Question-Answering Approach to 3D Information Retrieval System

Taweesak Sangkapreecha [1] and Pataraporn Sangkapreecha [2*]

[1] School of Architecture, Bangkok University

[2] School of Communication Arts, Bangkok University

[*] Corresponding author. E-mail address: pataraporn.s@bu.ac.th

Abstract

During the last decade, a common practice in architectural and interior design has been involving the creating of digital three-dimensional (3D) models during a design process. While a digital modelling of design artefacts is much easier and faster than before, managing and retrieving of 3D model information is still a challenging task. This study aims to review ongoing researches in 3D information retrieval systems and to highlight needs for developing of better methods to facilitate a 3D model retrieval and extraction process. A systematic review method was employed to identify, appraise and synthesise all the empirical evidences. This examine relied upon qualitative analysis of documentation from the literature and ongoing researches. The results reveal that the number of researches acknowledges that current 3D information retrieval systems still have drawbacks on a precise interpretation of what a user needs. This study thus suggests that the current systems can greatly benefit if they take advantages that a question-answering search system offers. The proposal model of a feasible architecture is presented in the discussions section. This model extends a common 3D information retrieval system by utilising a general architecture of question-answering system.

Keywords: Question-Answering Approach, 3D Information Retrieval System, Architectural and Interior Design, Systematic Literature Review

Introduction

Architects and interior designers have extensively progressed their work from only producing of two-dimensional (2D) drawings to generating of digital three-dimensional (3D) models in the last decades. A recent practice has shown that these groups of designers have been exploiting digital 3D models for various purposes. Particularly in a design visualisation area, there has been an increase in the number of using digital 3D images and models to present design idea to clients. With the versatility of current 3D modelling and editing tools, producing of digital 3D models is much easier and faster than before. As a result, a number of digital 3D models and components of design projects are rapidly growing. Management and access to those large repositories has become increasingly complex and time consume. This study thus aims to review ongoing researches in 3D information retrieval systems and to highlight the needs for developing of better methods to facilitate a 3D model retrieval and extraction process.

Three-dimensional (3D) information model retrieval is a research area that has been rapidly expanding in recent years. The retrieval aims at developing of techniques that facilitates effective indexing, searching, and browsing of 3D models, mostly based on automatically derived model vector features. In general, a 3D information retrieval system can be categorised according to its query formats: keyword-based and content-based query. A keyword-based query approach is based on the notion of indexing keywords description of 3D model contents, which are either created by a human input or derived from low-level features of a 3D model.

Mezaris, Kompatsiaris, & Strintzis. (2003) described a framework for mapping low-level features to intermediate-level descriptors, which then form a vocabulary and term an object ontology. The ontology was used to allow the qualitative definition of high-level concepts that a user may use as a query. Min, Kazhdan, & Funkhouser (2004, pp. 209-220) proposed a text matching for 3D model approach by indexing spatial models linked from a web page, and extracting textual information from both a model file and contents on the page. They also created an additional text source by adding synonyms and hyponyms of a filename using the WordNet.

Manual annotation, however, for a large set of spatial 3D information is too limited or ambiguous depending on many factors: languages, culture, age, etc. In many cases, a keyword-based approach is likely to fail to retrieve effectively, such as when 3D models are not annotated at all or are described with unspecific keywords, when all keywords are too common terms, when relevant keywords are ambiguous to a user, or when keywords of interest are not known for an annotator. These issues motivate research on an area of content-based 3D information model retrieval, which attempts to retrieve 3D information by a model example. The ultimate goal of a content-based approach is to develop techniques that support effective indexing, clustering, and searching of 3D models based on automatically derived model vector features.

There is a trend towards tight integration of a keyword-based approach with a content-based approach. In general, the integration is archived by combination of a textual keyword and a 3D model example to be entered in a single multimodal query. This method can improve a search that focuses on specific subclasses of 3D models. The complete integration of keyword-based and content-based approaches can provide orthogonal ideas of similarity corresponding to both function and form of 3D models (Min, et al., 2004, pp. 209-220).

Method

Understanding the past to guide the future of a 3D model retrieval and extraction is the aim of this research study. The purpose of this study is to uncover the past, define the present, and suggest the future implementation of a 3D model retrieval system for assisting in architectural and interior design. In essence, a systematic review method was employed to identify, appraise and synthesise all the empirical evidences. The study relied upon qualitative analysis of documentation from the literature and ongoing researches, to indicate a feasible architecture of the system.

Initially, four online academic research databases, ACM Digital Library, IEEE Xplore, Springer Link and Science Direct, were scanned for relevant articles. ACM and IEEE were chosen as they contain conference proceedings from major conferences on human-computer interaction, information systems, and computer aided design. Springer Link and Science Direct cover several important journals from various scholarly fields. In order to increase the reliability of the search results and ensure that articles from various scholarly fields will be included, the search was repeated with Google Scholar. All searches were narrowed down to empirical studies reported in peer-reviewed full conference papers and journal articles. The phrases 'Database indexing and retrieval, '3D model retrieval', '3D information retrieval system' and 'Question-Answering system' were used in the searches in order to find articles that discuss 3D information retrieval systems in some manner. As a result, a total of 56 articles meeting the inclusion criteria were selected for the review. The selected articles were published in the years 1995-2015.

Results

Representations of 3D Information Model

Regarding to a representation of 3D model contents, it can be defined as a formal scheme for describing a 3D model or some aspects of the model altogether with rules that specify how the scheme is applied to the particular model. A typical content-based 3D model retrieval system treats a query of 2D/3D shapes or 3D models in a repository as a set of feature vectors, and then ranks the relevance between both in proportion to a similarity measure, which afterwards is calculated from the feature vectors. In other words, these vectors identify contents of a 3D model.

The feature vectors of a 3D model can be classified into two categories; global features, and local features. The first category includes statistical moments of a boundary or a volume of a 3D model (Zhang, & Chen, 2001), volume-to-surface ratio (Corney, et al., 2002, pp. 65-74), a shape histogram (Ohbuchi, Otagiri, Ibato, & Takei, 2002, pp. 265-274), spherical harmonics (Kazhdan, Funkhouser, & Rusinkiewicz, 2003, pp. 156-164), a voxelised model (Novotni, & Klein, 2003, pp. 216-255). Since global features characterise the overall characteristic of a 3D model, the methods, mentioned above, are not very discriminative on the details of a 3D model. The second category is a 3D shape spectrum (Zaharia, & Preteux, 2001, pp. 133-145), 3D shape contexts (Kortgen, Park, Novotni, & Klein, 2003). These methods exploit various approaches that take into an account of a surface within nearby points on a boundary of a 3D model. While the first category evaluates only the pure geometry of a 3D model, in contrast, the last category uses a number of graph-based methods to extract a geometric meaning from linked components of a 3D model. These graph-based methods include model graphs (El-Mehalawi, & Miller, 2003, pp. 83-94), spectral graphs (McWherter, & Regli, 2001), Reeb graphs (Hilaga, Shinagawa, Kohmura, & Kunii, 2001, pp. 203-212), skeletons graphs (Sundar, Silver, Gagvani, & Dickenson, 2003a). While these methods broadly support for retrieving of 3D models with a particular shape, yet they may include the retrieved result with other 3D models having different functions and meanings.

Table 1 Two categories of feature vectors methods

Methods of defining the feature vectors of a 3D model (for query)	
Defining global features	Defining local features
1.Statistical moments of a boundary	1.3D shape spectrum
2.Volume-to-surface ratio	2.3D shape contexts
3.Shape histogram	3.Model graphs
4.Spherical harmonics	4.Spectral graphs
5.Voxelised model	5.Reeb graphs
	6.Skeletons graphs

Unlike exiting approaches which use either global or local features, Wang, Lin, & Tang (2014, pp. 128-139) introduced 3D model retrieval approach which utilises both global feature-based and local feature-based techniques in 2D hybrid views of 3D models as well as the query sketch. In this work, global features are extracted and represented as integrated descriptors while local features are extracted by adopting an improved bag-of-features method and represented as distributions of compact visual words. Although Wang, et al. (2014, pp. 128-139)'s proposed approach achieves better performance than previous approaches, sketch-based 3D model retrieval for realistic inputs is still a very hard problem.

A similarity measure is a key issue in content-based 3D model retrieval. It quantifies the similarity in contents between a pair of 3D models by computing distances between the pairs using a dissimilarity measure. Before performing a dissimilarity measure, the representation of a 3D model must be converted into a feature vector or a relational data structure (Iyer, Jayanti, Lou, Kalyanaraman, & Ramani, 2005, pp. 509-530).

In a feature vector conversion approach, feature vectors are represented as points in feature space in a database. Thus, the similarity between two vectors corresponds to the distance between the points in the space. The Minkowski Distance is commonly used to compute a distance between two points (Iyer, et al., 2005, pp. 509-530). However, when having two noisy corresponding points, a dissimilarity measure— like the Hausdorff distance (Veltkamp, 2001, pp. 188-197), is often used instead. For polygons and polylines, the Cumulative Angle Function or Turning Function can be used to match either the whole length of polylines (Arkin, Chew, Huttenlocher, Kedem, & Mitchel, 1991, pp. 209-215), or a partially match— in cases that the curves of polylines do not have the same length (Vleugels, & Veltkamp, 1999, pp. 575-584). Another technique to find a dissimilarity match of a part of curves is the Fréchet Distance (Alt, & Godeau, 1995, pp. 75-91), which works by continuous measuring the distance along corresponding curves to find the maximum length within these measured distances. To find the similarity of corresponding patterns, the Reflection Metric (Hagedoorn, & Veltkamp, 1999) can be used to define finite unions of curves on a surface in the feature space. Fry (1993) presents his Transport Distances as a metric for computing and matching of 3D model contours. This technique involves comparing an unknown edge to a set of standard edges, and deciding of which known edge has matched. In a relational data structure approach, it bases similarity metrics on relational data structures such as graphs. There are various approaches to use graph matching methods for comparing structures of 3D models; including Attributed Graphs (El-Mehalawi, & Miller, 2003, 95-105), Model Signature Graphs (McWherter, & Regli, 2001), frequency histograms (Peabody, & Regli, 2001; Lu, He, & Xue, 2009, pp. 495-499), Skeletal Graphs (Sundar, Silver, Gagvani, & Dickenson, 2003b, pp. 130-139). For a topology matching, the similarity between nodes is calculated based on the similarity of the node attributes with a support of the Multi-resolution Reeb Graphs (Hilaga, et al., 2001, pp. 203-212; Barra, & Biasotti, 2013, pp. 2985-2999)

Table 2 Example approaches and techniques used to find a similarity between 3D models

Approaches to a similarity measure (for ranking)	
converting into feature vectors	converting to a relational data structure
1.Minkowski Distance	1.Attributed Graphs
2.Hausdorff distance	2.Model Signature Graphs
3.Cumulative Angle Function	3.Frequency histograms
4.Fréchet Distance	4.Skeletal Graphs
5.Reflection Metric	5.Multi-resolution Reeb Graphs
6.Transport Distances	

Recently, Lu, Wang, Xue, & Pan (2014, pp. 703-713) proposed a combined 3D model distance measure that uses disjoint information. In their method, feature histograms are employed as the 3D model descriptor, and the disjoint information between two feature histograms is calculated to measure the distance between the two 3D models. However, the features used in their work were mainly based on the Euclidean distance in which when dealing with non-rigid deformed shapes, these features are no longer valid. Xiao, Feng, Ji, & Zhuang (2014) introduce the model to reduce the computational cost when dealing with the newcome 3D models. They proposed a fast view-based 3D model retrieval framework comprising an Unsupervised Multiple Feature Fusion (UMFF) algorithm and Online Projection Learning (OPL) algorithm. UMFF exploits the complementary information between different visual feature and hold the embedded data geometric structure information. Yet, their approach is an unsupervised learning algorithm that ignores the labels of 3D models. Unlike the most exiting view-based 3D object retrieval methods that use one single feature to describe a 3D object, Zhao, Yao, Zhang, Wang, & Liu (2015, pp. 110-118) proposed a feature fusion method via multi-modal graph learning for view-based 3D object retrieval. Different visual features, including 2D Zernike moments, 2D Fourier descriptor and 2D Krawtchouk moments, were extracted to describe each view of a 3D object. The Hausdorff distance was computed to measure the similarity between two 3D objects with multiple views. Their method based on multi-modal graph learning greatly out performs the methods using one single feature, because the complementation of different features are sufficiently explored.

Semantic Gap in 3D Representations

The similarity measure only captures certain facets of a 3D model; it excludes actual meaning of a 3D model - semantics of the model. This issue is regarded as a semantic gap, which can refer to differences between system-perceived and user-perceived similarities of 3D models. Many methods have been proposed to reduce a semantic gap. They generally fall into two categories depending on the level of user involvement in a retrieval process— a relevance feedback and a statistical classification method.

A relevance feedback method is an iterative and interactive refinement of the original formulation of a query. The essence of a relevance feedback is that it allows a user to interact with a retrieval process by providing additional information that the user thinks to be relevant to the query. According to a user feedback, the model of a similarity measure is dynamically adapted to give an approximation of a user's perception subjectivity. Three strengths of a relevance feedback are discussed in Baez-Yates, & Ribeiro-Neto (1999), which include: (1) It shields a user from the inner details of a retrieval system, (2) It brings down a retrieval task to small steps which are easier to grasp, (3) It provides a controlled setting to emphasise some features and de-emphasise others. Though a relevance feedback has been an active topic of research in text-based retrieval and in content-based image retrieval, it has hardly been explored in 3D model retrieval. There are works that combine a relevance feedback with the Support Vector Machine (SVM) (Elad, Tal, & Ar, 2001, pp. 97-108). Bang, & Chen (2002) propose a new approach to a relevance feedback by using the feature space warping, which can adjust data points of a 3D model in a controlled manner responding to a user feedback. A method combining the supervised and unsupervised feature extraction is presented in Leifman, Meir, & Tal (2005, pp. 865-875). While using a relevance feedback has shown considerable improvements in retrieval precision, user interaction has remained a main concern in information retrieval research. A main concern is that a relevance feedback may add burden to a user, especially when a user has to provide more and more information to filter a non-relevant result from the retrieval.

A statistical classification technique is a method for training for each semantic concept to semantically index a 3D model. A number of 3D models are grouped into semantically meaningful clusters by using low-level visual features, which allow a user to conduct searching in semantically adaptive manner. Zhang, & Chen (2001) propose a statistical model by using the active learning, an approach typically used in a machine learning system, to improve efficiency of hidden annotations in a content-based 3D retrieval system. Hou, Lou, & Ramani. (2005, pp. 155-164) use the Support Vector Machine based clustering to associate a user's query with the system's semantic concept, thus the search engine only performs a search within corresponding clusters and applies semantic concept to the retrieval by using those clusters. There has been work on attaching conceptual labels to a 3D model. Bustos, Keim, Saupe, Schreck, & Vranić (2005, pp. 345-387), for instance, introduces the Taxonomy based method in his feature extraction, which labels a 3D model according to its categories, names and concepts.

Many view-based 3D object retrieval methods, such as Elevation Descriptor (ED) (Shin, Lee, & Wang, 2007, pp. 283-295), Bag-of-Visual-Features (BoVF) (Ohbuchi, & Shimizu, 2008, pp. 411-418; Eitz, Hildebrand, Boubekeur, & Alexa, 2010; Bronstein, Bronstein, Guibas, & Ovsjanikov, 2011; Lavoué, 2012, pp. 931-942), Spatial Structure Circular Descriptor (SSCD) (Gao, Dai, & Zhang, 2010, pp. 1142-1151) have been proposed these years. For 3D model matching, the 3D model comparison is based on the comparison of features from different views. Gao, et al. (2010, pp. 1142-1151) proposed a SSCD method to compare two 3D models. The SSCD can preserve the global spatial structure of 3D models, and it was in variant to rotation and scaling. All spatial information of 3D model can be represented by an SSCD which included several SSCD images. But they mainly focused on employing these methods in various ways for supporting 3D model query inputs.

There are two main stages for 3D model retrieval, including 3D model representation and 3D model matching. Most of existing works focused on 3D model representation methods. For 3D model matching, the comparison between two groups of 2D views is modeled as a weighted bipartite graph matching (Gao, & et al., 2010, pp. 1142-1151). The weighted bipartite graph is built with the selected representative 2D views, and the proportional max-weighted bipartite matching method is employed to find the best match in the weighted bipartite graph. However, the relationship among views from different models are neglected. Other methods such as Gao, Wang, Tao, Ji & Dai (2012, pp. 4290-4303)'s work involve a learning process.3D model retrieval can be considered to find the 3D model which can reconstruct the query views most accurately. Thus, it can be modeled as a reconstruction task based on sparse coding. In Wang, & Nie (2015)'s work, a view-based 3D model retrieval algorithm based on weighted locality-constrained group sparse coding method was proposed. It employed to obtain the group based representation of query object, which captures the inter-group and intra-group relation to some extent. Then, the reconstruction residuals are combined according to the weights of query views.

Bridging the Semantic Gap: a Question-Answering Based Approach

Using natural language interface to access relational databases is not new and can be traced back to several decades ago. A Question Answering (QA) system draws on concepts and results from many fields, including natural language processing, information retrieval, and human computer interaction. The field of a QA system is concerned with a building computer system that understands a user's query in a form of natural language, and uses the query to retrieve information needed from a set of documents, which can then return the extracted parts of the information to a user. For more information about the early natural language systems and the discussion of a recent state of Question Answering (QA) research, see Hirschmann, & Gaizauskas (2001, pp. 275-300)

Many researchers view the modern QA task as a combination of two natural language processing tasks; namely an information retrieval (IR) and an information extraction (IE). An information retrieval (IR) system allows a user to access a collection of documents based on its semantic content (Baeza-Yates, & Ribeiro-Neto, 1999). Two main functions of an IR system are indexing and retrieval (Chowdhury, 2004). The former function responds to the extraction and storage of information that are most representative to the semantic content of a document. This extracted information is referred as an indexing term, which may include isolated keywords, and the relations between the keywords. The latter function reacts to a user's information query by matching the query with its information indices, built by an indexing function. If the retrieval system evaluates documents to be relevant to a user's query, then a matching function retrieves a set of the documents and presents them to the user. The typical performance measures of information retrieve system are precision and recall (Rijsbergen, 1979). The precision is a fraction of retrieved documents that is relevant. The recall is a fraction of relevant documents from the retrieved collection. The precision and the recall are real-valued variables in the interval [0,1]. An ideal system should turn 1-value for both measurements, which indicates that it returns all of the relevant documents and only the relevant documents.

An information extraction (IE) system is a method for filtering information from large volumes of documents. Grishman (1997) defines an IE system as the identification of instances of a particular class of events or relationships in a natural language document, and the extraction of the relevant arguments of the event or relationship. Information extraction involves extracting facts of different granularity, ranging from named entities, entity relations to complete event. It reduces a large collection or information that is relevant to facts into a structured representation. The structured representation can also be referred as a template. Each template contains a number of slots that an IE system would have to fill. The Named Entity recognition is a specialised module of an IE task (Bikel, Schwartz, & Weischedel, 1999, pp. 211-231). It can be used to identify phrases in documents that refer to entities like places, or persons, and extract their semantics. In a QA system, an IE system can be supported with a detailed question analysis, which helps the system to understand a type of an entity that it is looking for.

Integrating of textual and visual information in a QA system

New alternative research directions in a QA system have been explored in the last few years.

These works attempt to go beyond a factoid question of a conventional QA system- such as who, what, when and where questions. A multimedia question-answering system is one of those new approaches. The recent developing of multimedia QA system has been focusing on video. For example, Yang, Chaisorn, Zhao, Neo, & Chua (2003) propose a news video question answering system that returns the relevant news video fragments as the answers, supplemented by the text version of latest news, summarized to the duration constraint as specified by a user. Cao, Robles-Flores, Roussinov, & Nunamaker, (2005) developed a video QA system that captures semantic contents by converting them from soundtracks into transcripts. The transcripts are associated with the video segments, which support automated question answering by natural-language-based and pattern-based approaches. Unlike others, in which the textual information derives from video transcripts, STOP (Katz, Lin, Stauffer, & Grimson, 2004, pp. 113-124) uses an object recognition technology to isolate the moving objects from a stationary background scenery, and extract high-level events from the video input to generate symbolic representations that capture the semantics of those events.

Developing the shared data representation between textual and visual information is apparently the first step to solve an inter-media barrier problem. In the field of a computer vision, two types of representations are widely used. One is a geometric representation in which data items are mapped to vector space. Another is a graph representation, which focuses on the relationship of items. On the other hand, in a linguistic domain, the representation that is commonly used is a symbolic representation, such as a component of a lexical conceptual structure. In general content-based information retrieval, various statistics models have also been applied to join the textual and visual information. Scarlo, Cascia, Sethi, & Taycher (1999, pp. 86-98) propose a method to capture textual statistics in a vector form using latent semantic indexing. The latent semantic indexing vector and the visual feature vector are then combined into unified vectors that allow for a content-based search of an image database. Westerveld, de Vries, van Ballegooij, de Jong, & Hiemstra. (2003, pp. 186-198) employ probabilistic models based on the Bayesian decision theory to combine both textual and image features, and then represent them in terms of one uniform framework. The output from various feature extraction is represented as a combination of two similar models: one for textual features based on language models for text and speech retrieval, and the other for image features based on a mixture of the Gaussian densities. In a question answering system, Katz, et al. (2004, pp. 113-124) develop a symbolic structure that bridges textual and video representations. The Video-surveillance QA system, called the Spot, uses a motion tracking technology that sends video data to an event recognition module, which then converts this input into a symbolic structure. Thus, a user's query in a form of a natural language question can be translated and matched against recognised events on a symbolic representation layer.

Table 3 Comparing information representation between a textual QA system, a 3D information retrieval system, and a multimedia QA system

Comparing Methods of information representation		
Textual QA system	3D information retrieval system	Multimedia QA system
1.Lexical conceptual structure	1.Feature vectors space	1.Latent semantic indexing and Feature vector
2.Latent semantic	2.Graph structure	2.Probabilistic Bayesian decision model

In recent years, Chali, Hasan, & Joty. (2011, pp. 843-855) presented the impact of syntactic and semantic information in the graph-based random walk method for answering complex questions. They applied tree kernel functions to perform the similarity measures between sentences in the random walk framework. They also extended their work further to incorporate the Extended String Subsequence Kernel (ESSK) to perform the task in a similar manner. Wu, et al. (2012, pp. 135-143) proposed a graphical multimedia query language called GMQL, which is developed based on a semi-structured data organisation model. This method combines the advantages of graphs and texts, making the query language much clear, easy to use and with powerful expressiveness. Dwivedi, & Singh (2013, pp. 417-424) proposed taxonomy for characterising QA systems. They suggested that question answering techniques, based on linguistic approach, statistical approach and pattern based approach will continue to remain in sharp focus, receiving attention of a large number of Question Answering System researchers.

Discussion

3D Information Model Meets a QA System

From the intensive review on researches in a 3D information retrieval system and a QA system in this study, it is undeniable that the current 3D information retrieval system can substantially benefit from the state of art that a QA system offers. To combine both systems, this study draws a conclusion by proposing a model of a feasible architecture to extend a common 3D information retrieval system. The proposed model applies a similar approach to the work of Scarlo, et al. (1999, pp. 86-98). The proposed model contains a system that is able to pre-compute the visual statistics of a 3D model in collections based on the low-level feature vector extraction. This includes both a global and a local feature vector; although a set of certain feature analysis methods have to be further defined. Later, the textual descriptions of a 3D model and the annotations of its components are converted into vectors based on the latent semantic indexing similarity to the model proposed by Yang, et al. (2003) and Cao, et al. (2005). After that, the latent semantic indexing is combined and assigned to form descriptors, which later are compared with user's query vector space.

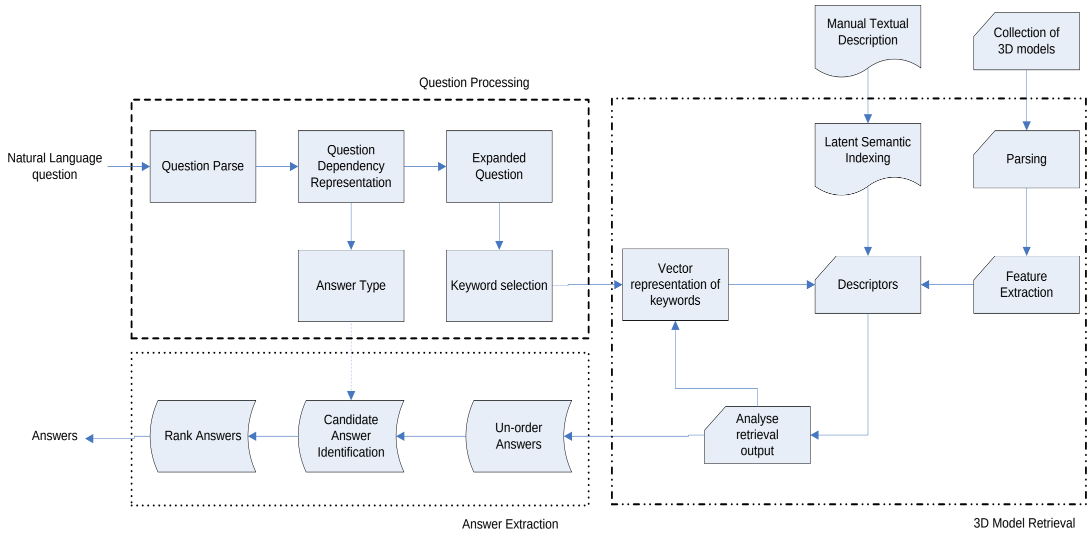

Figure 1 The proposed architecture model for a 3D information retrieval system that is integrated with general architecture for a QA system.

From the proposed model (see figure 1), the question processing is responsible for capturing a user's input in a form of natural language. To understand the semantics of the question, the system performs several steps to segment, find a semantic structure, and expand the query as suggested by the general architecture of a textual question answering system. During this process, the semantic category of expected answers is derived by projecting the question dependency representation onto an answer type hierarchy encoding semantic structure form the WordNet. The matching process starts after a set of keywords has been selected from the expanded query or the question module. The system then projects the selected keywords into the latent semantic indexing space to obtain a vector representation of the keywords. The vector representation is passed to the descriptors whose output is then analysed. If the number of retrieval 3D models is too small, this means the selected keywords are too restrictive. In contrast, if the number of the models is too large, the keywords are too general. The process is repeated until no keywords can be added or removed, which can be either in a manual or an automatic way. After that, the collection of retrieved answers or 3D models is passed through the process of answer extraction to identify the top rank candidate answers. Finally, the rank answers are presented to a user. Figure 2 demonstrates the proposal model of a feasible architecture.

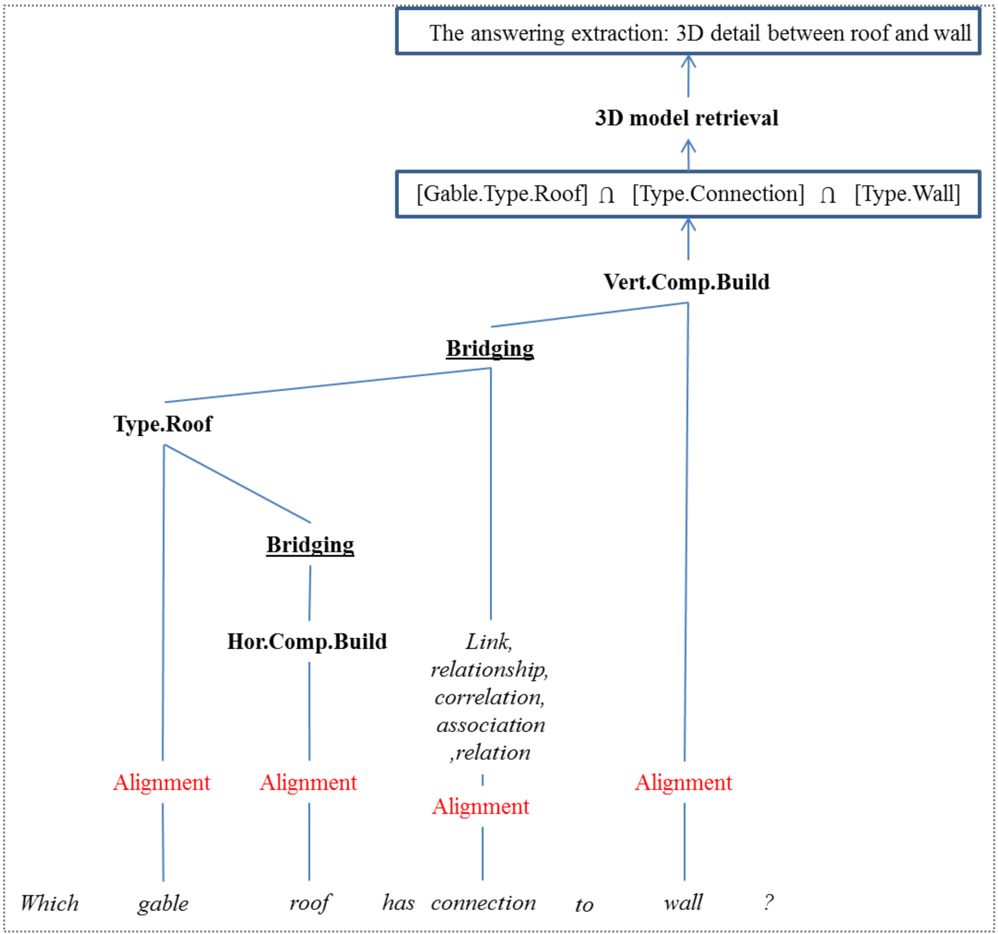

Figure 2 Process of mapping questions to answers via the question processing; passing keywords to the 3D model retrieval; and resulting from the answering extraction.

Conclusion

In design domain, the rapid development of 3D digital technology has led to an increasing volume of 3D model data and thus the volumes of digital 3D data archives are growing exponentially. The very large repositories of the digital 3D information bring about a challenging problem in retrieval and extracting the information for particular needs. Although the appearance of standards for retrieval 3D models has formalised the way information is retrieved and extracted, the use of model systems to effective indexing, clustering, and searching of 3D models based on automatically derived model remains relatively unexplored. This article intensively reviews ongoing researches in 3D retrieval systems. Despite a number of researches acknowledges that the current systems still have drawbacks on interpreting of what a user needs, this study discusses that the pitfall could be minimised and a common 3D retrieval system could greatly advantage if the system would have been integrated with a question-answering information retrieving system. As a result, this study proposes a new hybrid system model that is able to pre-compute the visual statistics of a 3D model in collections based on the low-level feature vector extraction. The proposed system would be a better alternative and would contribute to other researchers on this field of study.

References

Alt, H., & Godeau, M. (1995). Computing the Frechet distance between two polygonal curves. *International Journal of Computational Geometry & Applications*, *1&2*(5), 75-91.

Arkin, E., Chew, P., Huttenlocher, D., Kedem, K., & Mitchel, J. (1991). An efficiently computable metric for comparing polygonal shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *13*(3), 209-215.

Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. New York: ACM Press.

Bang, H., & Chen, T. (2002). *Feature space warping: An approach to relevance feedback*. USA: IEEE ICIP.

Barra, V. & Biasotti, S. (2013). 3D shape retrieval using Kernels on Extended Reeb Graphs. *Pattern Recognition*, *46*(11), 2985-2999.

Bikel, D. M., Schwartz, R. L., & Weischedel, R. M. (1999). An algorithm that learns what's in a name. *Machine Learning*, *34*(1-3), 211-231.

Bronstein, A. M., Bronstein, M. M., Guibas, L. J., & Ovsjanikov, M. (2011). Shape Google: geometric words and expressions for invariant shape retrieval. *ACM Trans Graph, 30*(1), 1-20.

Bustos, B., Keim, D. A., Saupe, D., Schreck, T., & Vranić, D. V. (2005). Feature-based similarity search in 3D object databases. *ACM Computing Surveys*, *37*(4), 345-387.

Cao, J., Robles-Flores, J. A., Roussinov, D., & Nunamaker, J. F. (2005). *Automated Question Answering From Lecture Videos: NLP vs. Pattern Matching*. In *the System Sciences, HICSS '05. Proceedings of the 38th Annual Hawaii International Conference* (pp. 1-8). Hawaii: IEEE *Computer Society*.

Chali, Y., Hasan, S. A., & Joty, S. R. (2011). Improving graph-based random walks for complex question answering using syntactic, shallow semantic and extended string subsequence kernels. *Information Processing & Management, 47*(6), 843-855.

Chowdhury, G. G. (2004). *Introduction to modern information retrieval*. London: Facet Publishing.

Corney, J., Rea, H., Clark, D., Pritchard, J., Breaks, M., & Macleod, R. (2002). Coarse filters for shape matching. *IEEE Computer Graphics and Applications*, *22*(3), 65-74.

Dwivedi, S. K., & Singh, V. (2013). Research and Reviews in Question Answering System. *Procedia Technology*, 10, 417-424.

Eitz, M., Hildebrand, K., Boubekeur, T., & Alexa, M. (2010). Sketch-based 3D shape retrieval. In *Conference ACM SIGGRAPH 2010 Talks* (p. 5). Los Angeles Convention Center: ACM, the Association for Computing Machinery.

El-Mehalawi, M., & Miller, R. (2003). A database system of mechanical components based on geometric and topological similarity. Part II: indexing retrieval matching and similarity assessment. *Computer-Aided Design*, 35, 95-105.

El-Mehalawi, M., & Miller, R. A. (2003). A database system of mechanical components based on geometric and topological similarity. Part I: representation. *Journal of Computer-Aided Design*, *35*(1), 83-94.

Elad, M., Tal, A., & Ar, S. (2001). Content based retrieval of vrml objects - an iterative and interactive approach. *EG Multimedia*, 39, 97-108.

Fry, D. S. (1993). *Shape Recognition Using Metrics on the Space of Shapes*. MA, USA: Harvard University, Cambridge.

Gao, Y., Dai, Q., & Zhang N. Y. (2010). 3D model comparison using spatial structure circular descriptor. *Pattern Recognit*, *43*(3), 1142-1151.

Gao,Y., Wang, M. Tao, D., Ji, R., & Dai Q. (2012). 3D object retrieval and recognition with hypergraph analysis. *IEEET rans Image Process*, *21*(9), 4290-4303.

Grishman, R. (1997). Information Extraction: Techniques and Challenges. In M. T. Pazienza (Ed.), *Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology* (pp. 10-27). Berlin: Springer-Verlag.

Hagedoorn, M., & Veltkamp, R. C. (1999). Metric pattern spaces. *Technical Report UU-CS-1999-03*, Padualaan: Utrecht University.

Hilaga, M., Shinagawa, Y., Kohmura, T., & Kunii, T. (2001). *Topology matching for fully automatic similarity estimation of 3D shapes*. In the 28th Conference on Computer Graphics and Interactive Techniques (pp. 203-212). Los Angeles: ACM SIGGRAPH.

Hirschmann, L., & Gaizauskas, R. (2001). Natural language question answering: The view from here. *Natural Language Engineering*, 7, 275-300.

Hou, S., Lou, K., & Ramani, K. (2005). SVM-based Semantic Clustering and Retrieval of A 3D Model Database. *Journal of Computer Aided Design and Application*, 2, 155-164.

Iyer, N., Jayanti, S., Lou, K., Kalyanaraman, Y., & Ramani, K. (2005). Three-dimensional shape searching: state-of-the-art review and future trends. *Computer-Aided Design*, *37*(5), 509-530.

Katz, B., Lin, J., Stauffer, C., & Grimson, E. (2004). Answering questions about moving objects in videos. In M. Maybury (Ed.). *New Directions in Question Answering* (pp. 113-124). Cambridge, Massachusetts: MIT Press.

Kazhdan, M., Funkhouser, T., & Rusinkiewicz, S. (2003). *Rotation invariant spherical harmonic representation of 3D shape descriptors*. Switzerland: Eurographics Association Aire-la-Ville.

Kortgen, M., Park, G. J., Novotni, M., & Klein, R. (2003). 3D shapematching with 3D shape contexts In *the 7th Central European Seminar on Computer Graphics* (pp. 55-66). Budmerice: Slovakia.

Lavoué, G. (2012). Combination of bag-of-words descriptors for robust partial shape retrieval. *Visual Comput*, *28*(9), 931-942.

Leifman, G., Meir, R., & Tal, A. (2005). Semantic-Oriented 3D Shape Retrieval using Relevance Feedback. *The Visual Computer (Pacific Graphics)*, *21*(8-10), 865-875.

Lu, K., He, N., & Xue, J. (2009). Content-based similarity for 3D model retrieval and classification. *Progress in Natural Science*, *19*(4), 495-499.

Lu, K., Wang, Q., Xue, J., & Pan, W. (2014). 3D model retrieval and classification by semi-supervised learning with content-based similarity. *Information Sciences*, *28*(10), 703-713.

McWherter, D., & Regli, W. (2001). *An approach to indexing databases of solid models*. Philadelphia: Department of Mathematics and Computer Science.

Mezaris, V., Kompatsiaris, I., & Strintzis, M. G. (2003). *An Ontology Approach to Object-Based Image Retrieval*. Paper presented at the IEEE International Conference on Image Processing, ICIP Barcelona: Spain.

Min, P., Kazhdan, M., & Funkhouser, T. J. (2004). A Comparison of Text and Shape Matching for Retrieval of Online 3D Models. *Research and Advanced Technology for Digital Libraries*, 3232, 209-220.

Novotni, M., & Klein, R. (2003). 3D Zernike descriptors for content based shape retrieval. In *proceedings of The 8th ACM Symposium on Solid Modeling and Applications* (pp. 216-225). Germany: University of Bonn Institute of Computer Science II.

Ohbuchi, R., Otagiri, T., Ibato, M., & Takei, T. (2002). Shape-similarity search of three-dimensional models using parameterized statistics. In *Proceedings of Pacific Graphics 2002* (pp. 265-274). Beijing: IEEE *Computer Society*.

Ohbuchi R., & Shimizu, T. (2008). Ranking on semantic manifold for shape-based 3D model retrieval. In *Proceedings of the 1st ACM International Conference on Multimedia Information Retrieval* (pp. 411-418). New York: University of Yamanashi.

Peabody, M., & Regli, W. (2001). Clustering techniques for databases of CAD models. *Technical* Report DU-MCS-01-0, Philadelphia, PA: Department of Mathematical and Computer Science, Drexel University.

Rijsbergen, C. J. V. (1979). *Information Retrieval: Butterworth-Heinemann*. USA: Newton, Massachusetts.

Scarlo, S., Cascia, M. L., Sethi, S., & Taycher, L. (1999). Unifying textual and visual cues for content-based image retrieval on the World Wide Web. *Computer Vision and Image Understanding*, *75*(1-2), 86-98.

Shin, J. L., Lee, C. H., & Wang, J. T. (2007). A new 3D model retrieval approach based on the elevation descriptor. *Pattern Recognit*, *40*(1), 283-295.

Sundar, H., Silver, D., Gagvani, N., & Dickenson, S. (2003a). Skeleton based shape matching and retrieval. In the Shape Modeling International 2003 (pp. 130-139). Seoul, Korea: IEEE Computer Society.

Veltkamp, R. C. (2001). Shape matching: similarity measures and algorithms. In *SMI 2001 International Conference on Shape Modeling and Applications (pp. 118*-197). Genova, Italy: IEEE Computer Society.

Vleugels, J., & Veltkamp, R. C. (1999). Efficient image retrieval through vantage objects. In D. P. Huijsmans, & W. M. ASmeulders (Eds.), *Visual Information and Information Systems - Proceedings of the Third International Conference VISUAL* (pp. 575-584). The Netherlands: Amsterdam.

Wang, F., Lin, L., & Tang, M. (2014). A new sketch-based 3D model retrieval approach by using global and local features. *Graphical Models*, *76*(3), 128-139.

Wang, X. & Nie, W. (2015) 3D Model Retrieval with Weighted Locality-constrained Group Sparse Coding. *Neurocomputing* , 151, 620-625.

Westerveld, T., de Vries, A. P., van Ballegooij, A., de Jong, F. M. G., & Hiemstra, D. (2003). A probabilistic multimedia retrieval model and its evaluation. *Journal on Applied Signal Processing, special issue on Unstructured Information Management from Multimedia Data Sources*, 2, 186-198.

Wu, Z., Xu, G., Zhang, Y., Cao, Z., Li, G., & Hu, Z. (2012). GMQL: A graphical multimedia query language. *Knowledge-Based Systems*, 26, 135-143.

Xiao, J., Feng, Y., Ji, M., & Zhuang, Y. (2014). Fast view-based 3D model retrieval via unsupervised multiple feature fusion and online projection learning *Signal Processing*, in Press, Corrected Proof, Available online 12 January 2014. doi: 10.1016/j.sigpro.2014.11.020

Yang, H., Chaisorn, L., Zhao, Y., Neo, S. Y., & Chua, T. S. (2003). *Video QA: Question answering on news video*. In *the Proceedings of the Eleventh ACM. International Conference on Multimedia (pp. 632-641)*. Berkeley, CA: ACM SIGGRAPH.

Zaharia, T., & Preteux, F. (2001). 3D shape-based retrieval within theMPEG-7 framework. In *Proceedings of SPIE Conference on Nonlinear Image Processing and Pattern Analysis XII 4304* (pp. 133-145). USA: San Jose.

Zhang, C., & Chen, T. (2001). Efficient feature extraction for 2D/3D objects in mesh representation. In *the International Conference on Image Processing* pp. 935-938). New York: IEEE *Computer Society*.

Zhang, C., & Chen, T. (2001). Indexing and retrieval of 3D models aided by active learning. In *the Proceedings of ACM International Conference on Multimedia* (pp. 615-616). Canada: Ottawa.

Zhao, S., Yao, H., Zhang, Y., Wang, Y., & Liu, S. (2015). View-based 3D object retrieval via multi-modal graph learning. *Signal Processing*, 112, 110-118.