HAL
open science

Capturing and Indexing Rehearsals: The Design and Usage of a Digital Archive of Performing Arts

Rémi Ronfard, Benoit Encelle, Nicolas Sauret, Pierre-Antoine Champin, Thomas Steiner, Vineet Gandhi, Cyrille Migniot, Florent Thiery

To cite this version:

Rémi Ronfard, Benoit Encelle, Nicolas Sauret, Pierre-Antoine Champin, Thomas Steiner, et al.. Capturing and Indexing Rehearsals: The Design and Usage of a Digital Archive of Performing Arts. Digital Heritage, Sep 2015, Grenade, Spain. pp.533-540, 10.1109/DigitalHeritage.2015.7419570 . hal-01178053

HAL Id: hal-01178053

https://inria.hal.science/hal-01178053

Submitted on 22 Dec 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Capturing and Indexing Rehearsals: The Design and Usage of a Digital Archive of Performing Arts

Rémi Ronfard,* Benoit Encelle,† Nicolas Sauret,‡ P.-A. Champin,†
Thomas Steiner,† Vineet Gandhi,* Cyrille Migniot,§ Florent Thiery¶
* INRIA, LJK, University of Grenoble, France.
† CNRS, Université de Lyon. LIRIS, UMR5205. Université Lyon 1, France.
‡ IRI, Paris, France.
§ University of Burgundy, France.
¶ UBICAST, Paris, France.

***Abstract*—Preserving the cultural heritage of the performing arts raises difficult and sensitive issues, as each performance is unique by nature and the juxtaposition between the performers and the audience cannot be easily recorded. In this paper, we report on an experimental research project to preserve another aspect of the performing arts—the history of their rehearsals. We have specifically designed non-intrusive video recording and on-site documentation techniques to make this process transparent to the creative crew, and have developed a complete workflow to publish the recorded video data and their corresponding metadata online as Open Data using state-of-the-art audio and video processing to maximize non-linear navigation and hypervideo linking. The resulting open archive is made publicly available to researchers and amateurs alike and offers a unique account of the inner workings of the worlds of theater and opera.**

***Keywords*—Digital archiving, hypervideo, video processing, audio processing, Linked Data, performing arts, theater, opera**

I. Introduction

This paper presents the results of a two-year research project dedicated to the creation of a digital archive holding the complete rehearsals of two recent productions of a theater play and an opera. To the best of our knowledge, this is the first time that exhaustive video recordings of theater and opera rehearsals are made available to researchers in the digital humanities at such a large scale. Quoting from a 2007 DVD release of theater rehearsals [1], we acknowledge that *"within the world of theatre the rehearsal room is a sacred space—the private domain where boundaries are pushed, risks are taken, mistakes made, vulnerabilities exposed and, at its very best, magic created. It is not a place into which the public is often, if ever, invited."* In our case, the directors of the theater and opera productions were a leading force to open this sacred space—digitally. This achievement was made possible by introducing a specifically designed, non-intrusive workflow for a natively digital archive from capture to editorialization, including collaborative annotation, audio and video processing, data modeling, and hyperlinking, which we describe in the present paper.

Archiving video recordings of live performances is important for preserving the cultural heritage of the performing arts but it is not yet a common or systematic practice, and theater remains known as *the ephemeral art* [2], [3]. This is likely to change in the near future, as high-quality video recording equipements are becoming more widely available to theaters. Many organizations around the world now actively produce and archive video recordings of theater performances. Recently, such efforts have been extended to the World Wide Web. The French National Institute of Audiovisual[1] (INA) allows viewers to download a large number of television broadcasts of historical theater performances. France Television has been experimenting with an augmented theater experience for the Web, showing five different versions of the same play at different stages of the production (including rehearsals) and presented in different styles [2].

In most instances, performances are recorded with a multi-camera, multi-take approach derived from broadcast television common practices. Having multiple viewpoints is especially beneficial for studying actor performances. One extreme example is John Gielgud's *Hamlet* starring Richard Burton in 1964, which used 19 cameras. One problem with the "multiple camera, multiple takes" technique used by television broadcasters is that it is costly and time-consuming to produce; and it generates huge volumes of data, which are difficult to organize into an accessible and usable archive. Or more frequently, the rushes are just ignored and only the edited version is kept as an archive.

In this work, we take a radically different approach, where we record the entire rehearsals from the same viewpoint. As early as 1966, Jean-Luc Godard remarked: *"Why do theater people never film their performances to keep them as an archive? This would be very simple: put the camera in the middle of the orchestra seats with a medium lens—not a zoom lens—because it would already be making choices and propose an interpretation"* [4]. Accordingly, our approach in this project has been to record the rehearsals from the director's viewpoint with the highest possible video resolution, and to augment them with detailed, high-level annotations that make it possible to quickly search, browse, and even re-edit the archive interactively. What we lose is the variety of viewpoints, which is typical for television broadcasts. What we gain is the ability to compare different versions, measure progress, and understand the rehearsal process by the virtue of having a single spatial reference. We also experimented with state-of-the-art computer vision techniques for varying the framings and shot sizes interactively by making use of a *virtual* zoom

[1] French National Institute of Audiovisual: http://www.ina.fr/
[2] http://nouvelles-ecritures.francetv.fr/theater-sans-animaux/

Figure 1: Three successive steps in the theater rehearsal process for *Cat on a hot tin roof*: (a) Table readings; (b) Early rehearsals in a studio; (c) Late rehearsals on stage (including dress rehearsals and technology rehearsals).

lens [5], [6], which makes it possible for researchers and amateurs alike to use the archive as *stock footage*, which they can re-frame and re-edit according to their own interpretation (this is described in Section V).

As a case study, we are releasing an Open Data archive of the complete rehearsals of two performances played in 2013: *(i)* a French translation and adaptation of *Cat on a hot tin roof* written by Tennessee Williams and directed by Claudia Stavisky, *(ii)* the baroque opera *Elena*, composed by Francesco Cavalli, directed by Jean-Yves Ruf and conducted by Leonardo García Alarcón. Details can be found on the Spectacle-en-ligne(s) web site [3].

The paper is organized as follows. Section II reviews previous work in using video as an archive for theater. Section III describes the context of theater rehearsals and introduces the key concepts that govern the archive. Section IV describes the workflow that was created to produce, annotate, and publish the archive. Section V describes usage scenarios and applications that were developed to support access to the archive by different audiences. Section VI concludes with limitations and future work.

II. State of the art

Audio and video recordings are useful resources in the field of performance studies [7], [8] and rehearsal studies [9], [10]. Yet, in practice, researchers have limited access to performances and must often refer to their own memories of the production or the rehearsals. Exceptions include previous work by theatre director Daniel Mesguich [3] and a recent production of *Hedda Gabler* with Cate Blanchett in the title role [1]. Even in those exceptional cases, only an edited version (an interpretation) of the rehearsals is made available to the public.

Audio and video processing: A useful feature for a large video archive is the capability to retrieve video segments based on which actors are performing and which segment of the play script these actors are performing. In order to relieve annotators from the burden of precisely annotating which actors are present in which frames and speaking (or singing) which lines of the play script (or score), we experimented with automatic audio and video processing methods. The topics of actor and speaker recognition are very active in the audio and video processing research communities, but little work has been dedicated to the special case of live performances. The stage is a complex and cluttered environment with complex lighting and vast dimensions, which makes automatic audio and video processing quite challenging.

One approach to indexing multiple theater rehearsals and performances is to automatically align all of them to the same play script. The *script-sync* feature in the Avid Media Composer editing suite is a well-known example of an off-the shelf solution for the alignment of video recordings with their scripts word-by-word. Such tools are based on Hidden Markov Models (HMM) and work by computing a forced alignment between the play script and the audio recording. Nevertheless they require very high quality sound recording and fail in the presence of music, reverberation and background noise. This makes them unsuitable for our case. Speaker change detection (also called speaker diarization) is an alternative method which provides line-by-line rather than word-by-word alignments and has been demonstrated on television shows [11] and theater recordings [12].

Hypervideo and interactive documentaries: The term *hypervideo* is commonly used to refer to *"a displayed video stream that contains embedded user-clickable anchors"* and annotations [13], allowing for navigation between the video and other hypermedia elements. In standard literature, hypervideo is considered a logical consequence of the related concept of *hypertext* [14], In contrast to hypertext, hypervideo necessarily includes a time component, as content changes over time. Therefore hypervideo has specific technical and aesthetic requirements, such as the need for video segmentation into scenes and objects. The opportunities for feature-rich semantic hypervideos are endless, only limited by feasibility and ease of creation.

III. The rehearsal process

Live performance rehearsals involve many actors and participants, not limited to theater actors, including the director and directing assistants, singers and musicians, lighting directors, sound directors, stage managers, technicians, *etc.* The archive and its metadata were designed to take all of them into account, based on a careful analysis of the experimental fields of theater and opera. Theater and opera turned out to be distinct experimental fields with important differences in the rehearsal process. The two experiments in our case were original productions of Tennessee Williams' *Cat on a hot tin roof* and Francesco Cavalli's *Elena*. Both of them were rehearsed during spring and summer of 2013, in a period

[3] http://spectacleenlignes.fr

of eight and six weeks respectively. The project was about documenting the entire rehearsal process, from the first *table readings*, where actors read the script together for the first time, all the way to the *opening night* (see Fig.1). However, the nature of the creative work diverges considerably from the *mise en scene* of a contemporary play to an ancient and obscure baroque opera, musically updated with a consequent adaptation work. Those aspects motivated us to design a generic system that could be deployed in such different contexts.

A. The rehearsal process

A performance rehearsal is a closed-door privileged moment where actors and director work in a protective intimacy that allows them a total commitment to their art. The idea of filming those moments of intimacy, moreover exhaustively, could cause discomfort at best, or even rejection at worst, from professionals considering as sacred the privacy of rehearsal. This aspect was taken into account in the design of the system, which was engineered to be minimally intrusive and maximally respectful of the creative process at work during the rehearsals.

Figure 2: Rehearsal of the opera *Elena*, Festival d'Aix-en-Provence, 2013. © by Pascal Victor / Artcomart.

B. Performances and discussions

The rehearsals unfold according to a relatively classic *modus operandi*: the acting and directing crew gathered daily in a studio room or on stage to work through certain parts of the script or of the partition. Based on observations of previous rehearsals by the same director, we were able to identify a recurring pattern in the rehearsal process: an alternation of performances during which actors play their roles, and discussions during which the actors and the director interrupt the performance and make comments and suggestions. This alternation of performance and discussion became a central element of the data model used to annotate the archive. We reckon that a director working according to a different methodology would have generated a very different corpus and annotation set.

IV. Archive construction and publication workflow

Digital heritage projects often focus on digitzation of analog materials and data, when others would focus on the usage and the valorization of digital resources. In our case, a specific feature of the project was the design decision to create a natively digital archive from scratch, taking into account its potential usages by different audiences. We therefore built a complete pipeline from the design and the making of the archive to its publication and usage scenarios. This section describes the details of the workflow and describes all the hardware and software components that enabled the construction and the publication of the archive.

A. Video recording

In the context described previously, the video recording required a non-intrusive and easy setup for non-technical operators. Actually, their profile must belong to dramaturgy and performing art in order to dedicate their attention to performance and directing description. Therefore, the apparatus was designed to be fully *plug & play* with a single record button to start a capture session. The collected corpus embraces the integrality of the rehearsals, from the "text walk-through" at the table to the final dress rehearsal, with about five to seven working hours a day, which in consequence result in likewise five to seven hours of recorded and annotated video every day.

The unit for video recording and annotating is typically composed of a PC laptop that is connected to a video camera mounted on a tripod and remotely controlled from the PC, of an omnidirectional microphone as input of a small sound mixing desk, in turn linked to the laptop. This system can be easily set up and wrapped up, is movable, and can fit in a studio room without burdens. Running on a dedicated Linux distribution, the unit executes and displays only one software: an interface through which the operator can *(i)* remotely control the camera and *(ii)* take notes synchronously. The technical specifications of the original video files are the format `.mkv`, the codec H264, and the resolution of 1920×1080 pixels (full HD).

B. Annotation

Simultaneously to the live video capture, the operator was asked to describe the rehearsals in his own words using the note-taking tool. The data produced during this annotation process were synchronized with the video stream and encoded into an XML format based on the Cinélab data model.[4]

Annotations of the rehearsals were further classified into two main categories: descriptions and interpretations. Description annotations contain contextual and objective information about the rehearsals (performances and discussions, with or without costumes, time, location, *etc.*), whereas interpretation annotations targeted the dramaturgy and aesthetics of the performances themselves—with free observations and comments on the creative work. The interpretations were further classified into sub-categories. Those sub-categories (see tables II and III) emerged progressively along the rehearsal process according to the point of view and the focus of the operator. In contrast, different categories emerged for theater and opera, reflecting both the different nature of these performing arts, and the personal point of view of each annotator.

The result of this synchronous annotation was an augmented video stream with semantic metadata, also considered

[4]Cinélab Data Model: http://liris.cnrs.fr/advene/cinelab/

	Cat on a Hot Tin Roof	Elena	Total
Number of videos	70	100	170
Total duration (h)	242	177	419
Total volume (TB)	2.47	1.60	4.07
Annotations	3,530	6,968	10,498

Table I: Size of the video corpus

as intra-video indices, making the video corpus easily searchable with a proper search engine.

1) Data model: Figure 3 describes the main elements of the data model used by the note-taking tool. This model results from a workshop analysis, gathering several kinds of attendees (researchers, producers, dramatists, film editors)—each kind having different requirements concerning rehearsals annotation capabilities. It focuses on annotations that have to be created in realtime (*i.e.*, during the rehearsal).

A session annotation is linked with a video file and represents a rehearsal session (typically lasting a half or full day). A session annotation is connected to a place (where the rehearsal takes place) and to a creation (linked to a work and castings). A session annotation is made up of chapters that represent performance or discussion times. A chapter contains contextual metadata (*i.e.*, with or without costumes, lights, sets and musical inserts) and includes a direct reference to the section of the work being performed or discussed (for theater, we used page and line numbers in the play-script; for opera, we used page and bar numbers in the musical score).

2) Practical usage of the model: Generally speaking, at the beginning of a rehearsals caption process for a given work, an operator first enters some time-stable data in the note-taking tool: the creation, work, actors, characters, rehearsals place. Additional data can next be entered or updated during a rehearsal session—*e.g.*, contextual data (costumes, sets, *etc.*) about a given chapter (inherited from a previous one, if any).

3) Annotation categories: The annotation process revealed a rich set of categories associated to interpretation annotations, reflecting the specific domains of theater and opera and the personal viewpoint of the annotator. Tables II and III present the richness of these categories and the potential of the archive for further research on genetic analysis of creative works.

C. Data cleansing and conforming

After the rehearsals, we centralized the data on a server for video encoding and annotation metadata conformation. Due to the tight agenda of the project, the annotation software did not implement a strict certification on data input. Therefore, as a preliminary step before the ingestion of the corpus into the applicative platform, manual cleansing and conforming were executed on both the data (checking audio and video quality) and the metadata (merging similar categories, harmonizing the text reference format, correcting misspellings).

D. Audio and video processing

As an optional step, we trained statistical models of the actors' voices and appearances to detect and recognize them

Category	Occurrences
alternative performance proposition	70
conflict	7
decision	14
detecting an issue	81
doubts	23
improvisation	16
intervention of the director	554
intervention of the first assistant	3
intervention of the technical staff	22
opinion of an actor	26
performance of an actor	331
questioning	4
reference to real world	3
rehearsal with text	1
repeating a scene	451
request from an actor	62
run-through	42
script adaptation	6
space issue	6

Table II: Annotation categories and number of occurrences for the theater play *Chatte sur un toit brûlant*

Category	Occurrences
accessories	292
address	247
articulation	357
breathing	25
character definition	725
characters relation	820
conflict	27
cooperation director/conductor	178
costumes	711
decision	154
decoration	14
emotion	134
gestures	346
improvisation	52
instrumentation	118
intonation	65
lights	221
moving	830
music dynamics	158
partition modification	166
phrasing	131
power relation	47
pronunciation	109
reference to camera	1
research	327
scenery setting	85
seduction relation	66
tempo	406
voice "color"	112
working on the text	104

Table III: Annotation categories and number of occurrences for the opera *Elena*

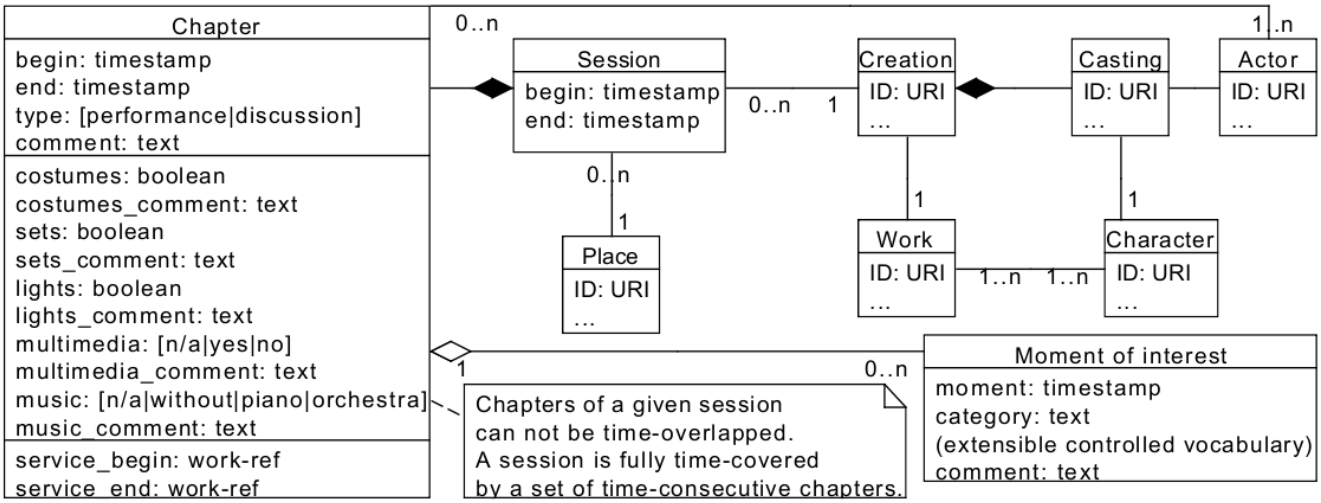

Figure 3: Main elements of the data model used by the note-taking tool

in the audio and video modalities. As explained later, this offers additional searching, browsing, and editing capabilities to our system, at the expense of (widely) increased processing time.

1) Learning actor voices and appearances: We used the ALIZE speech recognition system [15] for learning statistical models of each actor's voice from one example per actor. Each example is typically a ten-second extended speech line recorded during the first reading of the play. In parallel, we trained statistical models of each actor's visual appearance, using example images from at least eight views [16]. Examples are typically extracted from the first dress rehearsal, as the appearance models are dependent on costumes. Those models can then be used to automatically align each performance with the play script and to localize actors on stage and on screen in all subsequent dress and tech rehearsals.

2) Alignment between play script and performances: Using the ALIZE speech recognition toolkit, we built a play-script following algorithm where rehearsals are assumed to be generated by a probabilistic generative model. More specifically, each line in the play script is a state in a Semi-Markov Model (SMM) with a duration probability and a speaker voice probability. In our experiments, this methods outperformed state-of-the-art speaker diarization methods based on HMM, which do not have duration models.

3) Actor detection and recognition: In some selected scenes, we detected all actors present on stage in parallel based on their appearances using a method described elsewhere [16]. Despite the challenging environment, this method provided very good results for all actors in all tested sequences, where generic state-of-the-art tracking methods quickly drifted away from the target and were unstable both in presence of occlusion and fast movements. One drawback with the method is its computational cost, which made it impossible to perform actor detection and recognition on the entire archive. Selected scenes were chosen for evaluation and validation purposes.

E. Data publishing

The project built on the applicative platform *Ligne de temps*[5] to implement specific modules dedicated to fine-indexed, annotated video archives of performance rehearsals. *Ligne de temps* is originally a software designed for video annotation and indexation, inspired by the timelines usually used in video editing softwares, and was conceived for advanced film critique or pedagogy. It evolved into a Web platform with Web player and front-end annotations modules. For the purpose of the project, *Ligne de temps* was adapted to embed and display annotated video of performance rehearsals, taking into consideration the particular metadata and their potential usages.

1) Ingest: The ingest operation was made possible due to the previous step of data cleansing and conforming. It involved automatic referencing of the video files as media contents into the platform and the synchronization of the metadata (annotations) to the media contents. *Ligne de temps* provides an applicative environment to search, consult, edit, and publish the media contents and their metadata, therefore bringing the rehearsal annotation data from static to dynamic. However, theater and opera archives were separated in two distinct access points, due to usage scenarios and to differences in data format.

2) Search engine: A faceted search engine was designed to perform complex requests on the archive based on available metadata. Figure 4 displays the search interface for theater, divided vertically in two parts for different types of search operations: *(i)* left side: searching chapters by rehearsal context data such as date of rehearsal, character(s) on stage, lines of the script, type of chapters, *etc.*, providing a sortable list of chapters, *(ii)* right side: searching interpretation annotations by full text and/or by a closed list of categories, providing a sortable list of annotations. Thanks to the synchronization of metadata to the video stream, a search result list from either side can be used as a request to the opposite side. For instance, from the result list of annotations (right side), it is possible to extract all chapters with corresponding timecode (left side).

[5]Ligne de temps: http://www.iri.centrepompidou.fr/outils/lignes-de-temps/

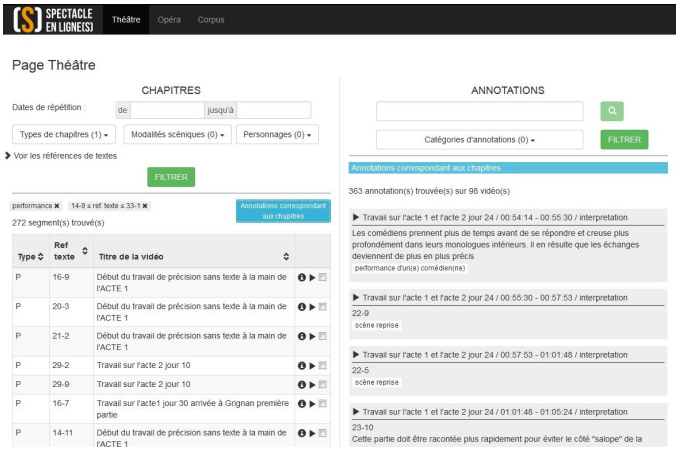

Figure 4: Multi-faceted search engine giving access to all archive materials. Returned video segments can be watched and navigated as hypervideos.

F. Linked Open Data compatibility

In order to foster and ease multiple uses of the corpus, it was important to make the collected information as accessible as possible. For this purpose, we chose to comply with the principles of Linked Open Data [17]. This implied that every piece of information was both identified by and accessible at a given URI, and that it was published using the RDF data model [18].

The Cinélab data model maps nicely into RDF, thanks to additional standards such as Media Fragments URIs [19] that are used to precisely anchor annotations to video temporal fragments. Furthermore, Linked Data has the common query language SPARQL [20], so building new services on top of our corpus can be achieved with relative ease by reusing standard components. Finally, we used an approach called Linked Data Fragments, proposed by Verborgh *et al.* [21], to make the data widely available while preserving scalability. Technical details can be found in a related publication [22].

But the benefit of Linked Data does not only lie in standardization. As its name implies, the key feature of this approach is to allow linking data across different sources on the Web, in order to combine it into otherwise inaccessible information. In the context of Digital Heritage, this notion is central, as data is often scattered in multiple institutions. While Linked Data is already widely used for museums or libraries,[6] its use with video has emerged more recently [23], [24], and as far as we can tell, has not yet been widely considered for live performance archives.

V. Usage scenarios

In parallel to the design and implementation of the previously described workflow, the project anticipated several usage scenarios of the archive, taking into consideration the potential needs of future users. A particular effort was made to address specific audiences with pedagogical scenarios, culture mediation scenarios, but also research and amateur scenarios. Some of these scenarios were implemented into demonstrators, others into mock-ups, allowing the partners of the project to freely explore prospective scenarios according to their interests. This prospective work was indeed central in the design of the continuous workflow in order to articulate the newly constructed archive with existing or prospective usages. This approach is one of the particularities of this project regarding researches in the digital heritage, *i.e.*, to have considered the archive as a living cultural object, even though the archive did not exist yet.

A. Quantitative and qualitative analysis

The massive aspect of the archive makes it difficult to apprehend. Both the nature of the archive, *i.e.*, video as a temporal object [25] and its total duration (419 hours for 170 video files) as well as the amount of annotations (10,498) demanded a comprehensive and cognitively acceptable representation. Taking the project's title literally [7], the archive can be viewed as multiple lines of rehearsals. We imagined a matrix where the columns are pages in the play-script (or the musical score) and the rows are days of rehearsal. The result is the distribution of the rehearsals according to the script or score over time.

B. Web components

According to our usage scenarios analysis, several requirements were identified concerning the archive access, archive visualization, and browsing capabilities. In order to make the rehearsal archive accessible to the broadest possible audience, we embraced the Web platform and generated online accessible demonstrators, enabled through native HTML5 video support in all major Web browsers. Concerning archive visualization and browsing capabilities, the multiplicity of usage scenarios highlighted the fact that end users must interact with different kinds of media or content (video captions, scripts or scores, and annotations) in different ways. As a result, components for displaying and/or interacting with these different kinds of content were firstly implemented and can be combined with each other to form adequate hypervideos that maximize the end-user's experience. Moreover, this hypervideo design approach—based on components combinations—fosters novel usages of the archive, as new components can be added and combined to older ones in order to fulfill new needs.

C. Fine-grained visualization and browsing

For some selected scenes, we offer a fine-grained visualization and browsing interface allowing to zoom both temporally and spatially through the archive by making use of the audio and video processing described in Section IV. Using audio-to-script alignments, we offer options to navigate a scene one line at a time and one rehearsal at a time. More precisely, hyperlinks are offered to move to the previous and next lines in any given rehearsal; or to move to the previous and next rehearsals in any given line; thus providing a "temporal zoom lens" to the archive. Because the source material was recorded from a distance and with a fixed lens, the archive presents a general view of the stage as a default framing. This has the advantage of never hiding anything from the audience. But

[6] See, *e.g.*, http://data.europeana.eu/ or http://data.bnf.fr/

[7] Spectacle-en-ligne(s) can be translated loosely to "online performances" or "lines of performance"

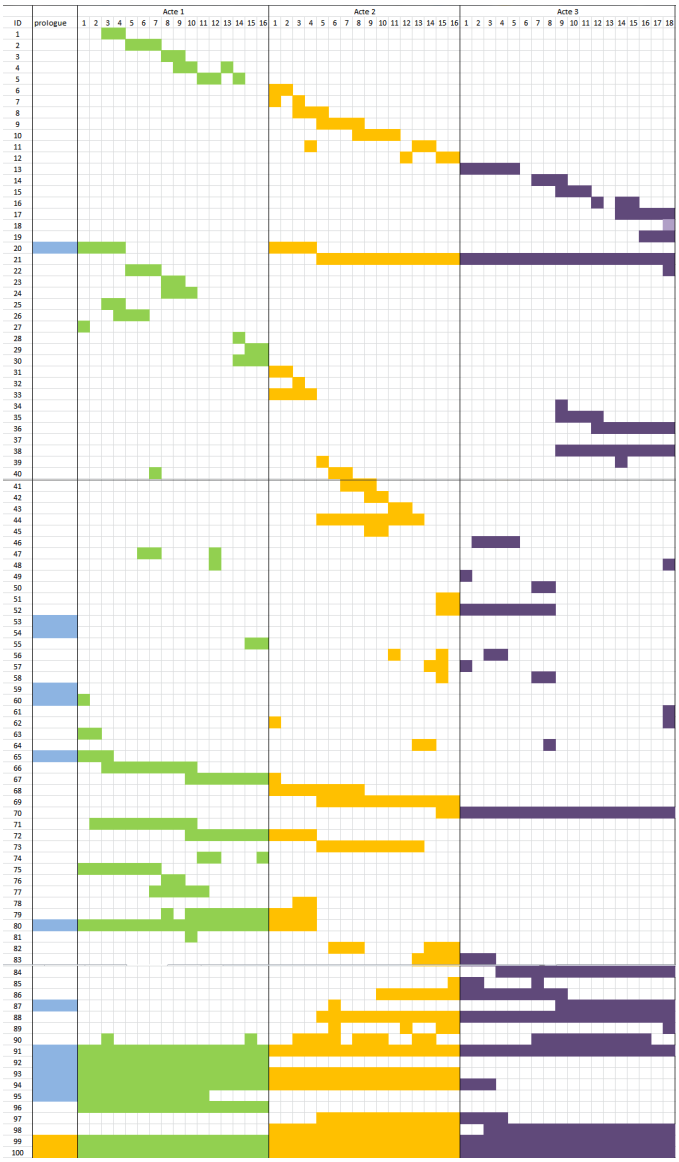

Figure 5: Matrix of distribution of the rehearsals according to the score of *Elena* over time.

one shortcoming is that the archive is not very pleasant to watch, especially on a small screen. To remedy this problem, we implemented a "spatial zoom lens" to make it possible to focus on the main actors in any given scene. Based on audio and video analysis, we generated alternative framings using methods described elsewhere [5], [6]. For example, during a monologue, we automatically compute a medium-shot of the speaker, and during a dialogue, we compute a two-shot that contains the speaker and his addressee (see details in Figure 6). This greatly improves the visibility of the main actors and renders it possible to focus on their gestures and facial expressions, as opposed to the general view, which is more useful to appreciate the choreography of their movements on stage.

Both the spatial and the temporal zoom lenses were assembled into an online application using Web Components. A demonstration video showcasing the application is available at http://spectacleenlignes.fr/hypervideo/ and the implementation is shared as open source. A screenshot of the application can be seen in Figure 7.

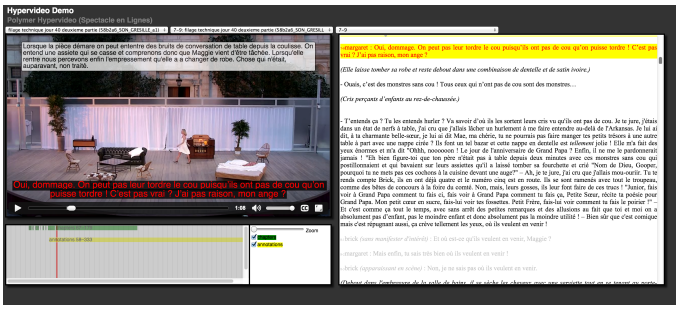

Figure 7: Generated hypervideo including subtitles, cue selectors, timeline, and connected play script showing a rehearsal scene from day 40. A demonstration video showcasing the application is available at https://youtu.be/GKmH8ucLgjo.

D. Mobile app for the creative crew

During the production of the archive, the creative crew reacted quite positively to the experiment and expressed interest in getting immediate feedback. While this had not been planned, we started designing mobile applications that they could tentatively use for *(i)* viewing the live feed being recorded on their smartphones and *(ii)* adding their own (signed) annotations and comments collaboratively. While not fully implemented, this feature was presented to the directors and their collaborators as mock-ups (Figure 8). These mock-ups were generally well received and are likely candidates as an addition to the existing system for future experiments.

Figure 8: Mobile app: extracts of the design mock-ups

VI. Conclusion

The research described in this paper was accomplished over a period of two years in close collaboration with the directors of the two documented productions and was contrary to all expectations well received by the performers and technical

Figure 6: Reframing of the videos during dialogues. Audio-to-text alignment temporally localizes the actors speaking and being spoken to. Video tracking spatially localizes them on screen. This figure displays the automatic reframing process obtained for several frames on the sequence $\mathcal{S}_4$. The arrows go from the speaker to the addressee.

crew. Filming the rehearsals from the director's point-of-view was a key element for this positive response. In future work, we would like to increase the video resolution of the recordings to 4K in order to better exploit the capabilities of our virtual zoom lens; and to experiment with stereoscopic 3D and 360° panoramic video which would allow to include reverse shots of the directors and their assistants, which are currently not recorded.

It is our hope that this archive illustrates the cultural heritage of the performing arts from a novel and interesting perspective. Audio, video and Web technologies were used to augment the raw video with fully integrated metadata and spatio-temporal zooming capabilities, which proved invaluable to make the archive attractive to researchers and amateurs. Nevertheless, the resulting archive is overwhelming and more focused future work is needed to better measure its impact and usefulness in the digital humanities.

Acknowledgments

This work was supported by Agence Nationale de la Recherche (ANR) project Spectacle-en-ligne(s) and Region Rhone Alpes project Scénoptique. All images and videos are reproduced courtesy of Célestins, Théâtre de Lyon and Festival d'Aix-en-Provence. Special thanks to Claudia Stavinsky, Jean-Yves Ruf, Leonardo Garcia Alarcon, Auxane Dutronc, Jérome Brunetière, Pierre Gandar, Pauline Lambert, Joelle Le Marec and the entire cast and crew of "Cat on a hot tin roof" and "Elena".

References

[1] I. Darling, *In the company of actors*. Shark Island Productions, 2007.

[2] M. Reason, *Documentation, Disappearance And the Representation of Live Performance*. Palgrave Macmillan, 2006.

[3] P. Bouchez, *Filmer l'éphémère: récrire le théâtre (et Mesguich) en images et en sons*, ser. Collection Information, communication. Presses universitaires du septentrion, 2007.

[4] J.-L. Godard, "Conversation," *Cahiers du cinema*, no. 177, April 1966.

[5] V. Gandhi, R. Ronfard, and M. Gleicher, "Multi-clip video editing from a single viewpoint," in *Proceedings of the 11th European Conference on Visual Media Production*, ser. CVMP '14. New York, NY, USA: ACM, 2014, pp. 9:1–9:10.

[6] V. Gandhi and R. Ronfard, "Interactive vertical editing of live-action video," in *Eurographics workshop on Intellligent Cinematography and Editing*, 2015.

[7] G. McAuley, "The video documentation of theatrical performance," *New Theatre Quarterly*, vol. 10, pp. 183–194, 5 1994.

[8] N. F. Bravo, "Filming theater: The audiovisual documentation as a substitute of the performance," Ph.D. dissertation, Brigham Young University, 2007.

[9] G. McAuley, "The emerging field of rehearsal studies," *Rehearsal and Performance Making Processes*, pp. 7–14, 2006.

[10] ——, "Not magic but work: making meaning in the theatre," *Theatre Research International*, 2008.

[11] P. Sankar, K. Jawahar, and A. Zisserman, "Subtitle-free movie to script alignment," *British Machine Vision Conference*, 2009.

[12] M. Caillet, C. Roisin, and J. Carrive, "Multimedia applications for playing with digitized theater performances," *Multimedia Tools and Applications*, Aug. 2013. [Online]. Available: https://hal.inria.fr/hal-00855102

[13] N. Sawhney, D. Balcom, and I. Smith, "HyperCafe: Narrative and Aesthetic Properties of Hypervideo," in *Proceedings of the the Seventh ACM Conference on Hypertext*, ser. HYPERTEXT '96. New York, NY, USA: ACM, 1996, pp. 1–10.

[14] T. Berners-Lee, "WorldWideWeb: Proposal for a HyperText Project," W3C, Proposal, 1990, http://www.w3.org/Proposal.html.

[15] J.-F. Bonastre, F. Wils, and S. Meignier, "ALIZE, a free toolkit for speaker recognition," in *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, vol. 1, March 2005, pp. 737–740.

[16] V. Gandhi and R. Ronfard, "Detecting and naming actors in movies using generative appearance models," in *2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, OR, USA, June 23-28, 2013*, 2013, pp. 3706–3713.

[17] T. Berners-Lee, "Linked Data," 2006, http://www.w3.org/DesignIssues/LinkedData.html, accessed July 15, 2013.

[18] R. Cyganiak, D. Wood, and M. Lanthaler, "RDF 1.1 Concepts and Abstract Syntax," W3C, Proposed Recommendation, Jan. 2014, http://www.w3.org/TR/rdf11-concepts/.

[19] R. Troncy, E. Mannens, S. Pfeiffer *et al.*, "Media Fragments URI 1.0 (basic)," W3C, Recommendation, Sep. 2012, http://www.w3.org/TR/media-frags/.

[20] E. Prud'hommeaux and A. Seaborne, "SPARQL Query Language for RDF," W3C, Recommendation, 2008.

[21] R. Verborgh, O. Hartig, B. De Meester, G. Haesendonck, L. De Vocht, M. Vander Sande, R. Cyganiak, P. Colpaert, E. Mannens, and R. Van de Walle, "Querying datasets on the web with high availability," in *The Semantic Web–ISWC 2014*. Springer, 2014, pp. 180–196.

[22] T. Steiner, R. Ronfard, P.-A. Champin, B. Encelle, and Y. Prié, "Curtains Up! Lights, Camera, Action! Documenting the Creation of Theater and Opera Productions with Linked Data and Web Technologies," in *International Conference on Web Engineering ICWE 2015*, Jun. 2015.

[23] D. Van Deursen, W. Van Lancker, E. Mannens *et al.*, "Experiencing Standardized Media Fragment Annotations Within HTML5," *Multimedia Tools and Applications*, pp. 1–20, 2012.

[24] T. Steiner, H. Mühleisen, R. Verborgh, P.-A. Champin, B. Encelle, and Y. Prié, "Weaving the Web(VTT) of Data," in *Proceedings of the 7th Workshop on Linked Data on the Web*, Apr. 2014.

[25] B. Stiegler, "The time of cinema." *Tekhnema*, vol. 4, pp. 62–114, 1998.