Berlin Remix – A Computationally Generative “City Film” Artwork

Jim Bizzocchi
Simon Fraser University, CA
bizzocch@sfu.ca

Berlin Remix is a computationally generative artwork that creates and presents a series of short films re-edited and re-mixed from the archetypal “City Film” documentary Berlin: Symphony of a Great City (Ruttman 1927). The “City Film” is an historical documentary genre. These City films thrived in the late 20s and 30s, and continue to this day. Each of them documented the daily life of an individual city. The central film in the development of this genre is Ruttman’s Berlin. Berlin Remix is a generative artwork that re-mixes the individual shots of the original film into a series of shorter films. Each of these shorter films differs from the others in specific content, cinematic style, or both. Each of the derived films shows a single facet or theme contained within the larger film. The final form of the artwork will run completely autonomously, generating the series of shorter films in real time. The artwork as exhibited at CSDH/SCHN 2019 required a moderate amount of human intervention to join the system’s output segments into short films.

Keywords: video art; generative art; film aesthetics; documentary film; digital media

Le Berlin Remix est une œuvre d’art informatiquement générative qui crée et présente une série de films courts remodifiés et remixés venant du documentaire de l’archétype « film de ville » (City Film) qui s’appelle Berlin: Symphony of a Great City (Ruttman 1927). Le « film de ville » est un genre documentaire historique. Ces films de ville s’épanouissaient pendant la fin des années 1920 et 1930, et aujourd’hui encore. Le film central du développement de ce genre est celui de Ruttman, nommé Berlin. Le Berlin Remix est une œuvre d’art générative qui remixe les images individuelles du film original en série de films plus courts. Chacun de ces films plus courts varie l’un des autres selon le contexte spécifique, selon le style cinématographique, ou selon les deux à la fois. Chacun des films dérivés montre une facette ou un thème unique, confiné dans le film entier. La forme finale de cette œuvre d’art fonctionnera de façon autonome, générant la série de films plus courts en temps réel. L’œuvre d’art, comme exposé...
à CSDH/SCHN 2019, a exigé une quantité modérée d'intervention humaine afin de transformer la production de séquences du système en films courts.

Mots clés: art vidéo; art générative; esthétics; cinématographique, film documentaire; média numérique

Generative art

_Berlin Remix_ is an exercise in generative art. The practice of generative art has a long history, and is a wide-ranging approach towards the making of art. Galanter notes that it predates the computer, and claims it is “as old as art itself” (Galanter 2003). Most definitions maintain that generative art is created through a relatively autonomous system, typically “constructed through computer software algorithms, or similar mathematical or mechanical autonomous processes” (Botha 2009). Generative art manifests across a variety of forms and media: music, writing, visual arts, moving images, and networked computers. Generative artworks can be analysed across a range of dimensions: entities incorporated, processes used, environmental interactions (if any), and sensory outcomes (Dorin et al. 2012). Generative artists vary considerably in their relative emphasis towards either the final output or towards the generative process itself. They also differ in the degree of autonomy of the generative operations within their works: one can categorize generative works as either “closed” systems (all elements and processes are self-contained) or “open” systems (permitting input or interactions external to the work). Some closed generative works are “recombinant”: they rely on the shuffling and rearrangement of existing content (predetermined text lexia, images, video clips, sound clips), rather than the ongoing creation of new content, such as computational text generation or CGI (computer graphic imaging) creation. The current art works of the creative team rely upon the expressivity and ongoing output variation of closed and recombinant generative systems. Boden identifies three models for computational creativity: combination, exploration, and transformation (Boden 2009). Our systems incorporate aspects of all of these models.
Non-digital examples of generative works include a variety of dada and surrealist games such as *The Exquisite Corpse* (Brotchie and Gooding 1991) or Tristan Tzara’s *To Make a Dadaist Poem* (Tzara 1920). Burroughs continued this tradition with his “cut-ups” in both text and cinema (Burroughs and Gysin 1978; Balch 1966). The most extensive exploration of analog generative narrative is probably found in the Oulipo ("workshop of potential literature") group (Wardrip-Fruin and Montfort 2003). A number of digital generative works link explicitly to the earlier literary tradition. Jim Andrews and Noah Wardrip-Fruin explicitly recognize their own grounding in Burroughs’s cut-ups: *On Lionel Kearns* (Andrews 2004) and *Two Textual Instruments: Regime Change and News Reader* (Wardrip-Fruin et al. 2003). The French Alamo, LAIRE, and Transitoire Observable movements are the direct computational descendants of Oulipo (Bootz 2012).

Hayles claims that “Generative art … is currently one of the most innovative and robust categories of electronic literature” (Hayles 2007). Lev Manovich is the theoretician most closely associated with the concept of computationally-driven “database narrative”. (Manovich 2001). Manovich’s *Soft Cinema* video artwork uses database and generative computation to build an aesthetic of recombinant cinema (Manovich 2003). There are many contemporary examples of engaging recombinant generative artworks. Montfort’s *Taroko Gorge* allows Monfort (and others) to enter selected text phrases which then output as a series of haiku-like variations (Flores 2012). Harrell has built a series of systems based on shuffling text and image to create expressive and socially relevant artworks, including GRIOT, GENIE, Renku, MEMORY, REVERIE MACHINE, and Chimeria (Harrell 2007; Harrell and Chow 2009; Harrell et al. 2014). Rettberg and Coover have created a series of striking cinematic generative video artworks: *Three Rails Live* (with Montfort), the large screen installation Toxicity, and the video Cave installation Hearts and Minds (Rettberg 2016). Canadian examples of recombinant generative audio-visual art include Stan Douglas’s video installations *Suspiria* (2003) and *Klatsassin* (2006), Guy Maddin’s *Seances* (Guy Maddin, Evan Johnson, Galen Johnson and the National Film Board of Canada, 2016), or the
large-scale urban screen generative video *Salmon People* by Julie Andreyev and Simon Lysander Overstall on the role (and vulnerability) of salmon in British Columbia First Nations culture (Andreyev and Overstall 2016).

In all these examples, generative artists create systems. Their systems, with varying degrees of autonomy, create the artworks. Within the domain of generative art, there are many different approaches and artistic goals. This project’s overall objective is to create and refine an autonomous computationally generative system that will output a stream of short films assembled from a database of video shots.

**The “City Film” as generative model**

The “City Film” or “City Symphony” is an historically important film genre. It thrived in the late 1920s through the 1930s, represented by over fifty films, spanning four continents, produced by dozens of filmmakers (Hielscher 2015). The form continues to attract filmmakers in the present day. Uricchio argues that representative American films from the early phases of this genre provided a critical link between the evolving documentary form and avant-garde cinema (Uricchio 1995). *Man with a Movie Camera* (Vertov 1929) certainly did the same, and is probably the most critically-acclaimed work in the genre. The central film in the genre, *Berlin, Symphony of a Great City* (Ruttman 1927) has received a more mixed critical reception. Despite the film’s prominence, Grierson and Kracauer complained that it lacked social analysis, and other critics—Mayer and Verdone—found it cold and impassive (Uricchio 1982). Uricchio disagrees strongly. He not only finds the film brilliant in construction and execution, he believes it does have both a heart and a social critique. He feels that the film at the same time provides a visually rich and kaleidoscopic view of life in Berlin, but also models and reveals the limited perspective of the city’s bourgeoisie.

The City Film is an ideal genre for our generative art for a number of reasons. First, this form provides an opportunity to address more significant themes and issues than our current ambient video work. Second, these films are computationally tractable, and will provide us with a foundational framework for building our own generative documentary artworks. Uricchio identifies the use of “multivalent” content
clusters in the construction of Berlin, and maintains that the film is a “catalogue of techniques, structures, and iconography”. In a similar vein, Manovich claims that *Man with a Movie Camera* is an archetype for “database cinema” (Manovich 2001). We not only agree with them, we believe that this characteristic extends to the entire genre. City Films typically do present a catalogue of urban life: buildings, transportation, commerce, and recreation, to name a few of the higher-level content categories found in these films.Parsed within these categories are the images of a wide variety of people engaged in a cross-section of social and cultural activities: working, travelling, sports, entertainment, consuming, gathering, etc. Uricchio and Manovich point out that these films are semantically taxonomic in nature, suggesting that the collections of visuals are amenable to a variety of sequences, combinations, and re-combinations – which can in turn be adapted to a variety of aesthetic goals and purposes. The same is true of the sounds and music that can be associated with the visuals. The city can be translated into a database of audiovisual experiences, and then formed and reformed into any number of distinct yet semantically coherent sequences and statements. It is as rich a ground as one can imagine for the aesthetics of recombinant generative art.

The database nature of the City Films means that they are computationally tractable. These films do not follow the traditional documentary model based on the conventions of linear continuity editing. The City Films rely on a more thematic montage approach to cinematic construction. There is no imperative to use film’s classic narrative model as articulated by David Bordwell and Kristen Thompson: “a chain of events in cause-effect relationship occurring in time and space” (Bordwell and Thompson 2003, p. 69). This traditional narrative model works well with the more linear logic of cinematic continuity construction. *Berlin Remix* relies on the less linear, more additive logic of cinematic montage construction: the conjuncture of a series of shots with similar thematic implications. The City Films are sometimes called “City Symphonies” for a reason. Since the montage approach dominates the City Film genre, these films tend towards the lyrical rather than the strictly linear.
**Berlin Remix and the DadaProcessor**

*Berlin Remix* is based on our generative video sequencing system—the “DadaProcessor”. The DadaProcessor uses a set of rules and processes to sequence video clips into a video stream. These processes select video clips based on their individual content tags, and sequence them according the rules we have encoded. The DadaProcessor’s sequencing model is based on the principles of cinematic montage construction, utilizing an additive logic rather than the more constrained linear logic of cinematic continuity construction. Continuity editing carefully sequences individual shots in order to create the illusion of continuous and naturalist space and time. This is a powerful tool for the creation of realistic narratives by human filmmakers. However, it is too prescribed and linear to easily allow for algorithmic non-human decision-making. Montage filmmaking uses a looser logic of construction—the additive principle. Cinematic montages usually involve a sequence of individual shots that repeat common concepts or themes, rather than a prescribed linear sequence with seemingly continuous time, space, and action. The guiding principle for montage is repetition and addition, not temporal linearity. A typical montage sequence involves three or more shots with similar, but not identical content. For example, a shot of a bus, followed by a shot of a train, followed by a shot of an airplane implies the more generalized concept of “transportation”, bypassing the need for temporal and spatial continuity. This three-shot montage sequence was used in the 1937 version of *A Star is Born* (Wellman 1937). Montage editing can therefore be seen as a relatively simple additive process: the repetition of related shots in order to signify a broader concept or theme: “bus” + “train” + “airplane” = “transportation”.

This is a computationally tractable sequencing logic, but only if the computer can recognize the specific content of each shot in order to select and join similar shots into coherent sequences. This in turn requires that each shot is tagged for content. Our shot database has over one thousand shots from the original *Berlin* film, and we have classified each shot for its visual content. The tagging structure is relatively complex, based on a hierarchical metadata scheme. The scheme has two levels: higher level “thematic” tags, and more specific “detail” tags. Since they are

---

*Art. 12, page 6 of 18*

*Bizzocchi: Berlin Remix – A Computationally Generative “City Film” Artwork*
derived from the visual content of *Berlin*’s shots, the tagging structure reflects the film’s scope and focus. The current version has nine thematic tags: economy, workers, buildings, government, transportation, cultural/social, people, time of day, and animals. There are sixty detail tags nested under these thematic tags. For example, the “worker” theme has the following detail tags: construction workers, industrial workers, office workers, service workers, and domestic workers. Another example is the “economy” theme with the following detail tags: wealth, poverty, stores, signs (commercial), offices, industry, construction, shipping/boats, machinery. In our system, each shot will have one or more thematic tags plus one or more detail tags, depending on what actually appears in the shot. For some examples, see Figure 1.

The DadaProcessor will emit a series of films drawn from the original *Berlin* shots stored in the shots database. These output films will be short, but they each have their own sense of visual and semantic flow. The tagging system allows the DadaProcessor to select and join separate shots into coherent cinematic sequences based on content and theme. The DadaProcessor will use a variety of operational “templates” to choose and position shots within these sequences. First, each template selects shots based on its own pre-determined theme and detail content tags. Second, the template builds a film segment by putting the selected shots into an order. Third, the template will specify the screen timing for each shot in the segment. The shot selection and

**Figure 1**: Examples of Shot Tagging.
ordering decisions are randomized, but are constrained by the template's content tag instructions. The shot timing decisions are locked into the template, and will result in segments with predetermined pacing. The pacing can also incorporate segments that accelerate cutting speed in order to build viewer interest. A template typically consists of several segments with different specific content and pacing instructions. Each template operation will result in a short, finished film between one and three minutes long. The system will shuffle among the specific templates. Because of the differences in the templates, each film has its own content tags and pacing instructions. Further, each template includes randomized shot selection and ordering decisions within the tagging constraints. This means that each of the short films emitted by the system will differ from the others in content, cinematic style, or both.

The content differences for the system’s output of short films are driven by the tagging instructions associated with each template. In **Figures 2** and **3** are two sample system templates and the types of films they will produce. The “Day in the Life” template in **Figure 2** uses shots with the theme tag “Time of Day” and these four detail tags “morning”, “noon”, “afternoon”, “night.” The detail tags operate

**Figure 2:** “Day in The Life” Template.

**Figure 3:** “Trains – Walk – Bikes” Template.
one at a time in the appropriate temporal order (morning to night), creating four segments of city life unfolding as the day proceeds. Each of these segments has a specified length (either 15 or 20 seconds), and the system will randomly choose the appropriate number of shots with the correct detail tag in order to fill that segment. The result is a short film portraying a slice of a single day in Berlin. This short film is consistent with the full content of the complete film – which is indeed structured from morning to night. It is in effect a miniature that reflects (but does not replicate) the same theme from the larger film.

The “Trains-Walk-Bikes” template in Figure 3 uses the “Transportation” thematic tag, and the detail tags “pedestrian”, “bicycle”, and “train.” This template also incorporates cinematic style settings for “acceleration”. The template uses progressively shorter and shorter cuts, giving an increased sense of pacing and interest as this short film proceeds. This piece gives a sense of the importance of transportation in the daily life of the city – once again, a simple idea that is one of the key concepts embedded within the larger film. For each of the templates, selection of specific clips from the Berlin Remix database of shots is randomized. The “Trains-Walk-Bikes” template will pick approximately 10 specific shots from a field of 93 shots in the database with the “bicycle” tag. This randomized selection means that a single template can produce a number of films—each with specific shot selections that are roughly similar but not identical to the others.

Current state of Berlin Remix

The Berlin Remix DadaProcessor system can output finished films in a semi-autonomous fashion. The system interface has a number of input mechanisms to shape shot selection and sequencing (see Figure 4). These include content selection variables (higher-level thematic content tag selection and lower level detail content tag selection) and cinematic style variables (including decisions on pacing and timing, scale and motion selection, and choice of transition).

A set of these input decisions defines a template for the system to produce output films. Currently, these variables are dependent on artist decision. This level of
functionality works well as a proof-of-concept of the system’s operational capabilities, but we need to go beyond this. The fully autonomous version of the system will have a number of these interface decisions encoded into discrete self-contained modules, which we see as templates. We are first building the system’s capacity for fully-autonomous template operation within this model. We will then incorporate the ability to select and implement from a set of these templates. The final artwork will automatically select a template, create and present a short film created by the template’s operation, and then select and implement the next template. The result will be a fully autonomous artwork that presents a series of short films drawn from the database of original Berlin shots. The output will be varied, due in part to the differences between the templates, and in part to the randomized shot selection and pacing instructions built into the operation of each template. See Figure 5 for a flowchart of the system’s design and operations.

The system will run continuously, emitting an ongoing series of short (2–4 minute) films in real time. Each short film will be different from the others in content, cinematic style, or both. We also believe that the system will be
stable. We have had considerable success exhibiting fully autonomous ambient video art in real time using a simpler version of the DadaProcessor. The system is built on the Maxx programming environment, and has proved to be reliable in its earlier versions – running for days and weeks at a time without breaks or problems.

The *Berlin Remix* experience

Ruttman’s original *Berlin: Symphony of a Great City* portrays a number of facets of Berlin in the mid 1920s. The original film starts in the morning on a train entering Berlin from outside. The film continues within the city itself, using time-of-day as an overarching organizing sequence. We see Berliners travelling to their jobs, working during the morning, breaking for lunch, working in the afternoon, eating dinner, partaking in a variety of recreational activities later in the day, and enjoying Berlin’s active evening night life at the end of the day. The participants cut across all segments of Berlin society: rich and poor, workers and clients, men and women, children, and both working and pet animals. It is a kaleidoscopic sampler of class, occupation, gender, and activity—reflecting the complex realities of a city embracing modernity and its contradictions.
The *Berlin Remix* experience reflects the core concepts of the original film. This is due in part to the duplication of content. All of the shots in Ruttman’s *Berlin* have been included in the *Berlin Remix* database. The relationship to the original film is further preserved due to the nature of the tagging structure, and the way the tags drive the individual templates. The specific tagging decisions involve a form of interpretation of the meaning of shots at a fairly basic level. Since these decisions are based on the direct visual content of the shot, the interpretive nature of the tagging is relatively grounded in Ruttman’s intended content.

The construction of the templates does involve more space for the intervention of the *Berlin Remix* authorial team. The templates contain and mix specific content groups from the original film. This content mixing is determined in large part by the higher-level thematic tags and lower-level detail tags embedded within each template. Meaning is constrained and channeled through the tags contained within the template, and instantiated in the specific shots selected by the system. Each of the system’s emitted short films is therefore an interpretation of one or more aspects of Ruttman’s work. In fact, it is possible to identify three different levels of creation within each short film emitted by the system: Ruttman’s original shot creation, the author’s template creation, and the system’s randomized selection processes operating within the template’s operations.

Of course, the authorial intervention of the system’s creators can move beyond Ruttman’s original intentions. We claimed earlier that the logic of cinematic montage was a relatively simple additive process, and that the selection and re-sequencing of Ruttman’s shots would therefore remain generally true to Ruttman’s intentions. However, the actual extent of that claim can be compromised by the power of cinema’s poetics. Eisenstein reminds us that the meaning of cinematic montage sequences can go beyond the meanings of the individual shots. “…a mouth + a child = ‘to scream’, a mouth + a bird = ‘to sing’, a knife + a heart = ‘sorrow’” (Eisenstein 1949). In the same way, we can create individual templates that transcend, subvert, or even contradict what seems to be Ruttman’s original intentions. For example, we have tags associated with gender (People – male, and People – female). One could
use these in combination with other tags to create a template that contrasts images of male privilege (male plus wealth, male plus recreation) with images of female suppression (female plus domestic worker, female plus service worker, female plus office worker). We haven't implemented such a template yet, but we will be working on this in the future. It is my intention that the output templates will generally be true to my own sense of Ruttman's work, but I am also interested in increasing the variety of output through templates that deviate from his norms.

**Design of generative art**

I wish to make a final point about the design process that built the DadaProcessor system and will lead to the set of Berlin Remix template modules. The design of this artwork explores an ongoing dialectic between artistic control and system variability.

```
artistic control <=> system variability
```

I am not a computer scientist, I am a video artist. My real goal is viewer experience, not system design. For me, the success of the Berlin Remix experience depends on the presentation of an ongoing series of short films that engage the viewer. Each film should offer aesthetic pleasure, intellectual interest, or both. At the same time, the set of films emitted should differ significantly, showing a range of content themes and cinematic treatments. This represents a difficult dual challenge for the design of a computational generative system. Any measure of artistic control limits the level of ongoing variation, so the entire process involves a balance between these two imperatives. At its most basic level, the inclusion of random processes within the templates is necessary for output variation, but any random decision making decreases artistic control.

It is possible to minimize this contradiction to some degree, and we are working on ways to do this. The basic method is simple, albeit time-consuming. As we increase the number and variety of effective templates, we will increase the level of variation in the system output without decreasing the quality the experience. We currently have models for five templates that we believe are effective and interesting. We are building more, and we believe that a set of twenty good templates will provide a reasonable level of output variation. Another strategy we can implement is to
increase the number of clips in the system’s shot database. For a future artwork, we will add the shots from Vertov’s *Man with a Movie Camera*, to the existing shots from *Berlin: Symphony of a Great City*. This new artwork will require some modifications to the existing tagging system. However, the films have significant thematic and content similarities, so this tagging modification will be minimal. The number of shots in the database will double, so the subsequent detailed output variation will be increased significantly.

In any case, the development of a series of effective templates is the heart of the dialectic between system variation and artistic control. Each template instantiates some of the creative decisions a human editor would make – such as a broad selection of shot content and sequencing, or determination of shot timing and pacing. However, our selection of specific shot content is not as controlled as a human editing process. Our templates select shot categories, not specific shots. Our design goal is that the cumulative effect of our system’s shot selections will have enough semantic and visual coherence to provide a sense of cinematic flow and thematic development.

My benchmark for artistic success is the ability of our system to approach the quality of human creators. We do not expect the system to replace or surpass the output of talented human artists. We do expect a consistent and reasonable level of artistic competence and associated audience pleasure in the system’s generative performance. To accomplish this, our challenge has been to “encode practice” - to identify the poetics of effective artistic creation and instantiate a version of these poetics in computational code.

**Acknowledgements**

The DadaProcessor design and the *Berlin Remix* artwork are my conception, but credit is also due to an excellent production team. Justine Bizzocchi is the Production Manager and Technical Director for this project. Casper Leerink and Paul Paroczai worked on the detailed software programming under her direction. Brandon Hoare did the shot breakdown and tagging, and edited a series of initial prototype films to help guide the design process. Arne Eigenfeldt provided input on approaches to
the system design. The concepts that drive this project were developed and refined with William Uricchio at MIT’s Open Documentary. This project is funded by the Social Science and Humanities Research Council of Canada. Additional funding and support have been provided by Simon Fraser University’s School of Interactive Arts and Technology, and by the University’s Faculty of Communications Art and Technology.

**Competing interests**

The author has no competing interests to declare.

**Editorial contributions**

Guest editor for DSCN Congress 2019 Issue: Barbara Bordalejo, University of Saskatchewan, Canada.

Section editor: Nathir Haimoun, The Journal Incubator, University of Lethbridge, Canada.

Copy editor: Shahina Parvin, The Journal Incubator, University of Lethbridge, Canada.

**References**


Harrell, Fox, Dominic Kao, Chong-U Lim, Jason Lipshin, Ainsley Sutherland, Julia Makivic, and Danielle Olsen. 2014. “Authoring Conversational Narratives in Games with the Chimeria Platform.” In *Proceedings of the 9th International*


*Two Textual Instruments: Regime Change and News Reader.* Accessed July 28,  

International Pictures. Motion Picture.