

INTEGRATING MULTIMEDIA DOCUMENTS IN 3D CITY MODELS FOR A BETTER UNDERSTANDING OF TERRITORIES

...

Commission XX, WG XX/YY

KEY WORDS: Multimedia, Urban 3D model, Billboards, Documents, Story telling

ABSTRACT:

Digital 3D representations of urban areas, through their growing availability, are a helpful tool to better understand a territory. However, they lack contextual information about, for example, the history or functionality of buildings. On another side, multimedia documents like images, videos or texts usually contain such information. Crossing these two types of data can therefore help in the analysis and understanding of the organization of our cities. This could also be used to develop document search based on spatial navigation, instead of the classical textual query. In this paper, we propose four approaches to integrate multimedia documents in a 3D urban scene, allowing to contextualize the scene with any type of media. We combine these integration approaches with user guidance modes that allows to guide the user through the consumption of these media and support its understanding of the territory. We demonstrate the usefulness of these techniques in the context of different projects within the Lyon area (France). The use of multimedia documents integrated into a digital tour allows, for example, the iconic buildings to be contextualised or to understand the evolution of a territory through time.

1. INTRODUCTION

Digital 3D representations of urban areas are becoming widely available. Multiple tools exist to visualize them such as Google Earth¹ or Cesium². By allowing to wander into virtual cities, these tools allow a better understanding of a territory. However, most of them are limited to a 3D representation of the buildings, without providing more information about their functionality, their evolution or other elements of context. Aside from 3D representations, multimedia (e.g. images, texts or videos) also plays an important role in the understanding of the urban landscape. Archival images, videos on the history of a district, job descriptions of certain industries or interviews in a district can provide important additional information about a territory. While these media have the potential to enrich virtual 3D cities, the link between those multimedia and digital 3D representations is rarely made.

Online mapping systems (e.g. Google Maps) allow to wander in a 2D or 3D view and pick information through labels or pins that simulate points of interest. By linking multimedia to a spatial position, users can find information more easily in a desired area. This can be seen as another form of search, rather than formulating a query on a web browser. As such, these tools lack ways to fully integrate the media into a 3D view of the city, and usually lack of explanatory media on the organisation of a district or a whole city, on its history, etc.

We propose a model that allows to combine a 3D view of the city in which the user can navigate with a collection of multimedia that provides additional information about the underlying 3D scene. We present new ways to integrate each multimedia document in the 3D scene as well as ways to guide the user through a collection of different documents. More precisely, we introduce four modalities of multimedia integration in a scene (shown in blue in Figure 1) that are presented in Section 3:

- The **3D geo-pinned multimedia** uses the principle of pins that target a point of interest and are linked to a multimedia document. Our extended pins gives an overview of the linked document.
- The **3D geo-web renderer** shows documents at given positions in the scene, always facing the camera such that the user can get the information from any place in the scene.
- The **extended document** superimposes images on a 3D model and complete them with textual data. The extended document is linked to a precise camera position.
- The **slideshow** allows to make a presentation in a view by placing documents in a 3D space.

A large number of documents linked in a scene do not necessarily guarantees a better understanding of a territory, as users might need to be guided through them in order to understand them. We thus present ways to guide the user through its consumption of documents in the scene. Our three user guidance modes range from a fully sequential order of consumption, to a free mode where the user can wander in the scene and pick the documents as he wishes. These modes and the way they can be mixed with the multimedia integration modalities are presented in Section 4.

The combinations of the integration modalities and user guidance modes open the door to a large variety of choices in the way to use multimedia documents in combination with 3D virtual cities. We demonstrate the potentialities of several of these combinations in the context of multiple projects developed in the city of Lyon (France) in Section 5. Such projects include a web-documentary that aims to deconstruct preconceived ideas about a territory or presentations about the future of a neighbourhood using a tangible model.

¹ <https://www.google.com/intl/fr/earth/>

² <https://cesium.com/>

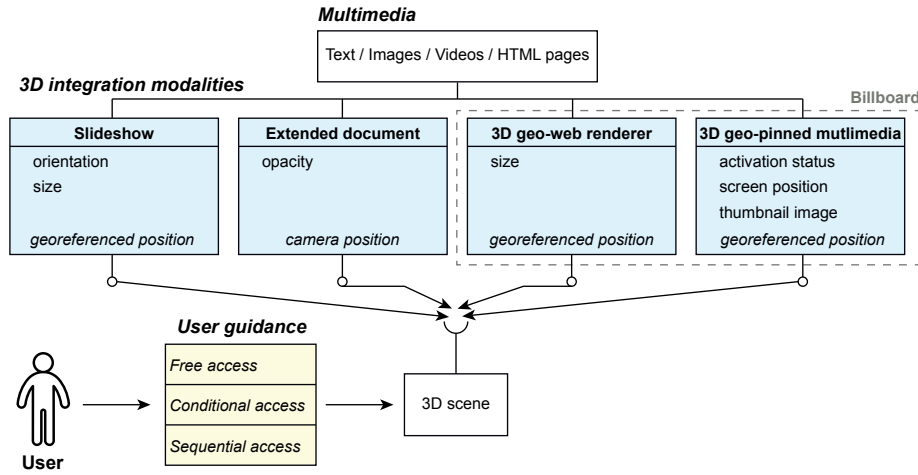


Figure 1. Architecture diagram of our proposed model. The multimedia can be integrate in the 3D scene using one of the multimedia integration approaches (in blue). Each approach is characterized by attributes that allow to parameterize the integration. The media are linked to the scene by either a georeferenced position or a camera position (shown in italic). The user can consume the integrated media in the scene through different guidance modes (in yellow).

2. RELATED WORKS

Many tools exist to navigate in a virtual urban environment (Bleisch and Dykes, 2015, Boutsis et al., 2019, Jaillot et al., 2021). Thanks to the navigation in the scene, users can visualize a city or any 3D geometry under different angles. This can greatly help to better understand an urban space by virtually navigating in it.

Other tools, such as web-documentaries, also allow to walk in an urban space and better understand it without making use of a 3D scene. These tools can target a wider audience as they are web-based and do not require any specific hardware. Their 2D representation of the digital territory is enhanced by additional information such as video (“Correspondances”³) or images (“Balade au merlan”⁴). However, these tools remain limited in their interaction because of the 2D representation of the world that they carry. We would like to take up this spatial contextualization of documents to the third dimension and integrate such documents in 3D walking tours.

Enhancing 3D scenes with multimedia documents have been used to show the evolution of a city through time (Jaillot et al., 2021, Chagnaud et al., 2016). Such documents can bring information about the state of a city at a certain date through images of buildings at different times or their chronology. These works have introduced several methods to visualize historical photo documents in a 3D scene, allowing to show the urban past but also the future planning of cities. However, they lack the possibility to provide additional information on these pictures to help reach a better understanding. They also do not allow to use any other type of media, such as videos of web pages, or to interact with them.

Other methods have tried to contextualise a digital model with quantitative data (Bleisch and Dykes, 2015) or text and images (Boutsis et al., 2019) in a scene. This gives another perspective on the data and a better understanding of a territory by linking information to a geospatial position. These methods are good examples on how a 3D representation and multimedia or data

can complement each other. However, they only integrate one type of content and do not give an overview of the different documents present in the scene, making the search for a specific document or data harder. We want to provide a new method of document search through a digital representation of a territory by allowing an explicit overview of the different multimedia available.

Rather than images or data, web pages can also be integrated in a 3D context to contextualise a 3D scene (Wijnants et al., 2015). This allows giving the user access to more content, such as external websites, while not losing the focus on the digital representation of a territory. The interest of integrating the web page directly into the 3D scene is also to support the relationship between the target and its related annotation objects (Seo et al., 2015). While we follow the same goal, we aim to provide a model that allow to integrate any type of multimedia in the scene and to bring information on a specific geometries like building or district (Samuel et al., 2016).

Multimedia documents can also have interesting metadata like their title, source, publication date, key content, tags. (Gan et al., 2014) gives an overview of the existing techniques used to visualize a multitude of documents both in 2D and 3D environment. However, they do not target geographic applications and the integration in 3D urban scenes.

While integrating multimedia document in the 3D digital representation of the city can greatly improve the understanding of such area, the user can quickly be lost in this multitude of information and not know where to start searching. Some works have shown the interest in guiding the user by making him follow a logical sequence of information. It also helps to engage the user and allow him to be more effective in its reading of the information (Othman et al., 2011). In our work, we combine methods to integrate multimedia into 3D urban scenes with ways to interact with a collection of documents, providing different levels of guidance to the user.

Existing methods of integrating multimedia do not satisfy our need to manage multiple types of document or to provide interaction with them. We provide several ways to integrate such

³ <https://correspondances.tv5monde.com/>

⁴ <http://unebaladeaumerlan.fr/>

multimedia in the 3D scene, depending on the type of such documents but also on their goal and content. These integration modes can be combined with user guidance modes that help reach better user experience and understanding of territory.

3. MULTIMEDIA INTEGRATION APPROACHES

Our objective is to integrate different types of multimedia documents into a 3D urban scene in order to make the representation more informative and interactive. The integrated documents can thus be videos (eventually 360 videos) as well as images, textual documents or web pages, as shown in Figure 1. They are all georeferenced and have a position in space. Although this position is a natural link with the 3D scene describing the city, integrating the media visually in a simple interface such that the user can mentally connect them to the 3D representation of the city is not easy.

We propose four methods to integrate such documents in the 3D scene, explained thereafter and shown in blue in Figure 1. The choice of one of them depends on the type of media but also on their meaning or their number. This integration makes it possible to have all the information on a district on the same web interface, easing the understanding of such areas. But it can also help to find documents or information by wandering in the digital representation of the city and browsing the multimedia that are available. Each of these integration modalities can be used in the 3D scene with different user guidance modes (free access, conditional and sequential) that will be described in Section 4 and are shown in yellow in Figure 1.

The first two approaches (**3D geo-pinned multimedia** and **3D geo-web renderer**) integrate a media into a 3D scene such that it is linked to a precise spatial location and always shown facing the camera (similar to the billboard technology). The **extended document** is linked to a camera position and superimposes the media to the 3D scene, while the **slideshow** projects a series of media onto a 3D plane in the scene. All of these methods aim to be integrated in a web-based library for 3D visualization of geographical data.

3.1 Billboard inspired approaches

These approaches consist in an arrangement of thumbnails directly shown in the 3D scene and targeting a precise spatial point. The thumbnail is always facing the camera such the user can access the information from any position in the scene. The two approaches differ in the way the document is displayed through the thumbnail: either directly or summarized in an interactive pin. In either cases, the integration is characterized by a georeference position (GPS coordinates).

The **3D geo-pinned multimedia** is based on the principle of pins that target a point of interest. We enrich the traditional pin by adding a thumbnail above it that gives a visual feedback on the nature and content of the multimedia. This pin is meant to be interactive: when selected, the linked document appears in 2D on the screen in front of the 3D scene (see example in Figure 2). The pin and its thumbnail are always facing the camera while the multimedia document is displayed as a 2D element that will not move with the camera. This approach allow to give the user a global vision of all the contents available in a scene. The user can then pick from this data and decide to get more detailed information about a place or building pointed out by the pins. The interactivity of the pin can be deactivated, allowing to

control what the user can access as content at a given moment. In this case, the user can still see the pin and its thumbnail, thus knowing that there is a document accessible, but will not be able to consume the linked document. In addition to its spatial position, the 3D geo-pinned multimedia is thus also characterized by its activation status, the position where to display the document on the screen and its thumbnail image (see Figure 1).



Figure 2. An example of the 3D geo-pinned multimedia approach. The pins show the available multimedia on the 3D scene while the selected document is shown on 2D at the left of the screen.

The **3D geo-web renderer** also points to a precise location in the 3D scene but the linked document will be directly shown to the user when he approaches it. Again, the document will be displayed always facing the camera so that the user can see it from any position in the scene. This way, the user can continue walking around in the 3D scene and mentally put this document in context. This is a naturally good approach to integrate web pages: the 3D geo-web renderer can point to an url, allowing the user to directly browse a web page from within the 3D scene. An example of the integration of a web page with this approach is shown in Figure 3. The size the media appears in the 3D scene is an additional parameter in this integration method.



Figure 3. An example of the 3D geo-web renderer approach. An html page is shown over the 3D scene.

These two billboard approaches are similar in the way they are integrated into a scene, they both target a precise location in space and allow to quickly see which content is accessible while moving into the scene. However, by summarizing the information into small pins instead of directly showing it, the 3D geo-pinned multimedia allows to display more document without cluttering the space. It could thus be a preferred solution when a large number of documents need to be shown in a small space.

3.2 Extended Document

The **extended document** approach displays a multimedia document in a 2D context on top of a 3D scene. The integration of the document in the 3D scene is done by modifying the position and orientation of the camera, rather than positioning the document in the 3D space (see Figure 4). In order to access different documents in the same scene, the user has to choose the documents he is interested in in a list. When selected, the camera will move to the right position and overlay the document on top of the 3D scene. In order to keep the link with the 3D scene, the camera will move smoothly toward the target position, allowing the user to locate more easily its position in 3D. Documents in this approach are thus characterized by the position and orientation of the camera, as well as the opacity of the element to be overlaid.

The extended document is composed of three distinct windows:

- The **navigator windows** (Figure 4, left-most window) which allows to navigate between the different documents with their titles and reference dates. This window also has a filter system for a better navigation among the documents .
- The **inspector windows** (Figure 4, right-most window) gives a preview of the document that will be superimposed in the 3D scene, as well as all additional information about it such as its description, source, and publication date. It allows the user to have a first understanding of the document and verify that it is the media that he wants to consume.
- The **visualization windows** (Figure 4, middle window) is activated when the user decides to engage the visualization in the inspector window. It will smoothly move the camera to the correct position and superimpose the document on the 3D view. It also allows to change the opacity of the document in order to compare it with the underlying 3D.

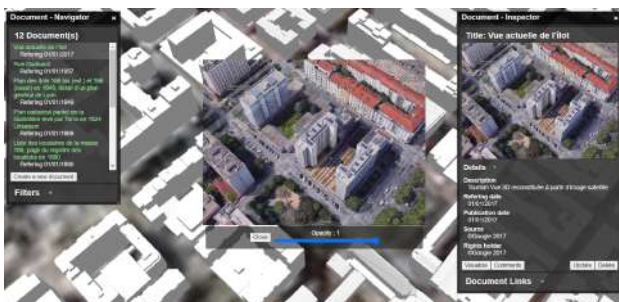


Figure 4. An example extended document with its three windows: the navigator window showing the list of available documents (left); the inspector window showing the metadata of the selected document (right); the visualization window that superimposes the document on the 3D scene (middle).

This approach will be preferably used with images, or animated images, as it allows to directly compare it with the 3D view of the scene. It can help users to realize the evolution of a district or a building, or observe it under another format.

3.3 Slideshow

The slideshow is anchored at a 3D position in the 3D scene and can display the documents on a 3D plane (see Figure 5). This plane is traditionally parallel to the ground but can be also set vertically to contextualize face buildings. The documents will thus not follow the camera as the user move into the scene. The slideshow is characterized by its center position (georeferenced position), its size and its orientation. This method overlays the 3D model of buildings on top of the documents, that can be images or videos of data layers on a district, or different states of a territory. For example, one could drop the image of the transport network of a city and see how it fits with the morphology of that city. The slideshow is made to integrate a series of documents: it simulates a presentation in the 3D world in which one can scroll through different documents, located at the same place, in order to illustrate problems or ideas on a district. The user can easily drag and drop different images or videos into the scene, allowing to quickly illustrate different ideas, but also to work in team and experiment different solutions and illustrations. A geographic mask is made available to prepare data directly in dedicated GIS information systems tools.



Figure 5. An example of the slideshow. The configuration window (on the left) allows to set its parameters. One image of the slideshow is shown on a ground plane in the 3D scene.

4. USER GUIDANCE

Our different multimedia integration approaches allow to integrate a large variety of multimedia in a scene, from an archive image to a web page. When becoming too numerous in one scene, the user could get lost among them thus breaking the goal of our approach to bring a better understanding of a territory. We thus design several ways to consume a collection of documents, by possibly guiding the user through their consumption:

- **Sequential access:** The documents are provided in a given order. As in a guided tour, this allows to build a narrative and guide very precisely the user through the documents. This mode limits the order of freedom of the user, it can only act on the time to spend on each document, but helps to build a better understanding of the documents.
- **Conditional access:** Going through some documents might be needed to access others, but the order is not completely constrained. Several documents are available at one moment in which the user can chose freely, but viewing some or all of them is needed to access the next ones.

- **Free access:** the user can consume the different multimedia in the scene as he wishes.

Most of our integration approaches can be used with any of these user experiences, although some are more naturally suited for some specific user guidance. The only exception is the slideshow, that is essentially made for sequential access. It allows a linear succession of images or videos on a 2D map and can be used in the case of a presentation to contextualise a district.

The extended document, because it has smooth transitions between documents and is not anchored at a specific position in the 3D scene, is well suited to be used in a sequential approach. It can be used to build a *guided tour* as a list of extended document where each of them will be contextualized in the tour with additional text. It can, for example, consist of a sequence of documents that show the evolution of a district with photos from different eras.

Both the 3D geo-pinned multimedia and the 3D geo-web renderer can be used naturally with any of the guidance. However the 3D geo-pinned multimedia is well suited for conditional access thanks to its activation status. It allows the user to see all the contents that is available in the scene through the pins, but the activation of some of the documents can be conditioned to the viewing of other ones.

5. IMPLEMENTATIONS

We experimented our different approaches on various use cases around the city of Lyon (France). Each of these use cases mobilizes one of our multimedia integration approaches, combined with a user guidance mode, allowing a wide variety of user experiences. These implementations are implemented within the UD-Viz library⁵, a JavaScript library allowing to visualize, analyse and interact with urban data.

5.1 Chemistry Valley: 3D geo-pinned multimedia

We demonstrate the use to the 3D geo-pinned multimedia in the context of a web-documentary, developed as part of a collaboration between a training center for the Chemistry industry Interfora⁶ and the Lyon metropolis. This web documentary aims to make the citizens rediscover the area of the Chemistry Valley (South of Lyon, France) through a 3D digital wandering where the user can interact with documents arranged in the 3D scene.

A large number of media, from different types were to be set in the scene. As examples:

- A photographic observatory⁷ of the Chemistry valley, consisting of a set of pictures of the chemistry valley from different angles.
- Additional layers as for instance bus lines, the atmospheric index, etc. can be used thanks to numerous data available via WFS (Web Feature Service) and WMS (Web Map Service) in the open data data.grandLyon.com⁸.



Figure 6. The 3D geo-pinned multimedia modality used in the context of the Chemistry valley. The pins show an overview of the different media available in the scene.

- Jobs descriptions located in the chemistry area. These pdf files provide information on workers in the area.
- Videos of stakeholder interviews in the Chemistry valley on the theme of employment and training.

The 3D geo-pinned multimedia is the best choice in this case since it allows to easily navigate between a large number of media, without cluttering the scene. Additionally, some of these media must be seen before others in the project specifications, naturally leading to using a conditional access mode. This also claims in favor of using the 3D geo-pinned multimedia.

Each interactive element in this scene can be configured through a JSON file that contains the geospatial position, the display thumbnail, its activation status, the text content and the document link. An example of such configuration file can be seen in Figure 7.

```
"episode-1-data":{
  "content-1":{
    "lock": false, // document activation
    "position":{ // coordinates in the scene
      "x": "1843554.77",
      "y": "5165405.73",
      "z": "220"
    },
    "imgUnlock": "./assets/img/Observatoire/cheminee.jpg",
    "imgLock": "./assets/img/Observatoire/cheminee.jpg",
    "text": "Vallee de la chimie - Observatoire photographique",
    "src": "https://umap.openstreetmap.fr/fr/map/vallee-de-la-chimie"
  }
}
```

Figure 7. An example of a JSON configuration file used for the 3D geo-pinned multimedia.

The conditional access allows to prioritise some contents, especially for the interviews that were created specifically for this experience. These videos were made to be seen in a given order in order to understand the whole issue. As a reminder, in the 3D geo-pinned multimedia modality, documents can be naturally deactivated. Their pin and thumbnail are still shown on the 3D scene but the user can not interact with them to access the

⁵ <https://github.com/VCityTeam/UD-Viz>

⁶ <https://www.interfora-ifaip.fr/>

⁷ <https://www.caue69.fr/1/page/10622/page>

⁸ <https://data.grandlyon.com/>

document. In order to better guide the user towards the accessible content, we symbolize the deactivation of the media by a lock image superimposed to the thumbnail. An example of a visualization of the Chemistry valley in this project can be seen in Figure 6. The atmospheric index is used as a color map on the ground of the 3D model while several pins show the available content.

The integration of multimedia in the 3D scene makes it possible to show the different professions that exist in the Chemistry valley, and help the users to know more about the Chemistry industry in this territory. To further this understanding, we have also integrated urban data into the scene by using geoservices⁹, spatial web services that make geodata available in a structured form. Finally, in order to show the valley differently, we have integrated WMS (Web map service) and WFS (Web feature service) which allow the display of georeferenced geometries such as building cadastres.

Thanks to the 3D geo-pinned multimedia method, we have brought another form of information search through the wandering in the 3D scene. The user can pick up interviews with the stakeholders of this territory and get information on where they are located and which company they belong to before even starting to watch the video. This makes it easier for him to choose the content that interests him.

5.2 Flying campus: 3D geo-web renderer

We use the 3D geo-web renderer method in the context of the project "IMUV/Flying campus"¹⁰. The aim of this project is to create a virtual space in which photographic exhibitions, conferences, or other events can be organised. It is an interactive space, floating above the city, where a variety of elements can be arranged and where several users can gather. Elements to be displayed can be of any types, including web pages. The use of the 3D geo-web renderer allow to integrate any type of media into the 3D space, such that the user can still be in immersion while consuming the content. One of the use case is a conference room in which a 3D geo-web renderer is placed to broadcast a video, the users can gather in the room and freely interact with the video. Figure 8 shows a user observing the city in 3D while looking at a map showing the organisation of the city.

We implemented the 3D geo-web renderer using the CSS3DRenderer technology (a web rendering method from the javascript library ThreeJS¹¹) that allows to include web pages, with which the user can interact, in the 3D scene. Our implementation consists in an overlay of two types of rendering: a CSS3DRenderer which shows the document to be displayed and the WebGLRenderer which is the 3D scene rendering of the walkthrough (Seo et al., 2015). We make a part of the 3D scene rendering transparent and render the CSS3DRenderer context behind the WebGLRenderer context. This allow the content of the document in the CSS3DRenderer to appear in the 3D scene.

By placing these billboards in the 3D scene of the flying campus, the user does not loose focus and has all the information in one web application. If the document to integrate is a web page, the user can freely interact with it as if it was facing the window of a web browser.

⁹ <https://geoservices.github.io/>

¹⁰ <https://demo.liris.cnrs.fr/vcity/flyingcampus/>

¹¹ <https://threejs.org/>

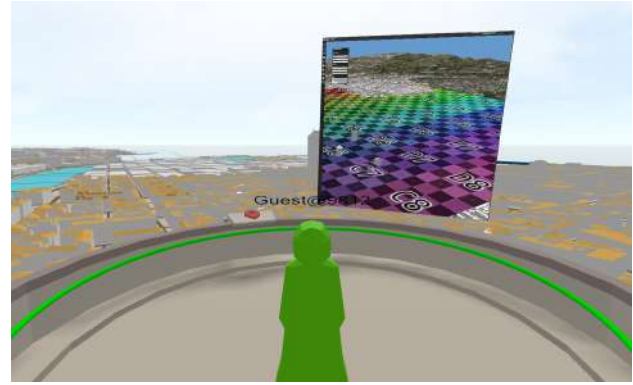


Figure 8. An example of the use of the 3D geo-web renderer in "Flying campus", a virtual gathering place. The user can still observe the 3D scene while the integrated multimedia is shown to the user.

In this context, the user is free to consume the content of these multimedia as he wishes and is not limited by an order to follow (free access).

Flying campus demonstrates the value of this multimedia integration as a way to virtually share multimedia experience and build participative experiences.

5.3 Historical guided tour: Extended document

In this project, we use the extended document as a way to show the evolution of a district over time. Our goal was to integrate a collection of archival photos from 1760 to 2017 in the 3D scene. Each picture shows how the district looked like at a specific date along with additional information. By allowing to superimpose the picture on top of the 3D scene, at its correct position, the extended document is naturally suited for this kind of use cases. By playing on the opacity slider, the user can more easily observe the morphological evolution of a building or a district between two periods. An example showing the historical organisation of a district with some additional information is shown in Figure 9



Figure 9. An historical cadastral plan is shown superimposed on top of the 3D scene by using the extended document modality.

To bring more coherence between these documents, we used the sequential access mode as a user guidance. We thus developed a **guided tour** in which the user can follow the chronological evolution of the district and better understand how it has evolved through the ages.

This use case allowed us to provide information about the evolution of a district thanks to archive images and the integration in the 3D scene.

5.4 Gratte-Ciel project: Slideshow

We use the slideshow approach in the context of an event organized during the Anthropocene week¹². Several multidisciplinary students (Geomatician, urban planner, information-communication) gathered on the issue of using a tangible or digital model as a technological tool for the intelligibility of territories. The students produced images or videos as an information layer over a district of Lyon, that they could contextualize with the 3D model of the same district. Three examples of such layers are shown in Figure 10. The images were produced such that they fit the underlying 3D geometry and can thus be superimposed with it. The slideshow method allows to see the geometry of building on top of these layers, allowing to better understand the interaction between those buildings and the data represented in the different layers.

The students could drop their visualisations onto the 3D map and animate a presentation by going through the different maps. They have proposed still image of geographic data or animated simulation with multi-agent model.

The aim of this project was also to show how tangible models could help in understanding urban projects. In addition to the 3D scene, we also provided the students with a tangible model of the neighbourhood that was built using white Legos. The white color allows to easily project information into the model. Rather than only seeing their maps in the digital 3D scene, students could project them onto the tangible model. Figure 11 shows an example projection on the model. We calibrated a video-projector such that the 3D scene in UD-Viz was matching with the tangible model, and we projected the slideshow module on the model. Using the slideshow on a tangible model allows to more easily understand and discuss about the interaction between the building and other data.

In these examples, we could observe the interest in adding an image to a 3D scene and having a way to present information linked to the 3D world. The mix of 3D modelling of buildings and videos or map backgrounds allowed to better understand the problematic of the district and the possibilities of the new urban project.

5.5 Building an home made application

All the described components are delivered as open source content available at <https://github.com/VCityTeam/UD-Viz>. Each of the demonstration project can be tested using the following links:

- Chemistry valley project using the 3D geo-pinned multimedia: <https://github.com/VCityTeam/UD-Demo-TIGA-Webdoc-ChemistryValley>
- Flying campus using the 3D geo-web renderer: <https://github.com/VCityTeam/UD-Viz/blob/master/examples/Billboard.html>

- Historical guided tour using the extended document approach: <https://github.com/VCityTeam/UD-Demo-VCity-Spatial-multimedia-db-Lyon>. The docker can be found at <https://github.com/VCityTeam/UD-Demo-VCity-Spatial-multimedia-db-Lyon>
- Gratte-ciel project with the slideshow approach: <https://github.com/VCityTeam/UD-Demo-Anthropocene-GratteCiel-UD-Demo-Anthropocene-GratteCiel-docker>. The docker can be found at <https://github.com/VCityTeam/UD-Demo-Anthropocene-GratteCiel-docker>

Some demos are provided as dockers to ease their deployment. All the source code is accessible such that the different integration methods can be exploited in other contexts.

6. CONCLUSION

Multimedia documents play an important role in understanding the urban landscape. We have presented several ways to integrate multimedia in a 3D scene in order to bring a better understanding of a territory. We also introduced different user guidance modes, going from a fully guided sequence up to a free consumption of media, than can be combined with our multimedia integration modalities. Different use-cases around real project allowed us to demonstrate the variety of user experiences that can be created with these approaches. The addition of multimedia in the urban digital model brings another dimension to it, and greatly improve the way user can understand and discover a territory.

We currently demonstrated use cases making use of a unique integration approach for all the media in the scene, but the different approaches can be combined in the experience. That opens interesting avenues for future research on how these different modalities can interact with each other and can greatly extend the possibilities of our model.

While our integration approaches can cover various scenario and multimedia types, it might be complicated to find the right integration method for a given use case. It is also not clear yet what are the exact effects on the user experience for each of these modes. Additionally, while our user guidance approach helps the user to find its way in presence of a lot of content, a filtering or grouping system would also prove useful. Being able to assign themes or time period to the different contents would allow the user to filter the different media, but also to group them in the scene and develop new modalities to interact with them. For example, linking the multimedia with a time period would add a fourth dimension to the scene, allowing to travel through time in the scene. The user could then observe the evolution of a district or a building by navigating through documents of different ages.

REFERENCES

- Bleisch, S., Dykes, J., 2015. Quantitative data graphics in 3D desktop-based virtual environments – an evaluation. *International Journal of Digital Earth*, 8(8), 623–639.
- Boutsis, A.-M., Ioannidis, C., Soile, S., 2019. Interactive online visualization of complex 3D geometries. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W9, 173–180.

¹² <https://ecoleanthropocene.universite-lyon.fr/a-quoi-revent-les-maquettes-247850.kjsp?RH=1633680335198>



Figure 10. Three examples of image layers created by students for the slideshow. These layers can be used one after another in a presentation.

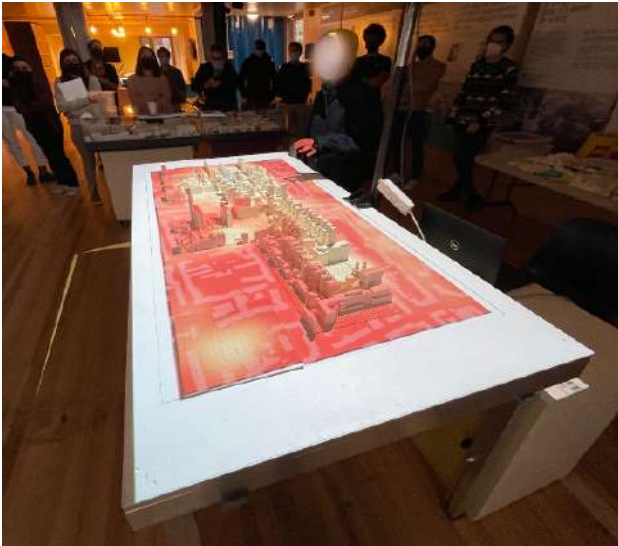


Figure 11. One image of the slideshow projected onto the tangible model.

Wijnants, M., Erum, K., Quax, P., Lamotte, W., 2015. *Web-mediated Augmentation and Interactivity Enhancement of Omni-directional Video in Both 2D and 3D*.

Chagnaud, C., Samuel, J. S., Servigne, S., Gesquière, G., 2016. Visualization of Documented 3D Cities. *The Eurographics Workshop on Urban Data Modelling and Visualisation, UDMV 2016*, Proceedings, Liège, Belgium, 87–93.

Gan, Q., Zhu, M., Li, M., Liang, T., Cao, Y., Zhou, B., 2014. Document visualization: An overview of current research. *Wiley Interdisciplinary Reviews: Computational Statistics*, 6.

Jaillet, V., Rigolle, V., Servigne, S., Samuel, J. S., Gesquière, G., 2021. Integrating multimedia documents and time-evolving 3D city models for web visualization and navigation. *Transactions in GIS*, 25(3), 1419–1438. Publisher: Wiley.

Othman, M. K., Petrie, H., Power, C., 2011. Engaging Visitors in Museums with Technology: Scales for the Measurement of Visitor and Multimedia Guide Experience. *Human-Computer Interaction – INTERACT 2011*, 6949, Springer Berlin Heidelberg, Berlin, Heidelberg, 92–99. Series Title: Lecture Notes in Computer Science.

Samuel, J., Périnaud, C., Gay, G., Servigne, S., Gesquière, G., 2016. Representation and Visualization of Urban Fabric through Historical Documents. E. D. L. ISBN 978-3-03868-011-6 (ed.), *14th Eurographics Workshop on Graphics and Cultural Heritage*, Proceedings of 2016 - Eurographics Workshop on Graphics and Cultural Heritage, Genova, Italy, 157–166.

Seo, D., Yoo, B., Ko, H., 2015. *Webized 3D experience by HTML5 annotation in 3D web*.