This article has been accepted for publication in IEEE Systems Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSYST.2022.3217655

V4Design: Intelligent analysis and integration of multimedia content for Creative Industries

Spyridon Symeonidis, Georgios Meditskos, Stefanos Vrochidis, Konstantinos Avgerinakis, Jens Derdaele, Maarten Vergauwen, Maarten Bassier, Ayman Moghnieh, Luis Fraguada, Verena Vogler, Yash Shekhawat, Leo Wanner, Montserrat Marimon, Kalliopi Valsamidou, Panagiota Koulali, Anastasios Tellios, Jolan Wuyts and Eva Lopez

***Abstract*—Nowadays vast amounts of multimedia content are being produced, archived and digitised, resulting in great troves of data of interest. Examples include user-generated content, such as images, videos, text and audio posted by users on social media and wikis, or content provided through official publishers and distributors, such as digital libraries, organisations and online museums. This digital content can serve as a valuable source of inspiration to the creative industries, such as architecture and gaming, to produce new innovative assets or to enhance and (re-)use existing ones. However, in its current form, this content is difficult to be reused and repurposed due to the lack of appropriate solutions for its retrieval, analysis and integration into the design process. In this paper we present V4Design, a novel framework for the automatic content analysis, linking and seamless transformation of heterogeneous multimedia content to help architects and virtual reality game designers establish innovative value chains and end-user applications. By integrating and intelligently combining state-of-the-art technologies in computer vision, 3D generation, text analysis, generation and semantic integration and interlinking, V4Design provides architects and video game designers with innovative tools to draw inspiration from archive footage and documentaries, inspiring and eventually supporting the design process.**

***Index Terms*—Creative Industries, Virtual Reality, Game Design, Architecture, multimedia analysis**

I. Introduction

CREATIVE INDUSTRIES, such as architecture and gaming, are one of the most growing sectors of the economy and great sources of growth. This sector is characterised by a strong innovation capacity, able to drive innovation in other sectors of the economy, through creativity, design and new organisational processes and business models. Beyond the undoubted impact on social, democratic and cultural wealth, creative industries are increasingly being acknowledged as drivers of innovation and economic transformation [1].

The creative sector is a highly competing and demanding sector, as it requires from designers to be constantly creative and multi-skilled, adapted to the latest technologies and coupled with strong presentational skills. At the same time, vast amounts of new online multimedia content are being generated and archived content is being digitised, resulting in great amounts of data of interest to the architecture and game design communities. This content is currently largely under-exploited, despite its great potential for reuse and repurpose.

As a result, it is very important to provide designers with innovative software that will help to enhance and simplify the designing process. Standard CAD applications such as AutoCAD or Rhino provide 3D modelling tools for digital design and visualisation, but also for importing, editing or adapting already existing 3D models and environments. However, there is a large amount of existing digital multimedia content that could be relevant for design, but is not yet accessible and in the appropriate file format. As such, the reuse and repurposing of digital content is mainly realised based on individual designers skills and a variety of non-interlinked heterogeneous tools.

In this paper, we present an innovative pathway to assist the creative industries in sharing content and maximise its exploitation. The V4Design (https://v4design.eu/) system enhances and simplifies the designing process by integrating technologies and tools for automatic content analysis and seamless transformation. In short, V4Design provides the ability to automatically reuse and repurpose existing visual and textual content from content providers and public web resources. It aims at exploiting state-of-the-art (SoA) digital content analysis techniques to generate 3D models, extract stylistic information from paintings and videos, localise buildings and objects within visual content, and integrate it with textual information so as to inspire and support the design, architecture, as well as 3D and Virtual Reality (VR) game industries. In the following sections, we firstly describe the concept and innovation of V4Design and then elaborate on the key back-end and front-end technologies that compose the proposed platform. Afterwards, we present the promising results of the evaluation, and lastly we conclude the work and briefly outline the future directions towards further improvements.

This work was supported by the EC funded project V4Design (H2020-779962, Corresponding author: Spyridon Symeonidis). We would also like to thank Kriszta Doczy (ArtFilms Ltd), Jesper Wachtmeister (SLRS Multimedia AB), Steffen Riegas (Herzog & de Meuron) for their support.

S. Symeonidis and S. Vrochidis are with the Information Technologies Institute - CERTH, Thessaloniki, Greece.

G. Meditskos is with the Information Technologies Institute - CERTH and the School of Informatics, Aristotle University of Thessaloniki, Greece.

K. Valsamidou, P. Koulali and A. Tellios are with the Department of Architecture and Urban Design, Aristotle University of Thessaloniki, Greece.

K. Avgerinakis is with Catalink Ltd, Nicosia, Cyprus.

J. Derdaele, M. Vergauwen and M. Bassier are with University of Leuven, Ghent, Belgium.

A. Moghnieh, L. Fraguada and V. Vogler are with McNeel Europe SL, Barcelona, Spain.

Y. Shekhawat is with Nurogames GmbH, Cologne, Germany.

L. Wanner and M. Marimon are with Pompeu Fabra University, Barcelona, Spain. L. Wanner is also with ICREA, Barcelona, Spain.

J. Wuyts is with Europeana, The Hague, Netherlands.

E. Lopez is with Deutsche Welle, Bonn, Germany.

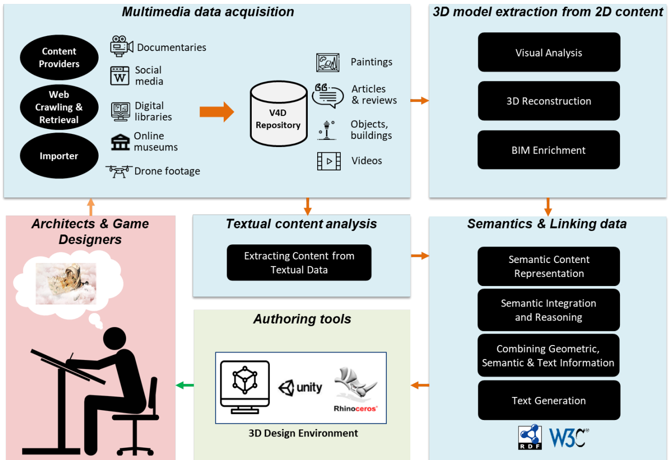

Fig. 1. The V4Design concept.

II. V4Design Concept and Innovation

The V4Design platform is a rich and sophisticated tool that facilitates designers throughout the whole design procedure, from the creation of an initial concept to a final design proposal. To achieve its goals, several different modules are combined and efficiently collaborating to create the content that reinforces the design process. The V4Design process is implemented as a user-oriented service, taking into account its ramifications and its different configurations.

The innovation centres on the functions related to the extraction of potential assets from media content, and their seamless integration in newly authored projects. Beyond the advances in technology that V4Design affords, the true added value lies in how it aids users to find, extract, and reuse complex assets that otherwise need to be rebuilt from scratch. This process by itself is highly innovative as currently no automated support for such exists in commercial tools or experimental frameworks related to architecture and game design.

Several tools in the market offer 3D reconstruction functionalities, such as 3DFlow Zephyr and Reality Capture. Compared to V4Design, these solutions do not offer many capabilities of content reuse and repurpose. The differentiating factor for V4Design is the data collection and analysis processes, along with the sophisticated solutions for semantically representing, aggregating and combining annotations coming from visual and textual analysis of digital content. To this end, apart from innovative 3D model reconstruction techniques, V4Design integrates computer vision and text analysis solutions to further process the content and extract annotations and dynamically enrich the generated 3D models with information, such as façade elements, named entities, opinions from online reviews and critiques. All the available information is semantically interlinked into rich knowledge graphs, promoting interoperability through established W3C standards, while text generation is used to create multilingual summaries of the assets and assist end users in consuming the information. Additionally, V4Design has been developed as a Rhino (https://www.rhino3d.com/) and Unity (https://unity.com/developer-tools) plugin, allowing users to create models directly in well-known design environments.

Figure 1 illustrates the V4Design concept. The processing pipeline starts with data acquisition, which retrieves freely available multimedia data (textual and visual) from several different sources. The textual content is fed to the text analysis module, whereas the visual content feeds the visual analysis and 3D reconstruction services. Next, knowledge graphs and semantic models are employed to properly capture and interlink the information, to generate natural language descriptions and to compile the semantically enriched 3D assets that are shared with the end users. The front-end of the V4Design platform consists of two authoring tools, one for architects and one for virtual-reality game designers. The architecture authoring tool is based on Rhino, a 3D CAD application and open development platform for creating plugins and custom components. The game tool is developed on top of Unity, a real-time development platform for creating 2D and 3D multiplatform games and virtual environments.

III. Content Analysis and Interlinking

A. Data acquisition

While numerous solutions have been proposed for standalone data collection tasks (e.g. Scrapy Python library for web crawling/scraping), none of them provides a complete system that can be applied on multiple and heterogeneous web resources. The data collection component of V4Design is a fully-fledged solution that gathers online data from various web resources. It is the starting point of a processing pipeline that generates the 3D assets, which are presented in the front-end tools. Several techniques are combined to accommodate for the different types of resources, such as web scraping and social media search. To reinforce the user-driven approach and to enable dynamic initialisation of the component, V4Design platform also contains the Importer, which receives data acquisition requests from users, and is also able to download videos and image collections from the user's computer to the platform's services, to be processed as raw data.

B. Content extraction from visual data

1) Localisation and texturing: The analysis of the visual content (video/images) is handled by the Spatio-Temporal Building Localisation (STBL) and the Spatio-Temporal Object Localisation (STOL) modules based on the output of the Scene Recognition (SR) module. SR extracts information about the scene depicted in the analysed image or video (e.g. village, skyscraper) and characterises the scene as "indoor" or "outdoor". Videos or images that are characterised as "outdoor" are further analysed by the STBL module, where buildings and building façade elements (e.g., walls, windows) are localised. The "indoor" videos and images are further analysed by the STOL module that detects interior objects such as table, chair, etc. For SR, STOL and STBL, SoA deep learning models were trained and extended, using datasets specialised on the V4Design's scope. The Texturing component applies a style from a user selected image (such as a painting or a building of a specific architecture) onto an automatically generated 3D model's texture, thereby creating a new textured asset. It is

This article has been accepted for publication in IEEE Systems Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSYST.2022.3217655

based on Cycle-Consistent Adversarial Networks and achieves better qualitative and quantitative results than SoA [2].

2) 3D Reconstruction: Automatic generation of 3D models using photogrammetry is handled by the 3D reconstruction (3DR) module [3]. Video and image files, as well as output from STOL, STBL, the shot detection and texturing modules provide important input to the 3DR task. Since the majority of the data we process is not captured for future photogrammetric use, it is important to implement mechanisms to filter bad/unsuitable data. For the reconstruction of videos, this encompasses the shot detection and frame extraction modules. Based on sum of absolute differences and Geometric Robust Information Criterion techniques, the most suitable shot and frames are automatically extracted from an input video, something that is nowhere employed in the existing commercial solutions. For the processing of image collections, a robust way of clustering images depicting the same scene is conducted with the use of a vocabulary matching strategy. Finally, the module conducts the ultimate steps of restyling the 3D model based on the outcome of the Texturing process.

C. Text analysis and generation

Content extraction from textual data is carried out by the text analysis (TA) component that analyses the textual material that is associated with or related to the assets of the platform [4]. The material originates from a variety of sources: (i) image captions and video descriptions; (ii) Wikipedia articles; and (iii) opinions conveyed in reviews and critiques from the Web. TA improves the concept extraction F1-score by nearly 10% compared to the strongest competitor (DBPedia Spotlight).

The multilingual text generation component is based on the grammar-driven FORGe generator [5], which has been adapted to create rich textual descriptions on the assets generated by the platform. In particular, content catered by four components is verbalised: visuals analysis, text analysis, opinion analysis and DBpedia [6]. The content reflects 45 architectural features (e.g. type of landmark, style, creator, etc.) and five metadata properties. The descriptions are generated in English, Spanish, Greek and German, with different levels of coverage.

D. Semantic integration

Semantic integration enriches V4Design with a semantic annotation layer by providing the appropriate knowledge structures to represent the data and store them into the Knowledge Base. The module also offers semantic retrieval capabilities by executing semantic queries, as well as a rule-based semantic reasoning module that extracts useful inferences. A unified model has been designed as an OWL 2 ontology (https://www.w3.org/TR/owl2-overview/), capturing V4Design information, using the Web Annotation Data Model, in order to promote interoperability and reusability [7] [8]. Existing schemata and ontologies, such as Building Topology Ontology (https://w3c-lbd-cg.github.io/bot/), have also been used to inherit useful classes and relationships. Using Linked Open Data, we assist in making richer descriptions (by serving the appropriate information to Text Generation module) and offering an intelligent conceptually based semantic search to the users (accessible through the V4Design tools).

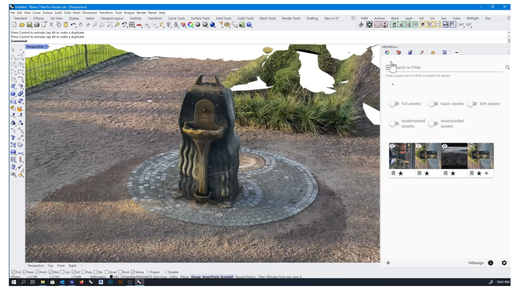

Fig. 2. Architecture authoring tool (Rhino plugin). On the left side of the figure, the UI of the Rhino Authoring tool is shown.

IV. AUTHORING TOOLS

One of the key technological achievements of V4Design is the implementation of two authoring tools that exploit and integrate novel content analysis techniques to generate 3D models, extract stylistic information from paintings and videos, localise buildings and objects of interest within visual content, and integrate it with textual information.

A. Architecture Design Tool

The Architecture Authoring tool serves as a content and process management environment through which users can access the V4Design platform and retrieve assets from its storage. It is a mature application developed on top of Rhino. It connects to the backend and implements data search, retrieval, download, import and export functions, including a seamless integration of downloaded assets with the Rhino editor. In addition, it facilitates the discovery and retrieval of 3D models for repurposing, and allows users to access raw data (shots and image clusters, with original text). The tool is published as a Rhino package and can be downloaded and installed from Rhino using the Package Manager. Figure 2 shows the architecture design environment in Rhino.

B. Game Design Tool

The Game Design tool is composed of: a) the Unity plugin, b) the VR tool. Each application contributes in a different way to the fast and efficient creation of VR games and experiences.

The Unity Plugin (Figure 3) has been developed for the Unity Game Engine and enables users to download assets from V4Design repository, upload videos for reconstruction, rate assets and use a set of prefabricated gamification templates (triggerables) that encourage and allow various forms of interaction. The VR tool enables the creation of new environments directly from inside the virtual world. It supports a subset of the functionalities of the Unity plugin and is implemented to help the developers import and place 3D models and recreate the environment on-the-go so as to enhance the final experience for the user. This application imports assets and questions from the V4Design databases and is compatible with any Steam VR compatible device, such as Oculus Rift.

V. V4DESIGN EVALUATION

In the initial prototypes of the platform formative usability testing was performed, whose goal was to support developers

This article has been accepted for publication in IEEE Systems Journal. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/JSYST.2022.3217655

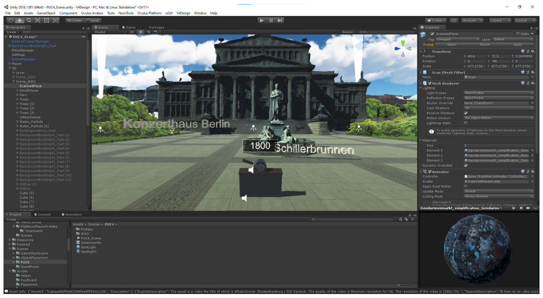

Fig. 3. Game design authoring tool (Unity plugin).

in further improving the system. Contrastingly, the evaluation of the final system followed the rules of summative testing; the integrated platform was assessed through well-defined measures i.e. effectiveness, efficiency and satisfaction. Effectiveness is defined as the extent to which the users can fulfil a task and achieve their goals, according to metrics such as the number of tasks performed and errors per unit of time. Efficiency depends on how users' effort relates to the accuracy and completeness of the results, e.g. in terms of the time spent to perform a particular task or to correct errors. Satisfaction expresses how satisfied the users are by working with the system, for example if the users feel "in-control" of it, or if they rate the system as "easier to use" compared to others.

For the final platform evaluation three workshops have been organised, where participants were performing design scenarios using V4Design and providing their feedback via questionnaires. The Aristotle University of Thessaloniki conducted an academic evaluation of the V4Design Rhino plugin based on architecture-oriented scenarios, with 20 participants. Herzog & de Meuron and McNeel conducted an evaluation/demonstration event for the Rhino plugin with 20 participants. For the evaluation of the Unity plugin, Nurogames with Deutsche Welle held an open day mostly with academic users, where in total 8 participants provided a valuable assessment of the platform in the game design scenarios.

In Figure 4, we present the aggregated results of the entire system evaluation. The values represent the percentage of the averaged Likert scores of all the answers (from every user, where values range from 1 to 5 and 5 shows the highest degree of approval), for the set of questions belonging to each of the aforementioned evaluation metrics. Apart from the system evaluation, the users assessed the quality of specific components and functionalities, in terms of satisfaction. Table I shows the mean Likert scores of the tested features.

Overall, the evaluation showed satisfactory results. Assessments indicate that the GUI and the interface of the different modules were considered as user-friendly and easy-to-use by users. Testers deemed V4Design as easier to use in comparison to its competitors and expressed that they quickly familiarised with it and are willing to use it in the future.

VI. Conclusions

In this work we described V4Design, a novel platform conceived for architects and game designers that is able to exploit the abundance of existing multimedia content towards automatically generating 3D assets and reinforcing the design process. The creation of these assets involves a sophisticated pipeline that acquires online data, reconstructs 3D models, and develops elaborate metadata, among other tasks. Authoring tools have been developed on top of well-known software to visualise the results and help end-users remain creative, innovative and competitive. Future advances on the underlying state-of-the-art technologies would further extend the impact of the proposed system in the creative industries.

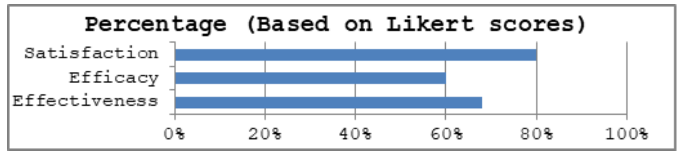

Fig. 4. Summary of V4Design system evaluation results

TABLE I
Satisfaction score per component

Component / Feature	Mean score
3D Reconstruction	3.6
Style transfer	3.4
Texture transfer	3.7
Textual summaries	3.2
Search functionalities	3.1
Create Asset-import video	3.9
Asset rating functionality	3.8
3D asset quality	3.0
Texture quality	3.4
Asset store access	3.8
Applying assets	4.0

References

[1] R. Boix-Domènech and P. Rausell-Köster, *The Economic Impact of the Creative Industry in the European Union*. Cham: Springer International Publishing, 2018, pp. 19–36.

[2] E. Batziou, P. Alvanitopoulos, K. Ioannidis, I. Patras, S. Vrochidis, and I. Kompatsiaris, "Cycle-consistent adversarial networks and fast adaptive bi-dimensional empirical mode decomposition for style transfer," in *2020 25th International Conference on Pattern Recognition (ICPR)*. IEEE, 2021, pp. 2360–2367.

[3] J. Derdaele, Y. Shekhawat, and M. Vergauwen, "Exploring past and present: Vr reconstruction of the berlin gendarmenmarkt," in *International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 2018, pp. 287–292.

[4] A. Shvets and L. Wanner, "Concept extraction using pointer–generator networks and distant supervision for data augmentation," in *International Conference on Knowledge Engineering and Knowledge Management*. Springer, 2020, pp. 120–135.

[5] S. Mille, S. Dasiopoulou, and L. Wanner, "A portable grammar-based nlg system for verbalization of structured data," in *ACM SAC 2019, Special Track on Knowledge and Language Processing*, 2019, pp. 1054–1056.

[6] S. Mille, S. Symeonidis, M. Rousi, M. M. Felipe, K. Stavrothanasopoulos, P. Alvanitopoulos, R. Carlini, J. Grivolla, G. Meditskos, S. Vrochidis, and L. Wanner, "A case study of nlg from multimedia data sources: Generating architectural landmark descriptions," in *Proceedings of the 3rd International Workshop on Natural Language Generation from the Semantic Web (WebNLG+)*., 2020, pp. 2–14.

[7] G. Meditskos, S. Vrochidis, and I. Kompatsiaris, "V4ann: Representation and interlinking of atom-based annotations of digital content," in *International Conference on Semantic Systems*. Springer, 2019, pp. 124–139.

[8] M. Rousi, G. Meditskos, S. Vrochidis, and I. Kompatsiaris, "Supporting the discovery and reuse of digital content in creative industries using linked data," in *2021 IEEE 15th International Conference on Semantic Computing (ICSC)*. IEEE, 2021, pp. 100–103.