics” with race statistics and maps, and “interactivity” with status updates and comments. In addition there might be several roles for authoring, intended only for the production team. Viewers may then start up with all roles hosted by one device, say the smart TV. However, as a tablet joins the presentation, responsibility for “infographics” and “interactivity” is dynamically transferred from the TV to the tablet. If the race is time-shifted, or navigated for the highlights, all views and roles are still kept in synchrony. In short, CM is a flexible platform for “secondary screen” and beyond.

Composite Media is derived from two distinct ideas:

1. map media content to time - or other relevant media dimensions - and synthesize media presentation based on the current positions on media dimensions
2. shared navigation along media dimensions

1. CM as linear media

Real-time synthesis of media based on user navigation is the foundation of linear media. For example, video presentations are synthesized from video frames and subtitle tracks. Flash[2] animations are produced by mapping vector graphics to a navigable event-timeline. Similarly, slide show applications map individual slides to a dimension of discrete slide numbers. More recently, popcorn.js[4] demonstrates that arbitrary web content can be synthesized in accordance with the progression of HTML5 video elements.

CM takes this familiar idea to its logical conclusion. Linear navigation is recognized as common to all linear media, and promoted as a fundamental component of CM. **In CM, media navigation is universally represented by a single concept; unidimensional motion. All CM presentations are driven by motion.** More formally; CM presentations are synthesized in real-time from media content referring to media dimension(s). This synthesis is at all times directed by motion(s) along those media dimension(s).

To implement motion, CM relies on the technical concept Media State Vector (MSV) [1]. The

MSV is based on the classical equations of unidimensional motion under constant acceleration, where motions are described by initial conditions [p,v,a,t] (i.e. position, velocity, acceleration, time). The MSV is simplistic, yet expressive. Crucially, the MSV supports common navigational primitives of linear media, including discrete steps (next, prev, goto), continuous motion (play, fast-forward), or even acceleration. With MSV's as a fundament, CM readily supports all these navigational primitives. In addition, CM supports mixed navigation styles as well as multi-dimensional navigation.

A central idea in CM is the decoupling of motion from media, and the promotion of motion as a resource in its own right. This gives CM significant power and flexibility. The decoupling of motion implies that motions can be shared between content sources as well as presentational (e.g. visual) components. A single CM presentation can thus be synthesized from multiple, heterogeneous content sources, while allowing content subsystems to operate independently, yet in a coordinated manner. Similarly, UI components directed by shared motions can appear as tightly interconnected parts of a single media presentation, without requiring UI components to communicate, or even be aware of each other. In short, motion decouples content from presentation, and allows CM presentations to recombine them in flexible ways.

In contrast, consider how rich, linear web-based media presentations are currently made. Frameworks like Flash, SMIL, popcorn.js or Prezi [2,3,4,5] all build linear media around similar concepts of time-based user navigation. Yet all of them rest on implementations of motion that are internal and custom to the framework. This makes it hard to combine such presentations with external resources. The HTML5 video element makes this easier by exposing its internal motion through an API, thereby allowing the popcorn.js JavaScript library to synchronize arbitrary web content to the progression of a video. Still, popcorn.js misses a grand opportunity by insisting that the source of motion must be video (or audio). By making motion an object in its own right, CM makes it easy to include multiple videos in a single presentation. Perhaps more importantly, CM allows the construction of continuous media presentations, without requiring the inclusion of continuous media at all.

In effect, CM implies the possibility of a universal API for motion, to which presentational components and frameworks would interface. For instance, a motion-enabled video element would accept external motion, and operate as a slave to that motion. We have verified that this is possible with current HTML5 video elements, even without optimizing the internals for this purpose.

2. CM as Online Media

Above we indicated the value of sharing motions in the context of single-device media presentations. However, the defining characteristic of CM is that motions can be shared across the Internet. In CM, motions (MSVs) are online objects. This allows clients across the Internet, including regular web browsers, to connect and closely mirror the motions of ongoing media presentations. In fact, the accuracy of MSV synchronization [1] implies that shared motions may be thought of as simultaneous, even if clients are separated by large geographical distances, or communicating over different networks carriers. This is possible by compensating for network latency and clock skew. Network latency is only evident as motions are actively changed, for instance when play or pause commands are issued.

Equally important, online MSVs are extremely lightweight and can be hosted by highly scalable services. This implies that the scalability of CM is likely to be limited not by motion sharing, but by the scalability of its content services. CM presentations may make use of a large number of MSVs. Some of these MSVs may be used in large-scale broadcast scenarios, where asymmetric media control is required. For instance, in the classical secondary screen scenario a TV provider can own and control a broadcast MSV, whereas the audience is restricted to read-only access on their smart TV and secondary devices. Symmetric media control is relevant in smaller social groups, for instance in social collaborative applications. Additionally, CM presentations may allow individual users to create and use private MSVs, and even switch back and forth between private and shared MSVs. Private MSVs are a natural fit for on-demand media, and since private MSVs are still online they can easily be shared among the devices of a given user. Opening a private session for a friend is simply a matter of sharing access and url of the MSV.

The online nature of CM is clear as both motion and content are online resources. Conceptually, CM presentations execute online, independent of its clients. In particular, a CM presentation may continue to play even without connected clients. Clients may connect and disconnect at any time, without disturbing the execution of the CM presentation or other clients. Joining clients may quickly learn the current state of the presentation, as they synchronize with its MSVs and (subsequently) starts loading the relevant media content. Still, despite its online nature, the cost of synthesizing CM presentations is paid locally by clients.

In short, by making motion sharable across the Internet, all the advantages of motion sharing is extended to the distributed setting. In particular, shared motion provides the intermittent glue required for loosely coupling content types, services, devices, UI components, and users, across Internet. The name “Composite Media” similarly refers to how motion-based media presentations are composed seamlessly (in real time), from online content sources, UI components, users, and devices.

We have conceptualized and implemented shared, online motion, and verified its utility as technical foundation for CM. This work was recently published under its technical name; The Media State Vector (MSV) [1]. The Motion Corporation is the world's first commercial company to offer shared motion as a hosted service.

Applications

Composite Media is to a large extent a new model for online media rather than a replacement of technology. Though the model is fundamentally new, implementation can still build upon existing technology and services while simultaneously open for extensions and novel applications. The following section discusses a few selected topics relevant for broadcast media.

Seamless interaction between providers and consumers

TV broadcasters often provide content from multiple backends such as the broadcast network, live online streams (possibly with a live-window of several hours), and on demand offerings for archived material. While this improves outreach and functionality, it also adds complexity for the end user. The user must for example often manually select which backend to use. Paused live streams can not be resumed if the live window is surpassed. Instead the user must manually relocate the show within the archived service, as well as relocating the offset where the show was previously paused.

When a Composite Media presentation is created around the various data streams and backends, this complexity can be hidden as the user interacts with the same presentation at all times. By consulting an online representation of user navigation (the MSV) clients can at any time select the appropriate backend as well as the correct media offsets. This way, seamless transitions between different technologies can now be done by the client, ensuring a coherent and pleasing service regardless of backend and choice of device. It also allows the user to move the viewing experience from one device to another by simply opening the presentation on a different device. There is no need to pause the content, communicate locally between devices or even for the devices to know about each other. On the same note, collaborative viewing of content can be enabled by allowing users to invite other users to use the same MSV. No additional support is required.

Web-based Content Production

We argue that CM is a generic, flexible and cost effective approach to online media. In particular, by providing a generic synchronization mechanism for the web, CM immediately allows HTML5 to become the standard authoring framework for time-synchronized second screen applications, as well as multi-device media applications in general. As MSVs synchronize all visual elements (on a single page as well as on multiple devices), visual elements do not need to communicate with each other in order to act coherently. This allows sophisticated UI's to be assembled from a set of simpler, general-purpose, motion-enabled UI components. The MSV's inherent support for time-shifting also implies that live CM productions may be reused without modification for later on-demand viewing. Furthermore, authoring may be live, or come from a mixture of prearranged material and live cues. Authoring may also be collaborative, allowing responsibilities to be shared within a production team.

Transforming existing HTML based visualizations to support MSVs and thus be integrated in CM presentations will often lead to simplifying the components as opposed to adding complexity. For example, a content provider can create a control element for CM presentations with buttons for play, pause and navigation, as well as sharing or storing bookmarks etc, all visualized for various devices. Any presentation (even just a single video element) can have that controller added to it's page, thus allowing the user highly recognizable and unified controls for all services, on all devices. This control element does not need to be integrated directly with any media object as this coupling is mediated by the underlying MSV.

Multi-device Applications

Composite Media is multi-device by design as it depends on shared motion. CM presentations are built from a set of data sources and a set of visual or interactive elements connected to these data sources. The UI elements visualize and interact according to motion, and will as such present the correct data at the correct time for each user. This also means that different visual components can be used on different devices to provide true multi-device experiences.

While multi-device CM has obvious applications within secondary screen scenarios, it can also expand more traditional services. As Smart TVs are more easily available, infographics can be generated by the TV as an independent HTML overlay. This will enable graphics based on current data even if time shifted, or customized data for the user, for the user's location or even integrating the infographic with social media sites, e.g. showing comments from Facebook friends, playing a quiz with friends on the TV etc. The potential of using CM to deliver more targeted commercials is also easily apparent, opening for time-sensitive game style commercials as well as localized or personalized commercials.

Summary

In short, CM is a highly flexible and holistic paradigm for online media, where devices, users and content providers can work seamlessly together. It fits perfectly with the webification of media content, the increasing abundance of smart devices and the flexibility and social expectations of end users. We have validated the central technical concept, shared motion, and have implemented a hosted service, designed for scalability and ease of use. With this we have made proof-of-concept demos targeting a variety of multi-device applications, including broadcast, secondary device, online education, online newspapers, distributed music orchestration, collaborative slide-shows, collaborative viewing, collaborative authoring, collaborative documentation, as well as interactive visualization of scientific data. We argue that CM reduces the complexity of online media by providing a flexible and unifying model.

[1]: “The Media State Vector; A unifying concept for multi-device media navigation”, Ingar M. Arntzen, Njål T. Borch, Christopher P. Needham, MoVid’13, Proceedings of the 5th Workshop on Mobile Video, ACM, pages 61-66, Feb 2013, Oslo, Norway.
[2]: Adobe Flash. www.adobe.com/flashplatform/
[3]: SMIL. Synchronized Multimedia Integration Language, www.w3.org/TR/2008/REC-SMIL3-20081201/
[4]: popcorn.js. The HTML5 Media Framework, popcornjs.org
[5]: Prezi, prezi.com