Time for a New Look at the Movies through Visualization

Ana Jorge[1, 2], Nuno Gil[1], and Teresa Chambel[1]

[1]LaSIGE, Faculty of Sciences, University of Lisbon, Portugal
[2]Faculty of Science and Technology, New University of Lisbon, Lisbon

Abstract — **Considered an important art form, a source of entertainment and a powerful method for educating, movies have great power to affect us, perceptually, cognitively and emotionally. Technological advances are making a huge amount of movies and related information available. Visualization has the potential to help handling this very interesting but complex richness. In particular, time-oriented visualization could help to capture, express, understand and effectively navigate movies along time: both the time when they were released, or viewed, and the time along which their contents are weaved, in each movie.**

In this paper, we present our recent work towards the inclusion of the time dimension in 2D and 3D visualizations, based on colors and tag clouds, at the movies space level, and down to the movies in an interactive web application to access, explore and visualize movies based on the information conveyed in the different tracks or perspectives of its content, especially audio and subtitles where most of the semantics is expressed. Our approach has the ultimate goal of helping finding and navigating information through interactive visualizations, and providing insights through analytical tasks or more ludic and artistic uses.

Index Terms — **Art, color, content based retrieval, interaction, 3D, tag clouds, time, videos, visualization.**

I. Introduction

Movies are considered an important art form, a source of entertainment and a powerful method for educating, having great power to affect us, perceptually, cognitively and emotionally, by combining diverse symbol systems, such as images, texts, music and narration. The movie industry is very active, and technological advances are enabling the access to movies over the Internet and interactive TV, making accessible a space with a huge amount of movies and related information. However, all the richness that makes this movies space so interesting, inside each movie and outside where in many ways movies relate to each other, comes with a challenging complexity to handle. One of the problems is the fact that video is not structured and changes along time, and so, accessing all the data of a video is often not an easy task. The other is the huge amount of movies available that have been released along time.

Visualization techniques could help to handle the complexity and express the richness in these information spaces, through intuitive and effective ways to convey meaningful information [1] inside and about the movies. Many researchers [2][3] have stressed the importance to develop methods for extracting and highlighting interesting and meaningful features in video to effectively summarize and view them. And time-oriented visualizations could help to capture, express, understand and effectively navigate movies along time: both the time when they were released, or viewed, and the time along which their contents are weaved, in each movie. And this is the focus of this paper.

The importance of visualization in conveying knowledge is undisputed, and having its roots in scientific reasoning, it has traditionally been viewed as an analytical tool for sense making [4]. But its inherent aesthetic qualities have also been associated with its effectiveness. It has been argued that the higher the aesthetic value of the visualization is, the more engaged the viewer is in trying to decode its meaning. And actually, many visualizations resemble works of art [5], especially since recent years when artists and designers stepped in to expand the conceptual horizon of infovis as artistic practice, as art based on data [4]. Although using the same basic techniques, their motivations and creations are different. Visualizations can be used to express a point of view, to induce the viewer to ponder in different aspects of our culture; as well as to analyze. Besides the more traditional disinterested analysis, it may be that much of the value of visualization comes from its ability to change attitudes. It is our intention, in the line of this work, to allow for both perspectives.

We are designing and developing MovieClouds [6] as an interactive web application, to access, explore and visualize movies based on the information conveyed in the different tracks or perspectives of its content, especially audio and subtitles where most of the semantics is expressed, and with a special focus on the emotional dimensions, either expressed in the movies and felt by the users. It adopts a tag cloud unifying-paradigm, from the movies space down to the individual movies. So far, movies space has been explored as a collection of movies based on their titles and overviews of their contents. Time has only been considered at each movie, through timelines of the different content perspectives.

In this paper, we present more recent work towards the inclusion of the time dimension also at the movies space level, and down to the movies, in 2D and 3D visualizations, based on colors and tag clouds. Time can be relevant to capture when movies were released, access them or get additional overviews and insights about patterns or tendencies in movies content along time. For now, and due to the amount of information available, at the movies space level, we experimented mostly with movie genres and ratings. But we intend to address more content based information, like at the individual movie level, to find out, for example, if the amount of screams was a trend at some decade or around some periods (e.g. around Halloween), if movies tended to have more romantic audio mood or use more slang in speech at some point, and how this relates with their genres and ratings.

This can provide support to more utilitarian uses of finding and navigating information and movies, to analytical tasks for getting overviews and understanding patterns, to get insights possibly in serendipitous ways [6][7], or making evident and draw attention to some aspects related with the way movies and our viewing habits evolved and reflected or have influenced our culture along time, in more artistic ways.

Section II addresses the conceptual framework that informs our approach, in terms of visualization principles and time-oriented visualization. Section III presents related work. Section IV focuses on the time-based visualization of movies in MovieClouds, at the level of movies space and individual movies. Section V takes a step further to reflect on the visualization options, based on principles presented in sec. II. The paper ends with conclusions and perspectives for future work in section VI.

II. Conceptual Framework

This section presents the main concepts relevant to inform our approach, in terms of visualization, their design principles, and especially the aspects that concern time-oriented visualization.

A. Visualization & Information Understanding

Visualization is defined as the use of computer-supported, interactive, visual representations of data to amplify cognition [8], helping people to understand and analyze data. It is about how 'seeing turns into showing', how empirical observations turn into explanations and evidence [9]. Actually, advances in data visualization have emerged from research rooted primarily on visual perception and cognition [1]. If human communication is based on a certain type of language, the more articulated the language, the more success in the message to be understood [10].

Shneiderman [11] stressed the importance to develop visual approaches matching the great skills in human perception in terms of scanning, recognizing and recalling images rapidly and also in detecting changes in size, color, shape, movement, or texture and in the identification of patterns, group or individualized items. Principles have been defined to inform strategies that improve information understanding. We present the most central principles, relevant to inform our own approach:

1. Simplicity/complexity: Maeda [12] argues that the important information must be preserved, and that simplicity is an aesthetic preference. The aim of the design is to achieve rich and understandable data, despite its complexity, as a balance: "How simple can you make it *versus* How complex must it be?" Maeda suggests the three key words: Shrink, Hide, Embody, while Tufte suggests: "to clarify, add detail." [13].

2. Layout: The overview, i.e., the way people see and understand data at first, is conditioned by the overall view of the layout so the displaying of the elements are of most importance [14][15][16]. An easy overview gives the user the sense of context and control, necessary for the user to find information [14].

3. Overview, Details, Zoom: It must be easy to choose information of interest and get details [11]. The visualization of complex and multidimensional info should allow a global overview that provides context, and also the possibility to clarify details without leaving the context. This allows comparing global and detailed views for a better understanding of the info [13].

4. Relate: People apprehend by comparison [10]. Thus, showing relationships among items eases information understanding by allowing these comparisons. It can also reveal, in one step, a range of options [13]. And through highlight techniques, attention can be drawn to specific items in a group of many [11].

5. Layering and Separation: It is of crucial importance the inclusion of different levels of information in order to maintain detail and coherence. Separated layers of info make possible the co-existence of complementary data without losing its meaning [13][15].

6. Color and Information: Tying color to information is as elementary and straightforward as color technique in art, "To paint well is simply this: to put the right color in the right place," (Paul Klee) [13]. Color allows to address multidimensionality, when used as a way to catalog, mediate, represent names, quantities, or relations, and as an aesthetical component.

7. Narratives of Space and Time: Temporal and spatial dimensions are usually hard to handle [13], but they allow to represent movement and info flow. Time can be mapped in space as a static representation, or be represented by dynamic visualizations. Having time controlling the evolution in other dimensions can be a solution to the problem of mapping multidimensional data into two dimensions. Also, the visual progression of the temporal data should be illustrated in a continuum mode, without gridding and framing [13].

B. Time-Oriented Visualization

Users are commonly interested in the evolution of data over time, the detection of trends and patterns that lead to insights, and to understanding the data [17][18][19][20]. Temporal relations enables us to learn from the past to predict, plan, and build the future, and to effectively deal with media that depend on time, as is the case of video. But visualizing time-oriented data is not an easy task. There are different and highly customized methods, because it is difficult to consider all aspects involved when visualizing time-oriented data, and time itself has many theoretical and practical aspects [17]. It might be strange the existence of more than one consideration about time, as the experience in the physical world shows only one physical space where we and all other things exist. Frank [21] states that, "The differences are not in the physical (objective) properties of time and space, but are found in the conceptual models for time or space people use." [21] for different tasks and contexts. [17] developed a systematic view that addresses three categorization criteria to help deal with time representation: Time, Data, and Representation.

B1. Time: what are the characteristics of the time axis?

The same amount of time may have distinguished modes of being considered, according to its granularity and on the purpose of the analysis. We consider the following sub-criteria:

Temporal primitives: time points vs. time intervals

A time axis can be composed of time points (instants in time), e.g. the day a film arrives in the market; or of time intervals (with an extent) e.g. 2h30m as the duration of the film, that can be specified by two time points or by a time point plus a duration [21].

Structure of time: linear vs. cyclic vs. branching

[17][21] distinguish three different structures: linear, cyclic, and branching time.

If events elapse on a linear time, it extends from the past to the future in an ordered manner. If the structure is cyclic, then events happen in a linear way but they repeat periodically in time. (e.g., the seasons of the year). In practical applications it is often useful to unroll a cyclic time axis to a linear time axis, the most common structures. Branching time is the property that shows several events happening simultaneously or in alternative scenarios. Andrienko et al. [22] add the time with multiple perspectives as the way to represent several points of view in the same length of time, e.g. different opinions about the same film simultaneously.

Absolute and Relative Time & Temporal Determinacy

Time can be absolute or relative if it is, or it is not, attached to a specific moment. In the first case, an interval of time starts at a specific point in time (anchored) and in the second, it doesn't (unanchored) [23]. Temporal determinacy is related to the duration of an event (e.g. a film) that can be of knowledge or not. Indeterminacy happens when there is no information about the time or the duration of an event [24].

B2. Data: what is analyzed?

Deals with the data that ties to the time axis. Like the time axis, it also has a major impact on analytical and visual approaches. [17] suggest the following sub-criteria:

Frame of reference: abstract vs. spatial

Abstract data are not per se connected to any spatial layout, while spatial data contain an inherent spatial layout, conditioned by natural or modeled realities.

Number of variables: univariate vs. multivariate

It concerns the number of time-dependent variables.

Level of abstraction: data vs. data abstractions

Tufte claims "Above all else, show the data". However, if larger data sets must be analyzed, new problems arise like overcrowded and cluttered displays. Data can be condensed, by deriving data abstractions, e.g. aggregated data is particularly useful to drive overview+detail interfaces. Feature visualization also follows the idea of computing data abstractions. In the context of time-oriented data a third derivable information unit is important: events, that can be user-defined or found by methods of Artificial Intelligence, lifting the data analysis to yet a higher level of abstraction.

B3. Representation: how is it represented?

The achievement of an efficient representation of information implies a good integration of the visual methods, analytics and those which are centered on the user. It must be taken into account, then, the data type (meaning and application) and the user (tasks and needs). [17] concentrate on two sub-criteria:

Time dependency: static vs. dynamic

Static representations visualize time-oriented data in still images, while dynamic representations that automatically change over time, e.g. animations.

Dimensionality: 2D vs. 3D

Distinguishes between 2D and 3D presentation spaces. The need for 3D is heavily discussed, some researchers argue that 2D is sufficient for effective data analysis, while others defend

III. RELATED WORK

Most relevant related work is presented, organized in visualizations, with a special focus on time, and approaches to the access and visualization of video.

A. Visualization with Time, Color and Clouds

Some of the most relevant visualizations are presented in Fig.1 and discussed next.

a- *ThemeRiver* [25] represents the number of occurrences in a collection of topics that are displayed as a flowing river.

b- *Flickr Flow* [26] represents the dominant colors in picture taken around a year.

c- *British History* (http://www.bbc.co.uk/history/interactive/timelines/british/index.shtml) presents historical facts of the United Kingdom since the Neolithic period. The user can overview the information or filter the information of interest.

d- *Portuguese Empire Expansion and Decline* [27] is an animated sequence of images which aim is the visualization of the expansion and decline of the Portuguese Empire.

e- *TaggedColors* (http://www.visualcomplexity.com/vc/project.cfm?id=41) uses colors and the tag cloud paradigm for the visualizations. It allows to search and visualize images from Flickr through their colors and words.

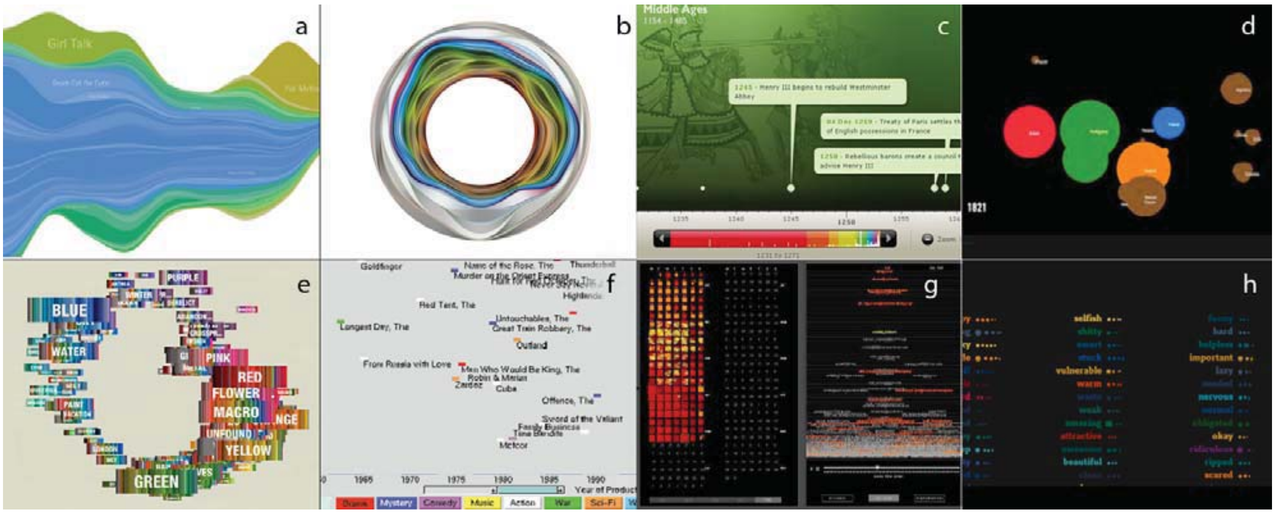

Fig. 1. Visualizations with time, colors, and clouds.

f- *FilmFinder* [28] presents the possibility of choosing films according to their popularity and gender through a panel of spots with the color of the gender.

g- *Post History* [29] allows the user to be aware of the mail correspondence history. When a contact name is highlighted, the corresponding emails sent by this person are highlighted.

h- *We feel fine* [30] scans the occurrences of "I feel" and "I am feeling" in blogs in order to look for human feelings. The feelings are assigned to colors and can be matched to a specific gender, age, country and location.

The visualizations in Fig.1 relate to the proposed work in different aspects in terms of time, data and/or representation:

Time: visualizations (a,b,c,d,f,g) implicate the flowing of time. (a,c,d,f,g) use linear time, along a timeline, while (b) adopts a cyclic time. The determinacy is, however, only in (a,c,f,g). Although (b,d) represent time, it cannot be measured.

Data: In respect to data (a,c,f,h) are multivariate and (b,d,e,g) deal with only one or two variables. All the visualizations have abstract frame of reference and (a,b,d) are data abstractions whereas in (c,e,f,g,h) the data is visible per se.

Representation: In relation to the time dependency (a,b,c,e,f,g,h) are static and (d) is dynamic, for (d) is an animation, while the first ones do not move, unless the user chooses to. The tri-dimensionality is not present in any of the referred visualizations.

All the visualizations seem to consider the seven principles referred earlier, in order to optimize the visualizations in respect to their functionalities, as well as in respect to the aesthetic aspects. Associations with colors have been made in all the visualizations. Color has been used as *label,* in order to code meaning (d,e,f,h); to speed recognition (a,c,d,e,f,g,h); *as measure* (a,d,e); *as a formatting tool* (a,d); *as representation* (a,b); *as beauty* (a,b,c,d,e,f,g,h).

B. Video Access and Visualization

Sites like IMDB (imdb.com) provides info about actors, directors, characters, genres, film ratings by users, etc. Other, like Netflix (.com) also allow accessing and watching movies through on-demand video and online rental. YouTube (.com) allows to publish and watch videos, search, comment, share, and get recommendations. But they do not support access to info conveyed in subtitles or audio tracks nor emotional info, and navigation does not go much further then through lists.

Film Finder [28], referred before in this article, allows the user to search for films based on their duration, genre, title, actors and directors, and present the selected movies in a starfield graphic of colored spots based on their date and popularity

Media Streams [31] allows to work on the structure of video content through the annotation, visualization and retrieval of video content. Informedia Project [3] is a digital video library that based on content processing and use of metadata to index, search, navigate, and visualize videos. It supports search by words and allows the analysis of video content, based on the image and speech of both current and archived content collections. Open Video [32] allows the retrieval, storage and research of video and its elements. It allows history, peripheral, and shared views, overviews and previews of the video data. It can save the history actions of the user referring to downloads, annotations and previews.

Videosphere (www.bestiario.org) representing TED's videos. It represents a 3D sphere video space with videos linked according to its semantic and allows the user to navigate inside and outside the sphere form. In our previous works: we provided interactive 3D visualization and navigation of video [33], to explore cultural and aesthetic properties of videos and videos spaces, using colors for categories, and temporal representation as image loops to summarize the videos at the video space level; and a 2D interactive application based on a physical particles system [34] that could visualize and explore videos based on their color dominance, rhythm and movement, from different views to allow the user to choose specific properties of interest. Static temporal representations included colored stripes with the dominant colors along the video, and slit-scans to capture in a still image, the movement in the video. Dynamic temporal representations included loops of main scenes in the videos. All these representations provided summarization views of the individual video, where the only content addressed was color and motion, and the temporal dimension was not explore to organize or view the video space level.

IV. Visualizing Movies in Time with MovieClouds

This section presents most recent work on MovieClouds concerning visualization in time, especially at the movies space level, but also down to the individual movie level. We have been working on techniques for the classification and exploration of movies based on different perspectives of their content (subtitles, emotions in subtitles, audio events, mood in audio, and felt emotions) and have some movies classified in the different perspectives, but we do not have yet a large amount of movies classified to fully explore realistic visualizations at the movies space based on these content properties. So, as a first approach, we are making our time based visualizations, at the movies space level, with information gathered from IMDB through its web service, in terms of movie title, time of release, genre and rating. Genre is a more qualitative and categorical information, resembling the emotional dimensions of the movies, while rating is quantitative, allowing to experiment with more variables of different nature, like we will have with the information about movie contents.

Next, we present the rationale behind our main design options, exemplified in the different visuali-zations that the users may navigate and choose from, in terms of the criteria: Time, Data and Representation, presented in section IIB.

A. Time: what are the characteristics of the time axis?

Temporal Primitives

At the movies space, we use time points as the date each movie is released, although we may aggregate info to represent for example all the movies released in the same month, year, decade, etc. (e.g. Fig.2-5). At the individual movies, subtitles, emotions in subtitles and sound events are represented as time points (dots on the 3 first timelines beneath the video in Fig.5c), while mood in audio and felt emotions are represented as time intervals (as colored rectangles made of little squares in the 4th and 5th timelines Fig.5c).

Structure of Time

At the movies space, we adopt both linear and cyclic time structures:

1. Linear as a timeline that can have different time scales (days, months, years, decades…). This allows to have a clear sequential perspective. Fig.2 presents examples of this structure: a) movies by genre and month, represented by genre color as their absolute amount, allowing to know which months have more or less movies, while the relative percentages in b) allow for better perception of the proportion of the genres along time. In c) Tag clouds reflect the frequency (like a) and explicit the genres names, reinforcing the information also conveyed by the colors.

2. Cyclic as a 3D helix or helicoids that goes around every year and up to the next one. This matches a synesthetic spatial representation of time that reflects the cyclic nature of the years and allows to compare and find patterns around and along the years. This view can fold and unfold between linear and cyclic. Fig.3a,b exemplifies this view

3. Cyclic as a 3D cylinder, to represent separate years, that can be stacked for better comparison around the year. Fig.3c,d exemplifies this view for color and tag clouds representations of genres along the year.

4. Cyclic as a 2d circle, to represent separate years that can be put side by side for comparison. Fig.4 and Fig.5a) present different views of this type, with genres by month or by day, as aggregate information or individualized movies.

At the individual movie level (Fig.5c), each track timeline is linear, but since the different timelines present what is happening simultaneously at the different tracks or perspective of the movie content, we may consider it as branching time.

Absolute and Relative Time & Temporal Determinacy

Movies are represented by their absolute time of release, although in aggregate views, they may be less strict (e.g. the decade instead of the exact day). At the movie timelines, events are placed in their absolute time. Only time intervals (mood and felt emotions) have their duration determined.

B. Data: what is analyzed?

Frame of reference: abstract vs. spatial

MovieClouds' data is abstract since it is not connected to any spatial layout.

Number of variables: univariate vs. multivariate

It is multivariate, since there are several variables changing in time, about the movies (e.g. genre and rating of different movies) and along the movies (e.g. sound events and subtitles within the movies).

Level of abstraction: data vs. data abstractions

We adopt different levels of data abstraction. At the movies space, movies can be aggregated by different time scales as amount (absolute or relative (%)) of movies per genre (e.g. months and years in Figs.2 and 3) or represented as individuals (Fig.4d,e). Since a movie can have *n* genres, they count for *1/n* in each of the *n* genres (in the amount of movies in the aggregate view, or dot size in the individualized view), so the total equals the number of films, and the proportions are correct. From an aggregate view, it is possible to navigate to an individualized view, for more detail. At this view, 'on over' one of the dots highlights all the dots (genres) for the same movie. 'On click' gets the user to that movie (Fig.5c). At the movie level, tag clouds are used to aggregate and overview the movie content at a chosen track (mood in audio inf Fig.5c), and each content event (e.g. an audio event like a scream) can be accessed individually along the corresponding timeline. These movie overviews can also be seen at the movies space as an intermediate level of abstraction (content overview for a collection of movies). This is shown in [6], but in a way that is independent of time. These levels of abstraction provide the users with an overview+detail interface to zoom and navigate in and out, in accordance with their goals and tasks.

Fig.5 presents an example of browsing through the different levels of detail:

1. In Fig.5a), users have an overview of the movie genres around the year of 2011, represented as colored tag clouds (on the right). After selecting a genre (e.g. action), the other genres turn white, for this one to stand out (on the left). Now, if they select action movies in a particular month (e.g. Jan 2011):

2. they are presented with all the action movies that were released in that month, represented by title and information about genres and rating (Fig.5b). We explored different views for this information, adequate for different types of tasks. Genres (colored dots) and ratings (highlight) for a global view and relative ratings; and ratings with an additional scale for more accurate information. The order may also be by title or by rating, to ease comparison and search. From here, users may select a movie to explore further:

3. they are lead to the individual movie view, where they may watch the movie, and search, overview and browse through its indexed contents (Fig.5c).

C. Representation: how is it represented?

The achievement of an efficient representation of information implies a good integration of the visual methods, analytics and those which are centered on the user. It must be taken into account, then, the data type (meaning and application) and the user (tasks and needs). [17] concentrate on two sub-criteria:

Time dependency: static vs. dynamic

Most MovieClouds' representations are static in the sense that they do not automatically evolve in time. But they are interactive, allowing for exploration. The exception would be at the movie level, when the movie is playing, where the dots on the timelines and the corresponding tags on the clouds are highlighted when their time arrives (Fig.5c). We are planning to experiment some animations at the movie space level, to highlight chosen properties (e.g. the amount of western movies released) along time.

Dimensionality: 2D vs. 3D

Most representations are 2D, and they serve their purpose as such. The exceptions are the helix and cylindrical views, used to view information both around and along the years. Note that this view can be moved around and zoomed to see information from different perspectives. For example, in Fig.2b, looking inside and down the helix provides a spiral perspective that allows to see, to some extent, the information condensed in one view, without the need to go around.

Color Choice

Itten [35], Kandinsky [36] and Klee [37], among other theorists and artists developed theories of color that contemplate the emotion impact of color on people. Plutchik [38] directly related colors with emotions, and in general, brighter colors are associated with happiness, excitement and relaxing and dark colors to anxiety, boredom, sadness and negative emotions [39][40].

We believe that films, as well as colors, can have a great impact on emotions. It makes sense, then, to index films by emotions felt by the viewers, and adopting colors for their representation, what we did in the emotions felt track of MovieClouds (4th timeline in Fig.5c). Although the meaning of colors are subjective and for that some choices are empirical, there are some associations that are commonly accepted. Based on related work, we made our own mapping of colors to genders (Fig.4c):

1. *Red*, Action. The emotions expressed as associated with red are rage [38], action, excitement [36][41][42].
2. *Purple*, Thriller. Purple resembles the color of blood but in a darker and mysterious version.
3. *Orange*, Musical. This color is the second most lightened color of the spectrum, next to yellow.
4. *Yellow*, Comedy. Yellow induces joy, happiness, ecstasy [38][40] and stimulation [42].
5. *Green (bright),* Western. Connected with adventure and nature [36][38][41].
6. *Green*, Adventure. Along with the bright green this medium green induces, we believe, the "Indiana Jones" feeling.
7. *Blue (dark)*, Drama. Dark blue is sad [38], restful, quiet [42] and timid [36].
8. *Violet*, Horror. Violet represents the illness and disgust [36][38].
9. *Rose*, Romance. The *rose* flower brought to this color the passion connotation.
10. *Brown*, History. Sepia photos instantly induces an emotion related to ancient times.
11. *Grey,* War. This is the color of uniforms and vehicles of the military people, and is related with lack of joy.

The colors were chosen in order to fulfill an aesthetic property, with a concern about the harmony among them, as well as to serve as gender code.

Other Representation Options

To represent information, we consistently adopted tag clouds, dots, regions and colors. At the movie level, movies are represented by colored aggregated regions or tag clouds (for example to represent genres per month, in Figs.2-5), or individual colored dots per movie (Fig.4c,d). At the movie level (Fig.5c), overviews adopt tag clouds, and individual occurrences of the events adopt colored dots (audio, subtitles), or regions (moods and emotions). Color is used for emphasis, and can be used selectively to highlight only the tags or properties that the user chooses (e.g. Fig.5).

Tag clouds have been used as tools for aggregation, and also for analytical tasks, which are fun, engaging, aesthetic and creative [43]. They have been used mainly for data like unstructured text, where the frequency of occurrence is of importance. Information wise, video is also unstructured, and the extracted features may be represented by text, reflecting their semantics and frequency, and thus by tag clouds that can be used to make summaries or overviews of their content. At the movies space level, tag clouds can also be used to summarize semantics and frequency (e.g. most popular movie genres).

V. Reflections on the Visualization Options

We make a reflection on the visualization options in MovieClouds, with a special focus on the temporal dimension, based on the seven visualization principles presented in section IIA:

1. Simplicity/complexity: We consider that the important information is presented with different levels of detail that can be used to hide and show what becomes more relevant, as the user chooses to go up for overview or down for detail and the actual content. This, along with the consistent adoption of the unifying tag cloud paradigm, dots and colors, allows to make it more simple at each level. Overall, with the different levels, some complexity is added, but along with the power to access movies from very rich and unconventional perspectives based on their content.

2. Layout: The consistent layouts can provide the user with the sense of context and control, helping to find the information.

3. Overview, Details, Zoom: It is easy to choose information of interest and get details about the complex and rich information, from global overviews that provides context, with the possibility to clarify details in context.

4. Relate: The adoption of colors associated with the dots and tag clouds allows to identify information and relate overviews with the content at different levels, with highlights to have the chosen information stand out along time.

5. Layering and Separation: Different layers of information are provided. For example, from a view showing movie genre occurrences around the year in Fig.5a), it is possible to open a complementary view with the list of movies corresponding to that genre at that time (e.g. month) in Fig.5b), keeping the context, adding detail, and allowing to navigate down to the actual movies (Fig.5c).

6. Color and Information: Colors as used to convey information, especially with mapping purposes, but it also has an important aesthetical component. Actually, in the user evaluation of MovieClouds [6] the users appreciated the adoption of colors to help in the overviews and navigation and was found aesthetical.

Following the principles referred earlier: color *labels,* by segmenting the genres; *measures,* by making visible the frequency of the films plus their genres; *imitates reality,* when associating the flowing genre colors through the circular timelines to the image of a flowing river; and *enlivens* the layout.

7. *Narratives of Space and Time:* Time is mostly represented by static but interactive representations at the movies space with representations that adopt a continuum mode along time (without framing and gridding) helping to follow the time flow when viewing and browsing the information from the movies space to down to the movies. At the movie level, as the movie plays, selected information is dynamically highlighted in the content tracks, helping to make it stand out.

VII. Conclusions and Perspectives

We presented our most recent work towards the inclusion of the time dimension also at the movies space level of MovieClouds, and down to the movies, in 2D and 3D visualizations, based on colors and tag clouds. For now, and due to the amount of information available, at the movies space level, we experimented mostly with movie genres and ratings, that have different natures, qualitative or categorical, and quantitative, like those that are relevant within the movies. Different visualizations were explored with different, alternative or complementary purposes, adequate for different tasks. Visualizations were characterized in the dimensions found more relevant for time-based visualizations, and a reflection was made based on most central visualization design principles, that were taken into account in the proposed visualizations.

Next steps include refining and extending current visualizations, based on our goals, our reflections, and user evaluation, towards effective, rich, expressive and, in some contexts, artistic visualizations that can help to provide insights about movies and their impact on us. We intend to explore more properties, and especially content features, based on subtitles, audio, and emotions, at the movies space, as we get more movies processed in these perspectives, along a more significant number of years, even decades. The representation of series is also in our horizon. Their seasons and episodes could be represented in a way that captures their evolution. The system could also trace and present personal movies viewing histories, allowing users to reflect on the movies they tend to watch in different periods of their lives, and even compare the emotional impact the movies had on them along time, and relate that to their content.

ACKNOWLEDGEMENT

This work was partially supported by FCT through LASIGE Multiannual Funding, the VIRUS research project (PTDC/EIA-EIA/101012/2008), and a FCT fellowship under the UT Austin|Portugal Program.

REFERENCES

[1] S. Few, *Data Visualization: Past, Present, and Future*, IBM Cognos Innovation Center, 2007.
[2] G. Daniel, and M. Chen. "Video Visualization". In *Proc. of the 14th IEEE Visualization 2003 (Vis'03)*, 2003.
[3] A. Hauptmann, "Lessons for the future from a decade of informedia video analysis research", *Image and Video Retrieval*, pp. 595–595, 2005.
[4] F. B. Viégas, M. Wattenberg. "Artistic data visualization: Beyond visual analytics". In *Proc. of the 2nd international conference on Online communities and social computing (OCSC'07)*, Springer, 2007.
[5] A. Lang, "Aesthetics in Information Visualization", *Media Informatics Advanced. Seminar on Information Visualization*, 2008/2009. /www.medien.ifi.lmu.de/lehre/ws0809/hs/docs/lang.pdf
[6] P. Martins, T. Langlois, and T. Chambel, "MovieClouds: Content-Based Overviews and Exploratory Browsing of Movies". In *Proc. of Academic MindTrek'2011*, Tampere, Finland, 2011.
[7] T. Chambel, "Towards Serendipity and Insights in Movies and Multimedia". In *Proc. of Int. Workshop on Encouraging Serendipity in Interactive Systems.* Interact'2011, Lisbon, Portugal, September 5-9, 2011.
[8] S.K. Card, J.Mackinlay, B. Shneiderman. *Readings in Information Visualization: Using Vision to Think.* Morgan Kaufmann, 1999.
[9] E. R. Tufte, *Beautiful Evidence*, Graphics Press, 2006.
[10] E. Gombrich, *Art and Illusion.* Princeton University Press, 1959.
[11] B. Shneiderman, *Designing the User Interface*, 3rd ed. Addison Wesley, 1997.
[12] J. Maeda, *The Laws of Simplicity.* Cambridge: The MIT Press, 2006.
[13] E. R. Tufte, *Envisioning Information*. Graphics Pr, 1990.
[14] J. Cohen, B. MacWhinney, M. Flatt, and J. Provost, "PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers", *Behavior Research Methods*, vol. 25, no. 2, pp. 257–271, 1993.
[15] W. Horton, L. Taylor, A. Ignacio, and N. L. Hoft, *The Web Page Design Cookbook: All the Ingredients You Need to Create 5-Star Web Pages*. Wiley, 1995.
[16] M. Morris, R. J. Hinrichs, and M. E. S. Morris, *Web Page Design: A Different Multimedia,* 1st ed. SunSoft Press, 1996.
[17] W. Aigner, S. Miksch, W. Muller, H. Schumann, and C. Tominski. "Visualizing time-oriented data – a systematic view". *Comput. Graph*., June 2007.
[18] F. H. Post, B. Vrolijk, H. Hauser, R. S. Laramee, and H. Doleisch, 'The state of the art in flow visualisation: Feature extraction and tracking', in *Computer Graphics Forum*, 2003, vol. 22, pp. 775–792.
[19] S. F. Silva and T. Catarci, "Visualization of linear time-oriented data: a survey". *Proceedings of the First International Conference on Web Information Systems Engineering 2000*, vol. 1, pp. 310–319.
[20] M. Weber, M. Alexa, and W. Müller, "Visualizing time-series on spirals". In *Proc. of the IEEE Symposium on Information Visualization*, 2001, p. 7.
[21] A. Frank, "Different Types of 'Times' in GIS", in *Spatial and temporal reasoning in geographic information systems*, Oxford University Press, USA, 1998, pp. 40–62.
[22] G. Andrienko, N. Andrienko, U. Demsar, D. Dransch, J. Dykes, S. Fabricant, M. Jern, M.-J. Kraak, H. Schumann, and C. Tominski, "Space, time and visual analytics", *Taylor & Francis, vol. 24, no 10*, Oct. 2010.
[23] R. Snodgrass, "Temporal databases", *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*, pp. 22–64, 1992.
[24] I. A. Goralwalla, M. T. Özsu, and D. Szafron, "An Object-Oriented Framework for Temporal Data Models", in *Temporal databases: research and practice,* Etzion et al. (Eds)., vol.1399, Springer Verlag, 1998, pp.1–35.
[25] S. Havre, B. Hetzler, and L.Nowell, "ThemeRiver: Vizualizing Theme Changes over Time". *IEEE Symposium on Information Visualization (InfoVis), 2000.*
[26] F. Viégas and M. Wattenberg, "Flickr Flow". [Online]. Available: http://hint.fm/projects/flickr/. [Accessed: 23-Jun-2012].
[27] P. Cruz and P. Machado, "Visualizing the Portuguese Empire Expansion and Decline", in _*Proceedings of Artech'2010*, Guimarães, Portugal, 2010, pp. 76–71.
[28] C. Ahlberg and B. Schneiderman, "Visual Information Seeking: Tight Coupling of Dynamic Query Filters with Starfield Displays", in *Proc. of SIGCHI conf. on Human factors in computing systems*, Boston, Massachusetts, USA, 1994, pp. 313–317.
[29] F. B. Viégas, D. Boyd, D. H. Nguyen, J. Potter, and J. Donath, "Digital artifacts for remembering and storytelling: Posthistory and social network fragments". *Procs of the 37th Annual Hawaii International Conference on System Sciences*, 2004.
[30] J. Harris and S. Kamvar, "We Feel Fine", *We feel fine*, 2005. [Online]. Available: http://www.wefeelfine.org/. [Accessed: 26-May-2012].
[31] M. Davis, "Media Streams: an iconic visual language for video representation", *Readings in Human-Computer Interaction: Toward the Year 2000*, pp. 854–866, 2000.
[32] G. Geisler, G. Marchionini, M. Nelson, R. Spinks, and M. Yang, "Interface concepts for the open video project",in *Proceedings of the Annual Meeting-American Society for Information Science, 2001*, vol.38, pp.58–75.
[33] T. Rocha, and T. Chambel. "VideoSpace: a 3D Video Experience". In *Proc. of Artech'2008*, 4th International Conference on Digital Arts, Porto, Portugal, 2008.
[34] J. Martinho, and T. Chambel. "ColorsInMotion: Interactive Visualization and Exploration of Video Spaces". *Proc. of Academic MindTrek'2009*, Tampere, Finland, 2009.
[35] J. Itten, *The Art of Color: The Subjective Experience and Objective Rationale of Color*. John Wiley & Sons, 1974.
[36] W. Kandinsky, *Do Espiritual na Arte*, 7th ed. Dom Quixote, 2006.
[37] P. Klee, Paul Klee, *Escritos sobre Arte*, 2nd ed. Edições Cotovia, 2001.

[38] R. Plutchik, *Emotion: A Psychoevolutionary Synthesis*. Harpercollins College Div, 1980.

[39] C. Boyatzis and R. Varghese, “Children’s emotional associations with colors”, *Journal of Genetic Psychology*, vol. 155, pp. 77–85, 1994.

[40] M. Hemphil, “A note on adults’ color emotion associations”, *Journal of Genetic Psychology*, vol. 157, pp. 275–281, 1996.

[41] R. A. I. T. de Faria, *Arte e percepção visual: uma psicologia da visão criadora.* Cengage Learning Editores, 2004.

[42] N. Kaya and H. Epps, “Relationship between color and emotion: A study of college students”, *College Student Journal,* vol. 38, pp. 396–405, 2004.

[43] M. Wattenburg, and F. Viegas. “Tag Clouds and the Case for Vernacular Visualization”. *ACM Interactions*, XV.4.2008.

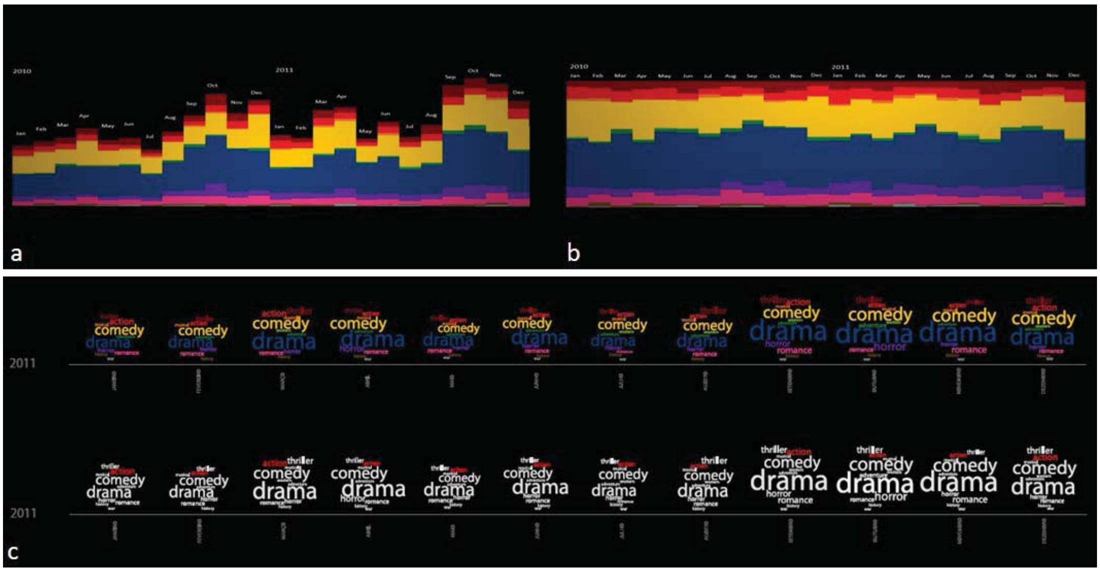

Fig. 2. Linear Time: a) Movies by genre and month in 2010-2011, represented by genre color (absolute amount); b) Same (in %); c) Movies by genre in 2011, represented by tag clouds, and color (top), or color (red) in selected genre: action (bottom).

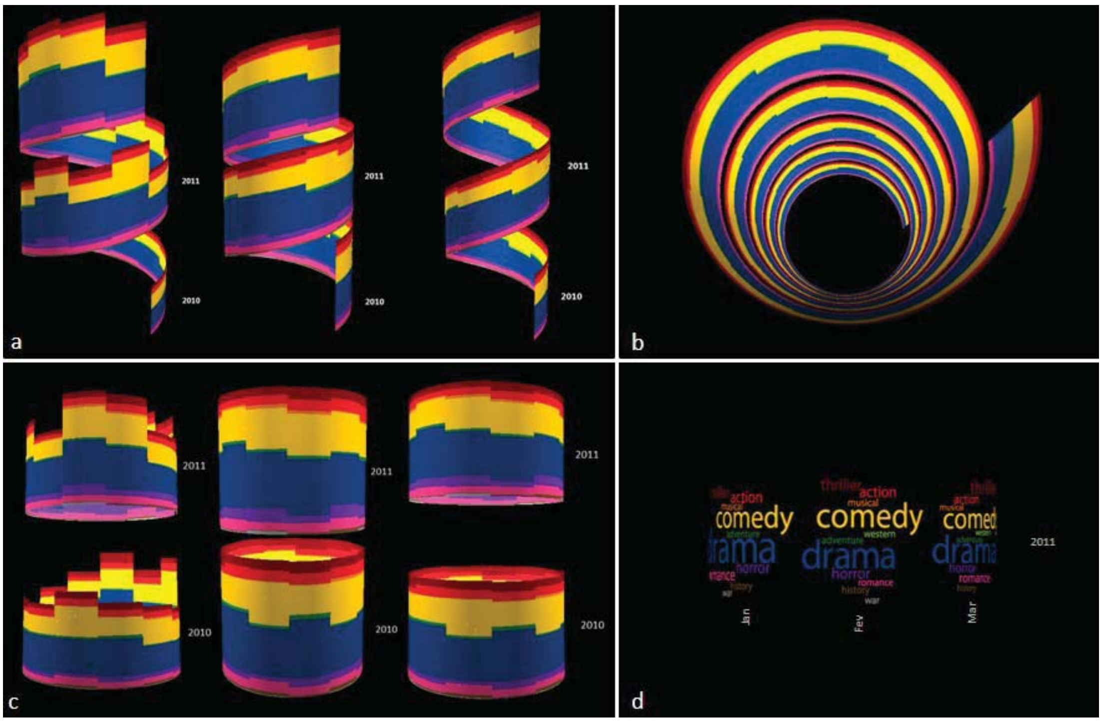

Fig. 3. Cyclic Time in 3D: Same info as Fig 1, represented by: a) helixes; b) View from top, showing a spiral around and along the years; c) Cylinders per year in color; d) Cylinder per year in colored tag clouds.

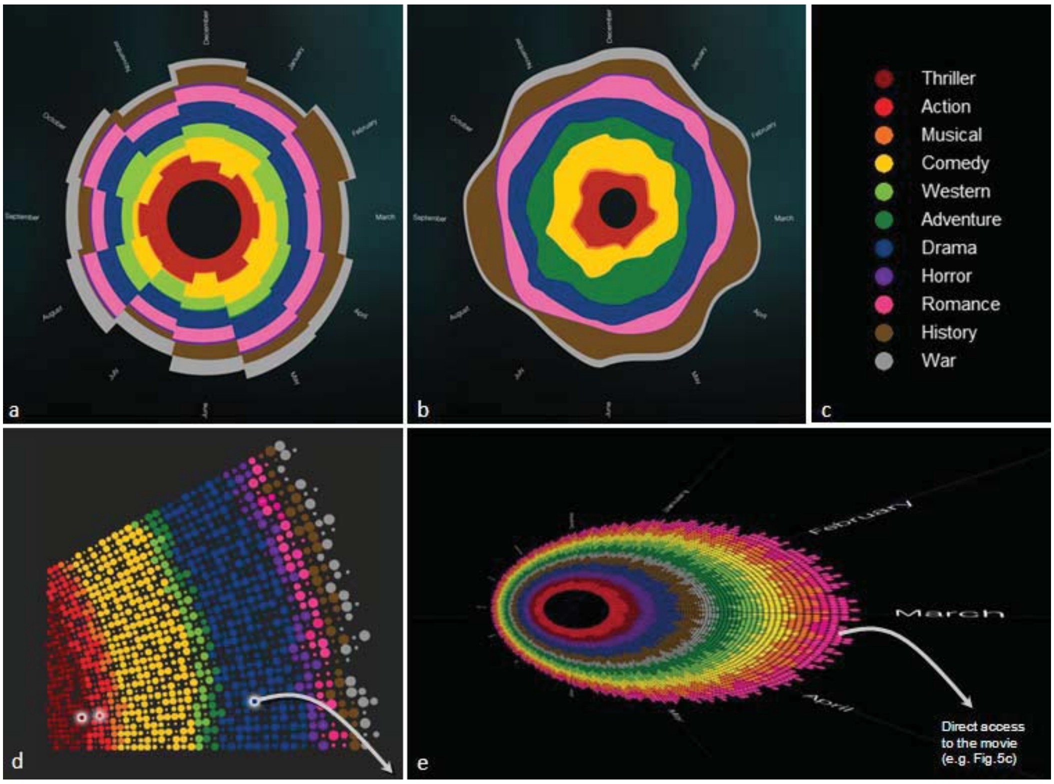

Fig. 4. Cyclic Time in 2D: a) Films by genres by month in color around a year; b) Films by genres by day; c) genre colors; d) Detail for one month in circle individualized by colored dots (genres in films), one day per column, ‘on over’ all dots (genres) of same film get highlighted, ‘on click’ access movie (Fig5c); e) Full circle in different genre ordering, in a perspective that allows to turn around for more detail on the chosen months, to help analyses (may have more than one year shown at the same time, for comparison).

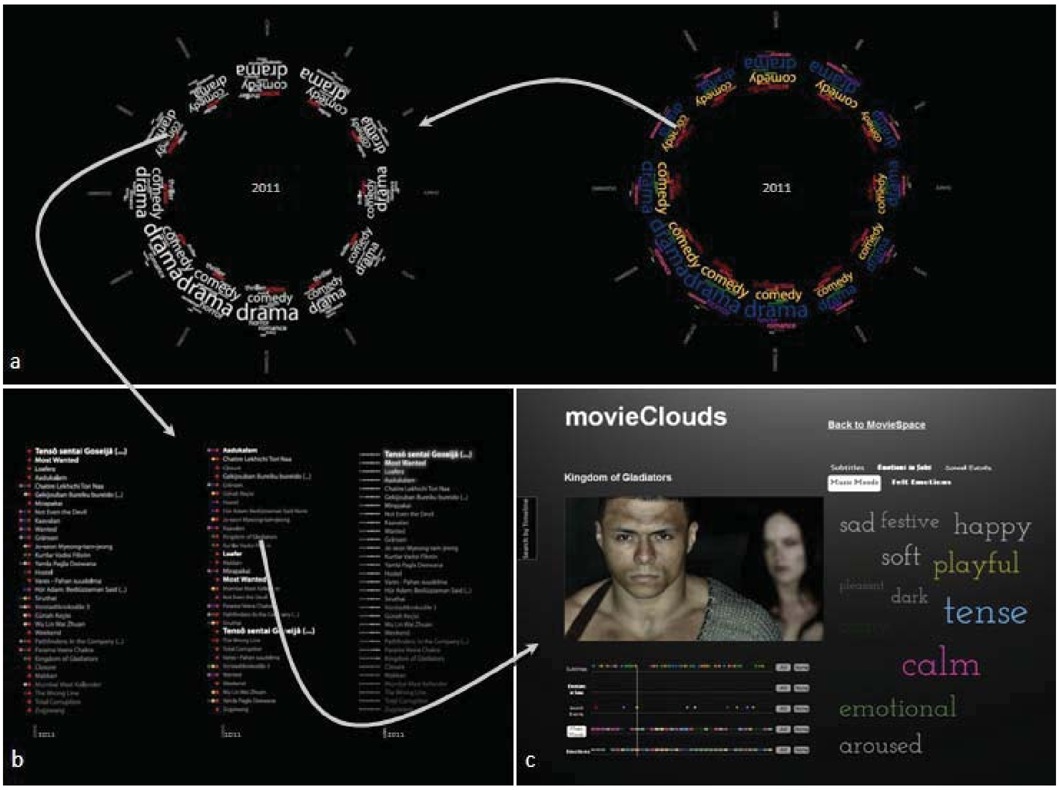

Fig. 5. From Movies Space to a Movie: a) cyclic view in tag clouds; b) 3 alternative views of the list of action films in the chosen month (genres colors and rating in brightness ordered by rating; same ordered by title; rating in brightness and scale; c) Accessing a chosen movie, with content clouds on the right (mood in audio track now) & track timelines beneath the video [6].