S.I. : VR AND AR SERIOUS GAMES

CrossMark

Semantic framework for interactive animation generation and its application in virtual shadow play performance

Hui Liang[1] · Shujie Deng[2] · Jian Chang[2] · Jian Jun Zhang[2] · Can Chen[3] · Ruofeng Tong[4]

Abstract

Designing and creating complex and interactive animation is still a challenge in the field of virtual reality, which has to handle various aspects of functional requirements (e.g. graphics, physics, AI, multimodal inputs and outputs, and massive data assets management). In this paper, a semantic framework is proposed to model the construction of interactive animation and promote animation assets reuse in a systematic and standardized way. As its ontological implementation, two domain-specific ontologies for the hand-gesture-based interaction and animation data repository have been developed in the context of Chinese traditional shadow play art. Finally, prototype of interactive Chinese shadow play performance system using deep motion sensor device is presented as the usage example.

1 Introduction

As one of the most famous folk arts, Chinese traditional shadow play has a long history and contains rich cultural elements. Mainly made of donkey's hide, flat shadow puppet consists of several parts and its joints are connected by threads. When playing, puppeteers use their hands to manipulate the flat shadow puppets through sticks attached on the puppets, and the shadows are projected on a simple rear-illuminated cloth screen to create moving pictures (Chen 2015), as illustrated in Fig. 1.

However, this intricate operating skill is a higher entry barrier for performers, which requires long-time special training. This valuable and fascinating art form is becoming less known to the public, or even dying nowadays. New technologies are required urgently to give new life to the historic art forms.

Our motivation is to develop a novel method to generate interactive shadow puppetry animation for the purpose of cultural heritage preservation and dissemination. The ease of use and user-friendly virtual interaction would attract more people's interest and engage young generations to generate shadow play animation without special trainings.

Developed for decades, as an essential technology, interactive animation is wildly used in the fields of 3D games and virtual reality, which offers users an immersive experience by providing natural and intuitive interactions with novel input/output devices. However, even though advanced theories, techniques and devices help a lot for generating animation, interactive animation production is still a tedious work and labour-intensive. Its inherent features of complexity—having to handle various aspects of functional requirements (e.g. graphics, physics, artificial intelligence, engineering, multimodal inputs and outputs)—throw out challenges to the current research.

One appealing solution is leveraging ontology to construct a systematic and standardized framework at a highly abstract and semantic level to provide a full view and understanding of the complex systematic procedure. Using structured terminology, ontological analysis could capture the core logic of complex system with natural language descriptions (Ramaprasad and Papagari 2009). In recent years, as a high-level conceptual specification, the notion

✉ Jian Chang
   jchang@bournemouth.ac.uk

1   Zhengzhou University of Light Industry, Zhengzhou, China

2   National Centre for Computer Animation, Poole, UK

3   Changzhou University, Changzhou, China

4   Zhejiang University, Hangzhou, China

Springer

Fig. 1 Chinese traditional shadow play performance

of ontologies has become increasingly important in the computer animation-related domains, such as semantic 3D content representation (Flotyński and Walczak 2013a, b, c, 2014), ontology-based 3D model retrieval (Li and Karthik 2007; Ohbuchi et al. 2007) and their usage in virtual scene and game environment (e.g. modelling the semantic information of the virtual game environment).

Additionally, animation data management has become a focal attention to animation production and a popular issue of interests including animation data archiving and reuse in a more efficient and user-friendly way. Because of the phenomenal growth of the animation data, efficient data management, such as structured data presentation and searching for particular information, has become a daunting task. Traditional text-based or content-based media data retrieval introduces problems of the semantic gap between the low-level description and the high-level semantic interpretation of multimedia object. Ontology-based semantic retrieval, as an appropriate way to represent structured knowledge bases, enables data sharing, reuse and inference, and also advances in bridging the semantic gap (Li and Karthik 2007; Ohbuchi et al. 2007). It provides us suitable technique on the animation assets management to facilitate data storage, organization, retrieval, reuse and repurposing.

For the purposes of heritage preservation, and also as a usage example, a prototype of hand-gesture-based interactive Chinese shadow play animation is generated based on the proposed framework and ontologies. It is expected that our method can provide guidance for building and rapid prototyping of diverse interactive animation/game applications.

Our work has the following contributions:

- Semantic framework for interactive animation generation.
- Domain-specific ontologies for interactive animation generation, including ontologies for hand-gesture-based interaction and animation data assets management.

- Prototype of interactive Chinese shadow play performance system.

The rest of this paper is organized as follows. Section 2 reviews the related works briefly. Semantic framework for interactive animation generation is proposed in Sect. 3. In Sect. 4, two domain-specific ontologies are implemented based on the framework in the context of the interactive Chinese shadow play performance. Prototype of virtual interactive shadow play performance system involving hand-gesture-based interaction and ontology-based animation data management is implemented in Sect. 5. Finally, Sect. 6 concludes the paper.

2 Related work

According to Gruber (1993), ontology is an explicit formal specification of a conceptualization, which is an effective tool to describe general concepts of entities as well as their properties, relationships and constraints (Grüninger and Fox 1995). Also as a practical application in information science and technology, ontology provides the vocabulary to define terminology and the constraints required by different applications that cover various fields, from knowledge engineering to software engineering, taking the advantage of the establishment of common vocabulary and semantic term interpretation (Borst 1997; Studer et al. 1998).

Traditional model retrieval is based on keywords or the shape properties similarity (e.g. geometry and topology) (Tangelder and Veltkamp 2008; Kazhdan et al. 2003), in which only the textual or physical information is utilized. The semantic information is ignored, and the efficacy of retrieval has not been satisfactory, however. AIM@SHAPE Network of Excellence (Falcidieno 2004) made first step of using semantics approach to describe and search 3D models to facilitate reuse of the animation contents (Falcidieno

et al. 2004). There are similar semantic search engines including 3D model Search Engine (Princeton University) (Funkhouser et al. 2003) and Google 3D Warehouse (2016). To describe media data such as images, audio, video and 3D objects, ontological solutions, such as the Ontology for Media Resources (2016) and the Core Ontology for Multimedia (COMM) (Arndt et al. 2009), are devised on the basis of standard models for data interchange on the web (e.g. RDF, RDFS, OWL and MPEG-7). To facilitate 3D content accessing, ontology-based approaches are used for semantic modelling of the different 3D models (e.g. geometry, appearance and behaviour). Ontology for virtual human was proposed in Gutiérrez et al. (2007), which incorporated semantics into the description of human shapes.

Another application is to model semantic information of virtual environment. Unlike traditional virtual environments design, semantic virtual environments maintain richer semantic information, in which the ontological model provides an abstract and semantic description that is adequate for computer processing (Otto 2005). Different aspects of the conceptual representation of the virtual world are modelled by ontologies as high-level and semantic description, including the environment structures, entities' behaviour and domain knowledge (Gutierrez et al. 2005). The concept of annotated environment together with structured representations of its contents and purposes was firstly proposed by Thalmann et al. (1999). An ontology-based framework with a cognitive middle layer and environment managed semantic concepts were presented by Chang et al. (2005). Studies (Bilasco et al. 2005; Mansouri 2005) described the semantics of a 3D scene focusing on the high-level description of objects or the composition of existing objects. There are also some ontological research works using semantic information to model the interaction of virtual environments, for example NiMMiT (Notation for MultiModal interaction Techniques), which is a diagram-based notation describing multimodal interaction (De Boeck et al. 2006). A contextual augmented reality environment was proposed in Ruminski and Walczak (2014), which consists of three elements: the trackables, content objects and interface. With the model of SCM (Semantic Content Model) designed for semantic representation of interactive 3D contents, environments were represented at different levels of abstraction: conceptual level, concrete level and platform level (Flotyński and Walczak 2013).

There are some ontological applications in the domain of computer games: for example, a game ontology that was developed based on game theory to facilitate interactions in multiagent system (Mirbakhsh et al. 2010). In Semantic Class Library designed for semantic 3D game virtual environment (Tutenel et al. 2009), 3D model can be classified and provide with additional semantic information, such as physical attributes, functional information besides 3D

representations. Many research works have been conducted to develop game-related ontologies, e.g. developing game content ontology to describe games characteristics, properties and design process (Leino 2010; Teixeira et al. 2008; Dubin and Jett 2015), using ontological framework as the guide to develop and employ game-based training (Anthony et al. 2009), using an ontology and a set of rules to represent game logics (BinSubaih et al. 2005), using ontologies as a solution to enable transfer of meaningful game information between different video games (Parkkila et al. 2015), using defined learning game ontology as evaluation methodology of serious games for teachers/trainers to choose and retrieve learning games (Ghannem 2014).

3 Semantic framework for interactive animation

As a complex system, interactive animation production involves various functional requirements, including interaction methods and animation data management. To provide a clear vision of this complex process, a three-layer semantic framework is proposed as a formal semantic foundation as illustrated in Fig. 2. Constructed at a highly abstracted level, this semantic framework can be easily applied to different VR applications for various purposes and is expected to facilitate the conceptual design process.

At the top abstract layer, the semantic framework provides a systematic and intuitive description with larger granularity for the design of animation generation. As a complex system, multimodal function combinations are involved in the animation procedure, such as UI, animation data management, game design, story scripts and system integration. In this research, two key components—UI and data repository, are our main concerns, instead of including all the related techniques. These components involve two most important aspects of interactive animation generation: the virtual interaction and animation data assets accessing. The domain layer is the ontological implementation of the upper abstract layer, at which the domain-specific ontologies are further defined to provide knowledge support for the development of particular applications. And then, using the knowledge provided by the developed domain-specific ontologies, various applications, such as interactive animations, games and animation database, could be finally developed at the generation layer.

In the component of UI, player interacts with computer system, which involves the use of various input/output devices and interactive modes, such as monitor screen, keyboard, mouse, touch screen, haptic device, motion sensor and tracking device. Animation data repository component provides data support, in which various animation data resource is maintained and managed for the purpose of

Fig. 2 Architecture of semantic framework for interactive animation generation
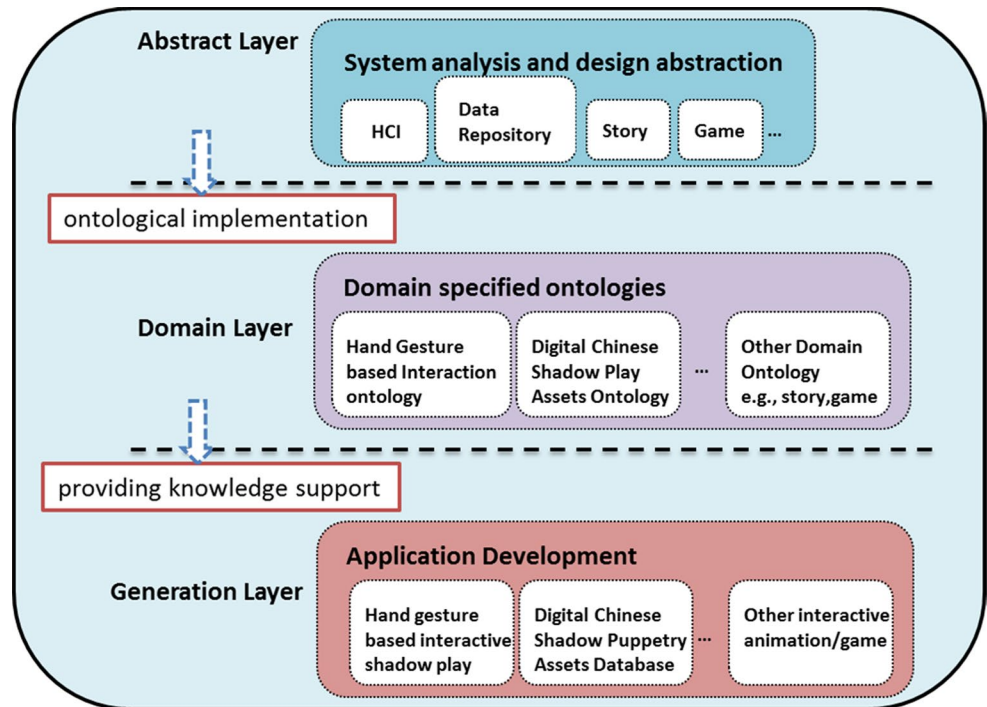
enhancing reusability. In Sect. 4, we will discuss its implementation by developing two domain-specific ontologies, which is mapped to the domain layer, and how the proposed framework can be employed to generate interactive Chinese shadow play animation, which is mapped to the generation layer.

3.1 UI component

As illustrated in Fig. 3, the UI component contains a set of structural concepts, which could be decomposed into three levels of semantic abstraction. Due to the multidisciplinary nature of UI, the interaction between the user and the computer system involves the use of various input/output devices and interactive modes, e.g. touch screen, haptic device and head-mounted devices. The method for designing and implementing novel computer interface and function usability finally provided to users is crucial to facilitate the development of virtual reality.

3.2 Animation data repository component

As the data resource for animation creation, the digital assets are multimodal involving audio, video, 2D image/textual, 3D models, motion files, scene files, etc., and they vary a lot. Targeting different topics and depending on the specific application domains, the contents of the digital data repositories are also different. From a semantic perspective, the repository could be abstracted and analysed with several sublayers. Taking 3D contents for example, they are

interpreted with four layers in this paper: geometry, structure, appearance and logic, as illustrated in Fig. 4a. Figure 4b presents an example of animation content, which is a digital character of the Chinese traditional shadow puppetry. As a higher level of expressiveness, it represents a brave warrior in the play of "The Emperor and the Assassin" from the aspect of the logic layer. The shape can be considered as a structure composed by different parts of shadow puppet presenting geometric information. Added with the appearance information, such as colour, pattern or accessories, a complex character is created carrying rich culture contents.

The conceptual framework can systemize and standardize the procedure of interactive game/animation synthesis and promote system integration in a semantic representation. The two components are abstract concepts of interactive animation generation. Hence, the framework enables conceptual design for various interactive animation applications.

4 Development of domain-specific ontologies

Our framework sets a semantic foundation to formally model the interactive animation production from the abstract high level. This section focuses on investigating the cases in which we apply this semantic framework to guide the implementation of the domain-specified ontologies. As a practical example, our method is presented in the context of a 2D Chinese traditional shadow puppetry performance scenario, in which players manipulate virtual characters to

Fig. 3 UI component
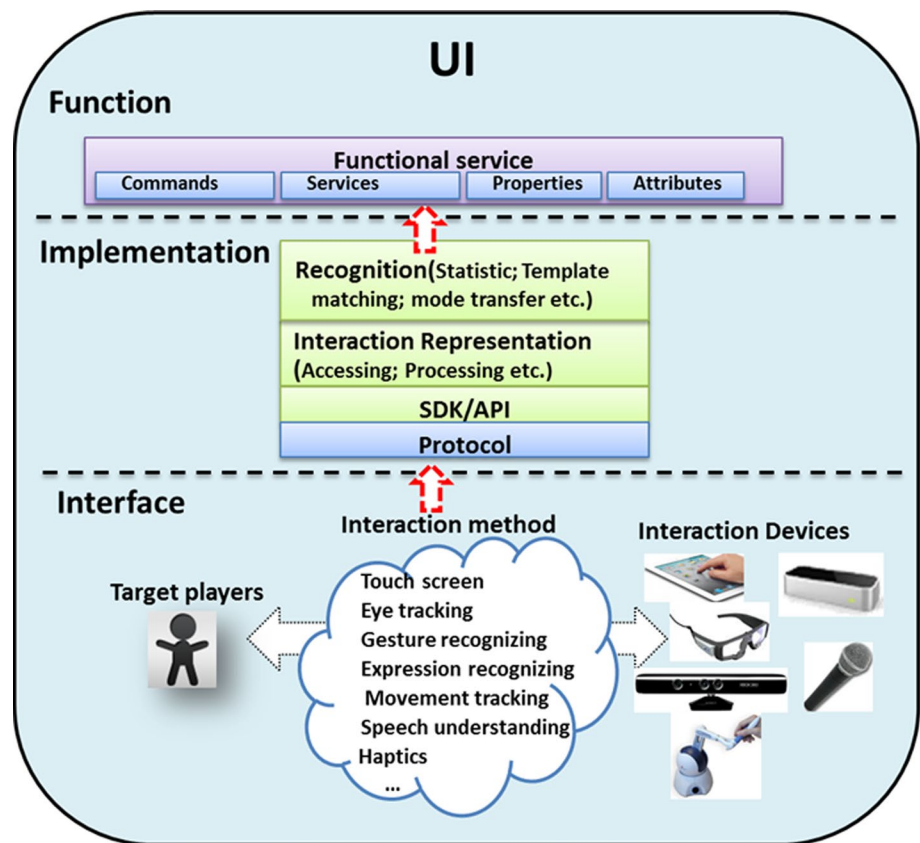
produce animation with hand gestures utilizing depth motion sensing device—the leap motion controller.

Using the Web Ontology Language (OWL 2012) together with the ontology editor and knowledge framework-Protégé (2016), the proposed semantic framework has been implemented as two domain-specified ontologies: Hand-Gesture-Based Interaction Ontology (HGBIO) mapped to the UI component, and Digital Chinese Shadow Puppetry Assets Ontology (DCSPAO) mapped to the animation data repository component. We take advantage of the OWL ontology structure in order to incorporate SWRL (semantic web rule language) (2016) rules, which can be visualized using SPARQL (a recursive acronym for SPARQL protocol and RDF query language) (2016) queries for knowledge accessing.

4.1 Hand-Gesture-Based Interaction Ontology (HGBIO)

Chinese traditional shadow puppet is made of several flat parts. Each part is linked together by joints. Thin sticks are attached to the key parts separately for controlling. Holding sticks in hands, puppeteer uses hand gestures to control puppet's movement and performs the play.

In our method, we adopt the traditional puppet design mode. Each digital puppet consists of ten parts linked by nine joints. Digital puppet's motion is controlled by player's hand movement and gesture using motion sensor device. During playing, player's hand position is mapped to the position of puppet in virtual environment through motion sensor after mapping. The animation is generated through a hand-gesture-based interaction. The detail of the motion control is described in Sect. 5 "Implementation".

The diagram of a part of the HGBIO is shown in Fig. 5a. The ontology defined beneath the root of every OWL ontology "owl:Thing" has three subclasses: "Human", "Device" and "Method". Each subclass (ancestor) contains various subclasses (descendant) of their own. Each one of these descendants contains individuals for the assignment of characteristics. A number of object properties are designed to give the ontology fundamental relational functionalities. The set of data type properties are provided to link OWL individuals with typical data values. The structure of hand gesture ontology is organized in a hierarchy consisting of three levels. The lower level presents motion tracking data accessed from the leap motion sensor controller including information of frames, list of detected hands and list of pointables (such as fingers or finger-like tools). Each frame contains the measured positions and the following information of the object detected by the controller. The device provides model of hands with five fingers of bone structures, and each finger is made of four bone joints. By
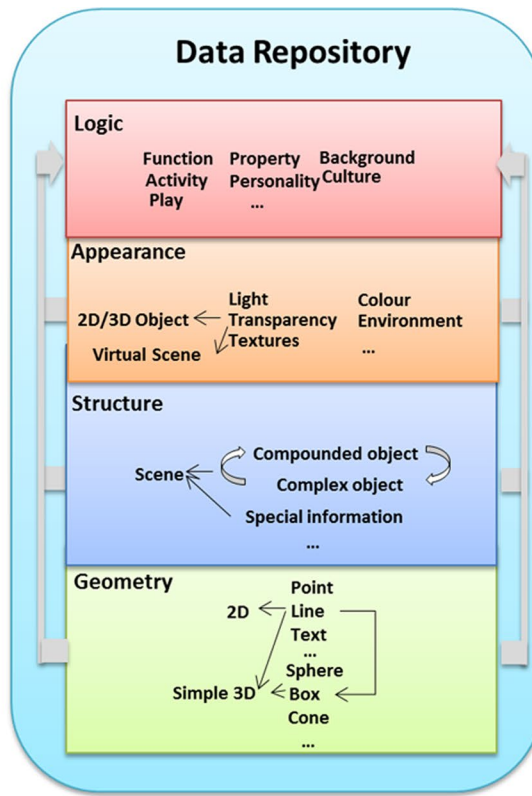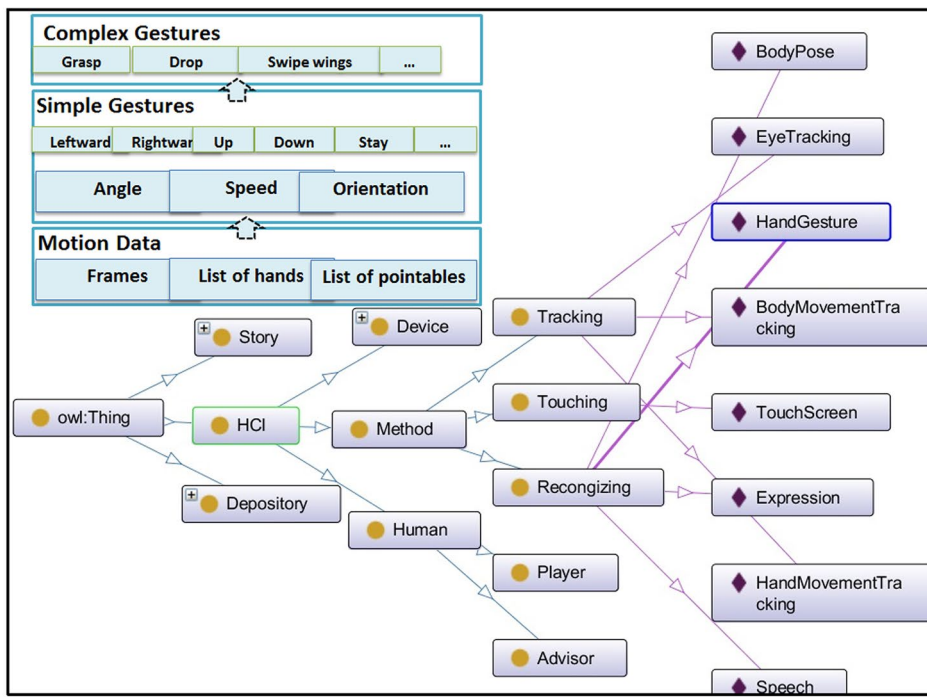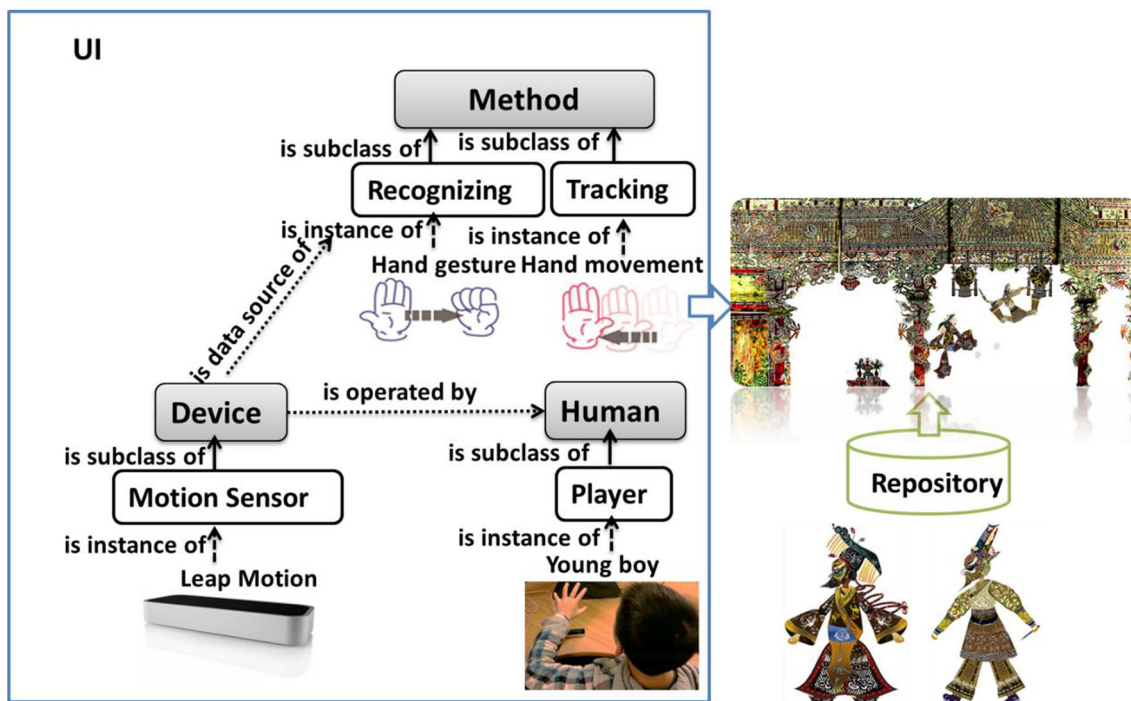
(a)
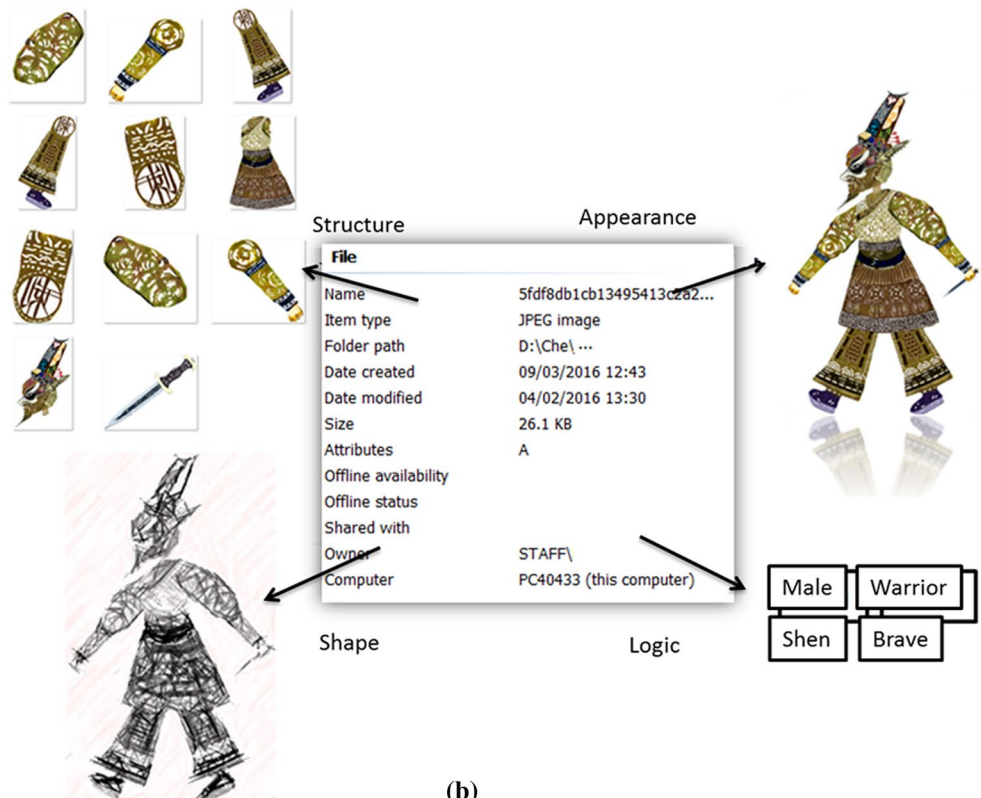
(b)

(a)

(b)

Fig. 5 Diagram of Hand-Gesture-Based Interaction Ontology (HGBIO)

using the positions and rotations information of each finger bone, we could design specific interactions. At the middle level, by calculating hands movement speed, the angle of movement and the orientation angle, we can define simple

hand gestures and describe the hand movement. For example, "Leftward" represents "Fly/run to the left", "Rightward" represents "Fly/run to the right", "Upward" represents "Fly/jump up", "Downward" represents "Fly/crouch down", and

"Stay still" represents "Hover in the air/stay unmoved". At the top level, we can develop more complex hand gestures which are the combination of simple hand gestures. These compounded gestures could meet player's further needs by depicting sophisticated interactive semantic meaning. For example, "Hand open to close" represents "Grasp", "Hand close to open" represents "Drop", and "rapid multiple finger taps" represents "Swipe wings".

Figure 5b shows a usage example of the defined HGBIO. The visualization displays the relationships among the selected class and other classes in its neighbourhood.

"Player" is a subclass of "Human", which has an instance a young boy, who is using his hand gesture to control the movement of an avatar in the virtual environment through leap motion controller. Through the object property "is operated by", the leap motion controller is connected with player to deliver the semantic information: "Device is operated by play for interaction".

There are a range of interactive devices that enable user interactions using hand gestures, such as leap motion and kinect. In this diagram, leap motion is an instance of class "Motion Sensor" which is a subclass of "Device". The essential attribute of leap motion is to provide data source to the class "Method", which is described by the object property "is data source of" that links two classes.

The sensor data provided by leap motion contain a diversity of information about hands and fingers including id, direction, position, orientation, relative information and status. "Method" is the core of the system, which has two subclasses in our case: "Tracking" and "Recognizing". Hand Movement is an instance of "Tracking", which is responsible for processing the player's hand movement and mapping the relative hand position to the movement of the avatar in virtual world. Recognizing function is in charge of identifying the pattern of hand movement and gestures implicating player's intention. Hand gesture is an instance of "Recognizing", which provides a natural and intuitive gestural interaction manner. Once a recognizable hand gesture is detected, recognizing function will trigger a pre-recorded animation or event (e.g. avatar grips a pebble or flaps wings).

4.2 Digital Chinese Shadow Puppetry Assets Ontology (DCSPAO)

Traditional Chinese shadow play has distinctive folk style and artistic characteristics. As the implementation of the domain layer of the proposed semantic framework, a domain-specified ontology is developed for traditional Chinese shadow puppetry; furthermore, a shadow puppetry animation data repository is also used as the digital assets for the animation generation.

The diagram of a part of the DCSPAO is shown in Fig. 6a. Defined beneath the root of every OWL ontology

"owl:Thing", DCSPAO has several subclasses including "Role", "Music", "Prop" and "Scene". And each of these subclasses has the superclass at the same time, also its descendant. For example, the human characters in traditional Chinese shadow puppetry play fall into four major roles—Sheng (main male role), Dan (female role), Jing (painted-face and forceful male role) and Chou (clown male role), which leads to four subclasses of the class "Role" correspondingly (Shadow Play 2016).

A set of object properties are added to create relationship among classes by linking relevant objects, such as "has Prop of", "has Scene of" and "has Music of". We also propose data properties which provide attribute descriptions for classes in the DCSPAO, including "name", "age", "rank" and "personality".

As an example, an ontological description of the character "Jing Ke" is illustrated in Fig. 6b, who is one of the most famous generals in ancient China and widely respected as a symbolic cultural icon of loyalty and righteousness in general worship. "Laosheng" (decent middle-/old-aged man with beard) is a subclass of "Sheng" (main male role). As its instance, "Jing Ke" has a set of inherited properties including name, age and personality. As a most famous character in shadow play, "Jing Ke" has unique personality and character symbolism, which is normally presented by particular props (e.g. his trademark weapon) and other dramatic elements. This specified relation between the class "Role" and the class "Prop" is determined by the object property defined in the domain ontology. For instance, through the object property "has Prop of", "Jing Ke" and his pole weapon named "Yu Chang" (an instance of the subclass of "Weapon") are connected, which is the same as the relationship between "Jing Ke" and his war horse (an instance of the subclass of "Animal" which belongs to other characters except human roles).

Since "Jing Ke" is a significant role in the classic shadow play "The Emperor and the Assassin", object property "has Scene of" can assign a particular instance ("Xian Yan Palace") of the class "Scene" to him, which implicates the significant battlefield in the play. As a key role in shadow play, music creates particular atmosphere and shape characters. Object property "has Music of" can be used to reflect the bravery personality by connecting the unique melody "Splendour Allegro" with "Jing Ke".

5 Implementation of the prototype of virtual interactive Chinese shadow play performance system

To demonstrate how the proposed semantic framework and developed ontologies can be employed to fertilize the design of interactive animation generation, the prototype
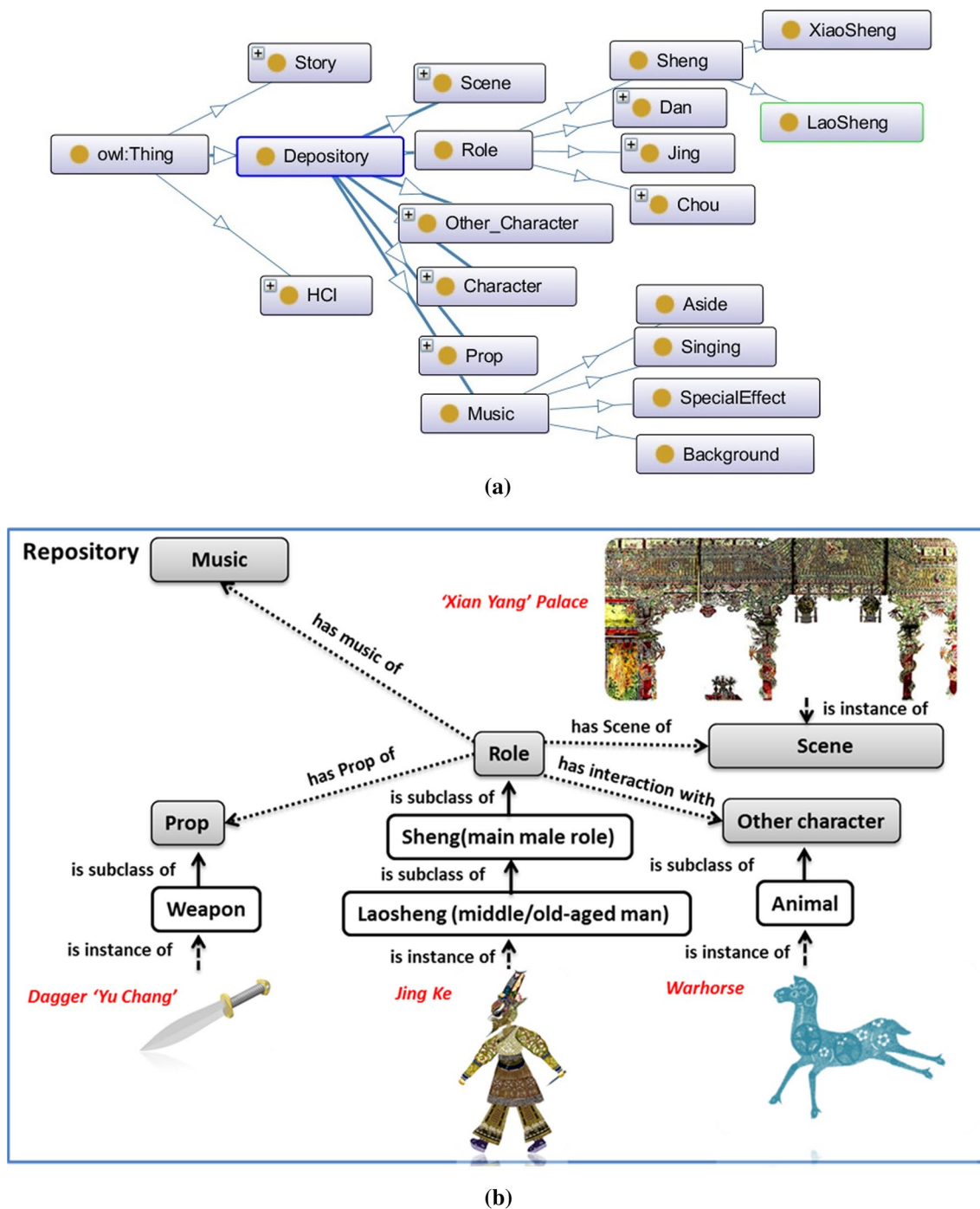
(a)

(b)

Fig. 6 Diagram of Digital Chinese Shadow Puppetry Assets Ontology (DCSPAO)

of an interactive Chinese traditional shadow play performance system is generated, which enables players to interact with virtual shadow puppetry using hand gestural control intuitively.

As the ontological implementation of UI component in the semantic framework, HGBIO has been used as a guidance to develop a novel hand-gesture-based interaction method. By utilizing depth motion sensing technology, the method tracks and recognizes user's hand and finger motions as input, which enables the users to use hand gestures to control the movement and manipulate animation of the digital shadow puppetry with natural interaction/ control and immersive experience. DCSPAO has been used to implement the animation data repository component of

the semantic framework. Using domain-specific information extraction, inference and rules to exploit the semantic metadata, the repository supports ontology-based retrieval, which improves searching performance by recognizing the animator's intent and contextual meaning of the digital assets in the context of Chinese traditional shadow play. More relevant results can be generated to meet the user's needs and also provide the capability of data reuse. Figure 7 gives an overview of the prototype architecture which has the following components.

5.1 Hand-gesture-based interaction

To provide natural interactive experience for the players, we design and develop a novel hand-gesture-based interaction by utilizing depth motion sensing technology as the interaction method, which allows them to use hand gestures to play with virtual shadow puppetries and manipulate them to interact with virtual items in virtual environment. Leap motion controller is used in our system as the interaction device to track hand gestures, which can provide a high-fidelity finger tracking through an infrared depth sensor. We utilize the leap motion SDK provided by the Leap Motion Co. as the API to access the motion data of hands and fingers from the device.

UI module mainly includes three submodules: input data processing, movement control and output. For more details of hands movement control mechanism and gesture recognizing, please refer (Liang et al. 2015a, b).

5.1.1 Input data processing

As illustrated by the UI component in Fig. 7, once the connection between the leap motion controller and our system is established, the deep images of user's hand will be achieved. The sensor data provided by leap motion controller contain a diversity of information about hands and pointables [such as fingers or finger-like tools defined by Leap Motion Co. (Gutierrez et al. 2005)], in the virtual scene, which is updated by frame and can be represented as follows:

$$SensorData = \langle FR, H, P, T \rangle \tag{1}$$

where FR is frame rate; H represents the set of hands detected and P represents the set of pointables; and T is timestamp. Hand data H mainly contain the hand id, direction and different value about the palm position and status:

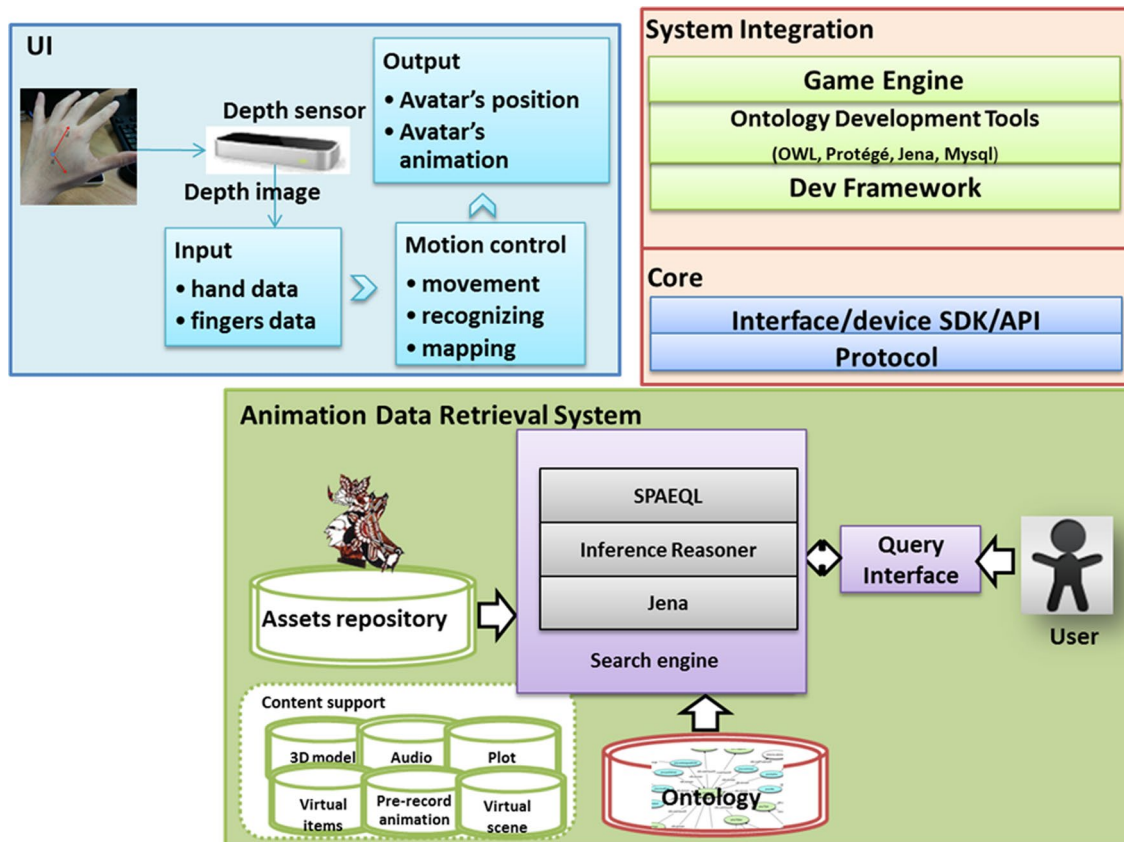$$H = \langle id, dir, palmInf \rangle \tag{2}$$

Fig. 7 Prototype architecture

And pointables data P include its id, direction and position information relative to the hand:

$$P = \langle id, dir, handid, positionInf \rangle \tag{3}$$

5.1.2 Motion control

To control the movement of the shadow puppetry, the most direct way is converting the translation of their hands position in leap motion coordinate system to the position of puppet in virtual environment. Considering the different scale between these two workspaces, a proper transformation matrix should be involved as a scaling into the coordinate translation. Once players' hands is not recognized by the leap motion device, puppet will keep its position in the last frame and then resume the movement immediately when hands can be detected again in the following frames. The virtual puppet's position in virtual workspace will move accordingly as shown in Eq. (4):

$$C_{avatar} = T \cdot S \cdot R_{hand} + C'_{avatar} \tag{4}$$

where C_{avatar} represents the virtual puppet's position coordinate in current frame, R_{hand} is the player's relative hand movement in real world, i.e. the displacement of hand positions in current frame and the previous frame, which can be obtained in Eq. (1), T is the transformation matrix between the player's workspace and the avatar's coordinate system, S is the scaling matrix, and C'_{avatar} represents avatar's position in the previous frame.

How to design hand gestures for the purposes of ease of use and intuition is of particular importance for interaction experience, which involves various considerations, such as the natural language, players' physical and memorial limitation, recognition accuracy of the input device and the assumption of the task in storyline. To provide a natural and intuitive interaction, we have designed our hand gestures vocabulary elaborately. There are in total two sets of hand gestures ("Steering" gestures and "Trigger" gestures), and each set is attached with one hand: "Steering" gestures are assigned for navigation tasks, and "Trigger" gestures are allotted for action performance.

To control virtual puppets' movement in the virtual environment, players just need to move their stretched palms intuitively. The virtual puppet's position in virtual workspace will move accordingly as shown in Eq. (4). This task is designed to be accomplished by a set of "Steering" gestures performed by one hand (the red hand), as illustrated in Table 1.

Besides the navigation, virtual puppets' actions are controlled by another set of pre-defined "Trigger" gestures performed by another hand (the blue hand), as illustrated in Table 2. We have designed characteristic actions for the Chinese traditional shadow puppetries, including walking, running and singing, and these actions can be triggered and

Table 1 Pre-defined "Steering" hand gestures
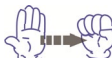
Gestures	Illustration
Leftward	
Rightward	
Upward	
Downward	
Hover/stay	

Table 2 Part of pre-defined "Trigger" hand gestures

| Gestures	Illustration
Grasp	 Hand open to close
Drop	 Hand close to open
Walk	 Single finger tap
Run	 Rapid multiple finger taps

performed corresponding to players' hand gestures. To provide real-time response and live visualization, these actions are integrated as a set of pre-recorded animation clips. Furthermore, these pre-recorded animations associated with puppets' particular actions can only be triggered by specific events of hand movement and hand gestures.

5.2 Ontology-based animation data assets retrieval

As one of the most important aspects of the interactive animation generation, animation data assets is a mixed-type animation database of digitalized traditional Chinese shadow puppets, which involves domain-specified knowledge including classification of performers and roles, visual/aural performance elements and repertoire, and depicts model's high-level concept. Two key submodules: Assets repository and retrieval systems are involved to provide animation content support. Using domain-specific information extraction to exploit the semantic metadata, the data repository supports data retrieval in a more intelligent way. Through automated

Fig. 8 Interface of retrieval system
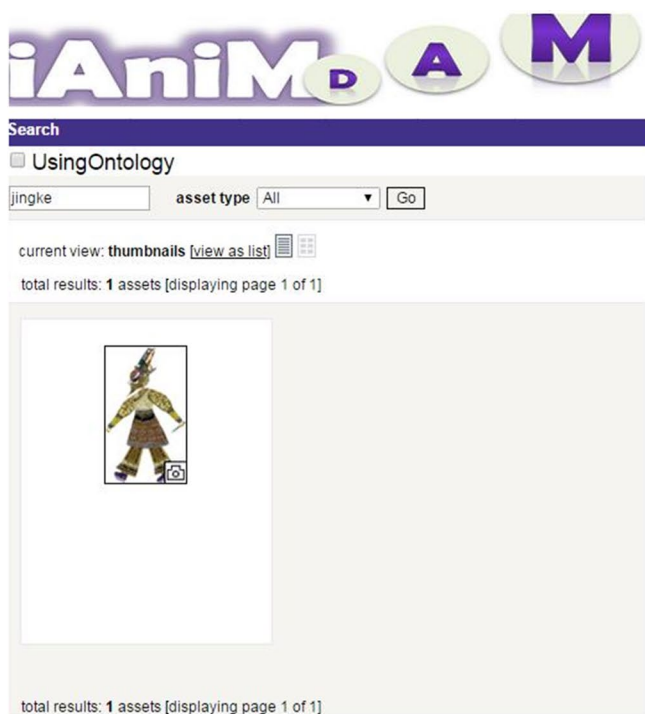
(a)
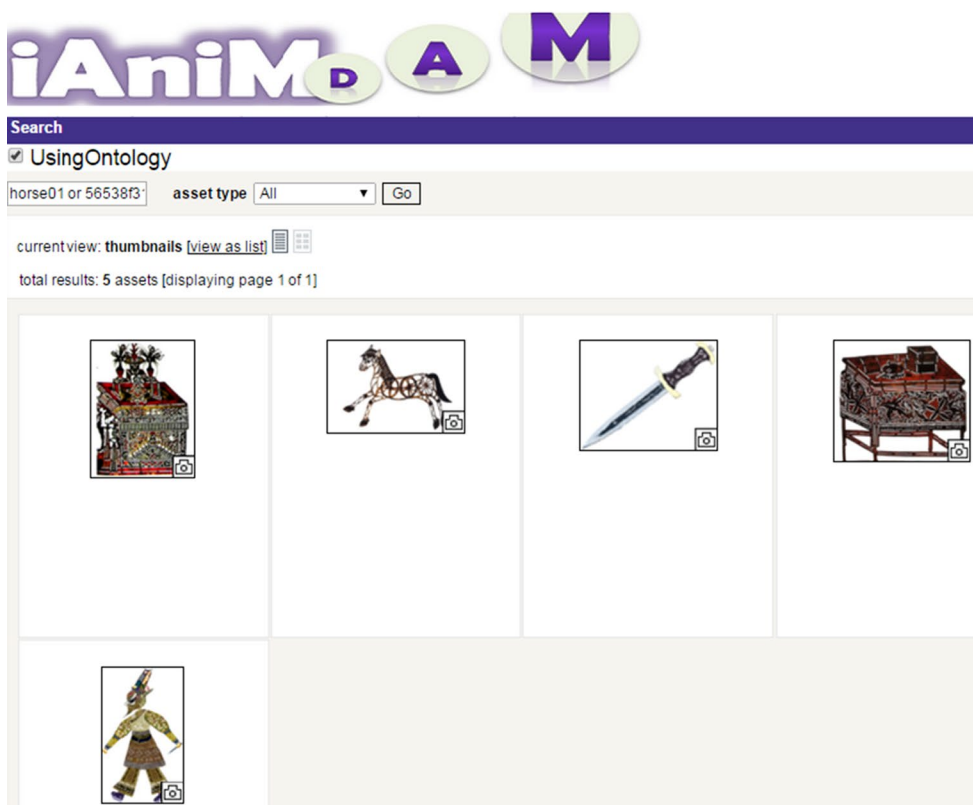
(b)

reasoning and inferencing with ontology-based annotations and rules, ontology-based retrieval makes possible an intelligent and productive manipulation of the digital assets in the context of Chinese traditional shadow play, which improves the retrieval effectiveness by recognizing the animator's semantic intent and contextual meaning.

Let us take the shadow puppet character "Jing Ke" as an instance to show all his relationship to other entities. "Jing Ke" is the hero of a famous Chinese traditional shadow play titled "The Emperor and the Assassin". As a brave fighter, "Jing Ke" attempted to assassinate the king of Kingdom Qin-"Qin Shi Huang" to avert the imminent conquest of his home country by Kingdom Qin. After intense fighting in the palace, however, the assassination attempt failed, and "Jing Ke" was killed on the spot.

Let us take the retrieval of the character "Jing Ke" for example. As illustrated in Fig. 8a, using traditional keywords marching search method, only the character "Jing Ke" itself was retrieved. However, for the character designer of the play, there is plenty of animation data existing to be referred to and reused as digital assets in the shadow puppet repository. If he wants to review all the subjects and objects that are related to the character "Jing Ke", through the object properties which are used to link to classes or individuals of another, such as "has Prop of", "has Scene of" and "has Music of", related entities will be retrieved after referring and reasoning, e.g. his weapon and the scene of the play. As the result, these retrieved related objects could provide more reasonable relevant feedback and wider choice to animators. From the result of ontological retrieval, as shown in Fig. 8b, we can find that not only the digital character "Jing Ke" itself is retrieved, other useful assets related to the hero are

also found from the knowledge base, which may satisfy the artist's potential requirements, such as the props affiliated to the play and his weapon used in the story.

Concept process of semantic reasoning is presented in Fig. 9.
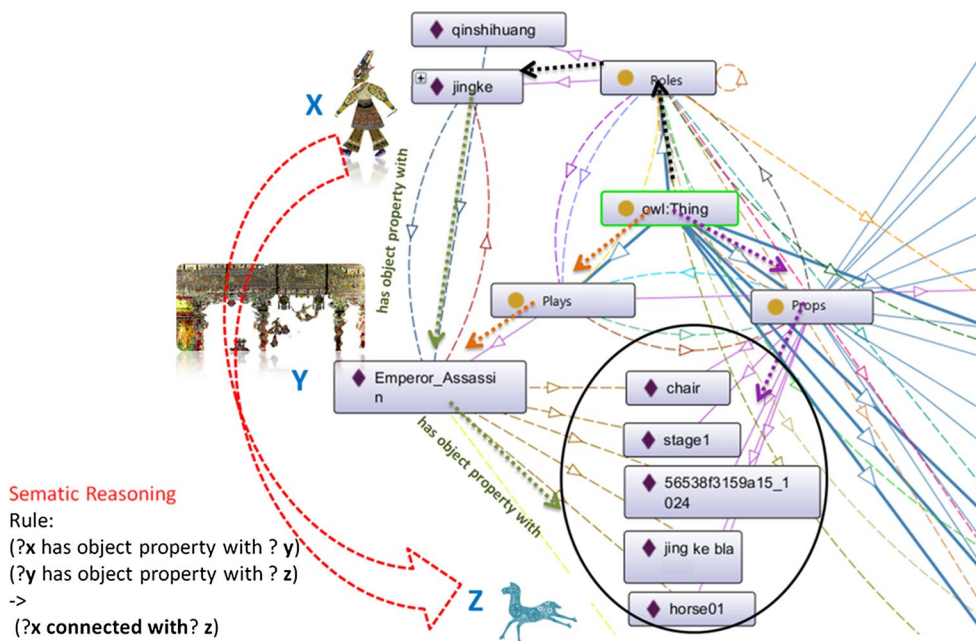
5.3 Result

5.3.1 System prototype

Finally, the prototype of the interactive traditional Chinese shadow play performance system has been generated. Shadow puppets' intensive motion is animated by player's hand movement and gestures through motion sensor device, as illustrated in Fig. 10a. The performance involves a set of hand gestures, which can only be successfully performed by specific hand movement and touch off puppets' related actions. Then, the pre-defined animations associated with puppets' particular actions can be successfully triggered. This prototype of a novel interactive Chinese traditional shadow play performance system provides players natural interaction/control and immersive experience.

Screenshots of generated digital shadow plays "The Emperor and the Assassin" and "Journey to the West" performed by an 8-year-old boy are shown in Figs. 10b and 11.
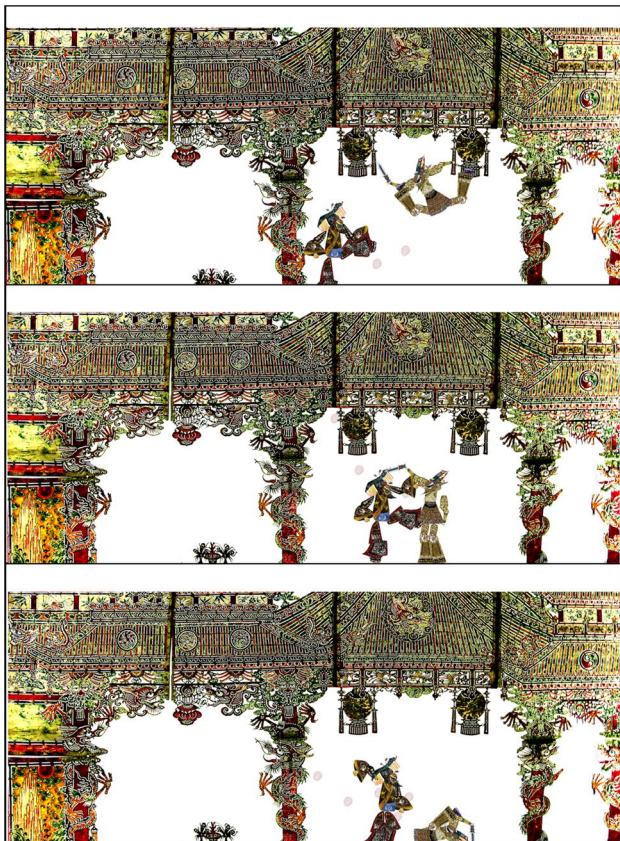
5.3.2 User test

To get qualitative feedback of our method, the feasibility verification has been conducted including both the user experience test of the ontology-based animation assets

Fig. 9 Semantic reasoning

Semantic Reasoning
Rule:
(?**x** has object property with ? **y**)
(?**y** has object property with ? **z**)
->
 (?**x** connected with? **z**)

(a)

(b)

Fig. 10 Scenario of digital shadow play "The Emperor and the Assassin"

retrieval and interaction comfort test of the prototype of the interactive shadow play system. We invited 7 target users to join the tests involving three animation designers and four 7–9-year-old young children (mean 8 years and 3 months old) with permission from their parents who were informed about the nature of the study and its purpose. Animation designers have good experience with computer animation design but little cognition of semantics or ontology

knowledge. Seven participants were divided into two groups to have two qualitative tests: the first group has 3 ontology-based assets retrieval tester (animation designers), and 4 young children took part in the second group for the interaction comfort test. After detailing the operation of system and 15-min training, two tests began. Qualitative tests were designed as follows:

Qualitative Test 1: It describes the feasibility of the ontology-based digital assets retrieval. We provided three participants with a list of animation data including different type of digital assets in the context of Chinese traditional shadow play, e.g. "Role", "Music", "Prop" and "Scene". Each of the animation designers was asked to use ontology-based retrieval to search 5 assets randomly and then make comparison with traditional keywords marching search method.

Qualitative Test 2: After the learning process, following the provided key story points, four children got engaged with the digital shadow puppets with great interests. Qualitative test 2 achieves testers' intuitive feelings on the interactive shadow play system through questionnaires. There are five criterions given—ease of use, handiness, naturalness, freedom of movement and effectiveness to describe the feelings with a scale of 1–5, indicating "Poor satisfaction" to "Excellently satisfied", and the higher the scores are, the more positive the feedback is.

5.3.3 Qualitative feedback

Qualitative Test 1: The three animation designers have input totally 15 different keywords to test the ontology-based retrieval and made comparison with traditional keyword-based marching method. All the three testers provided rather positive feedback. Not only the target data itself were retrieved, other useful assets related to the target were also found by conducting semantic reasoning. The relative wealth of retrieved material animation resources greatly satisfied the designers' potential creation requirements and also gave designers more creative inspiration.

Qualitative Test 2: Fig. 12 shows the mean score of each criterion about the interaction comfort. The ease of use, naturalness, freedom of movement, handiness illustrated positive feedback with score around four, especially the ease of use and naturalness. This is possibly due to the natural hand-gesture-based interaction which was elaborately designed for young players and provided intuitive interactive experience. However, the criterion of effectiveness received the lowest score which possibly was induced by the limitation of the device recognition accuracy. For example, folding the fingers or overlapping the hands could lead to depth sensor's false detection. Additionally, player's hand position relative to the location of the device is also crucial. Any less perfect position may affect the

Fig. 11 Scenario of digital shadow play "Journey to the West"

accuracy of hand tracking. From children's feedback, we also have observed that it is relatively easy for the player to use one hand to control one character. The novel interaction provides player a more immersive and interactive experience and enable him perform the story in intuitive and natural way. But manipulating two puppets simultaneously, each by one hand, are rather challenging. This kind of operation demands bimanual coordination skill. We hope this appealing interactive animation generation method may help people become intimate with their cultural heritage and promote traditional arts and culture among the young generation.

6 Discussion and conclusion

The construction process of interactive animation is complex because there are too many low-level details involved, for instance multimedia animation data management and data reuse, user-friendly interaction and integration. To provide a systematic and standardized semantic description, a semantic framework is constructed at an abstract high level in this paper. The ontological implementation based on the framework defined two domain-specified ontologies (Hand-Gesture-Based Interaction Ontology and Digital Chinese Shadow Puppetry Assets Ontology) to formalize the multimodal interaction method and the
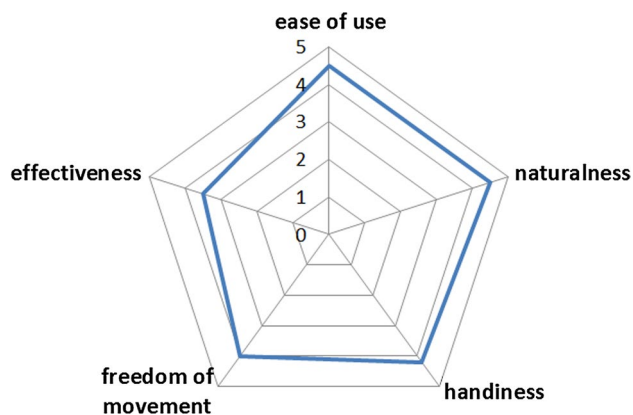
Fig. 12 Mean scores for each criterion of interaction comfort test (1 Poor, 2 Fair, 3 Average, 4 Good, 5 Excellent)

construction of the animation data assets repository, which finally leads to an interactive puppetry animation as a usage example.

Ontology is the core element of our research, by using which we provided a concept depiction for the complex animation production process, derived a semantic representation providing a more profound understanding of interaction and presented the specific domain knowledge of traditional Chinese shadow play art. Ontology is a high-level conceptual representation of the components in the animation production. The defined class hierarchy, object properties and data properties revealed how the low-level details involved in animation production are related to each other and how they collaborated and interacted with each other as well as the players and devices.

Our main goal is to utilize semantic/ontology concept to improve the reusability, extensibility and modularity of the interactive animation production and facilitate the development process. The proposed framework is flexible and extendable for many applications. In the usage example, we mainly worked with hand-gesture-based interactive digital shadow play, which provides a novel method to generate stylized traditional shadow play animation. This will contribute to the preservation of this cultural treasure of art. Based on the semantic framework, researcher can define various domain-specified ontologies and construct animation asset repositories depending on the context and application.

One limitation of the user interface comes from the leap motion sensor device, which can only detect hands and fingers within limited tracking area. Performance system will easily lose track if player's hand moves outside. To keep the consistency of the movement of the shadow puppetry, the puppet is kept still and stays where it is until the hand is detected again by leap motion when moving back. The user evaluation of our method is very brief and preliminary

examined in the test. We plan to examine more cases to affirm the conclusion at the next step. And also, a detailed evaluation of the semantic retrieval system will be carried out to illustrate the benefit of this data management approach further.

References

A free, open-source ontology editor and framework for building intelligent systems. http://protege.stanford.edu/. Retrieved 21 March 2016

Anthony M, Chandler T, Heiser E (2009) A framework for promoting high cognitive performance in instructional games. In: Interservice/industry training simulation and education conference (I/ITSEC), Orlando, FL

Apache Jena. https://jena.apache.org/. Retrieved 21 March 2016

Arndt R, Troncy R, Staab S, Hardman L (2009) Comm: a core ontology for multimediaannotation. In: Handbook on ontologies. Springer, Berlin, Heidelberg, pp 403–421

Bilasco IM, Gensel J, Villanova-Oliver M, Martin H (2005) On indexing of 3D scenes using MPEG-7. In: Proceedings of the 13th annual ACM international conference on multimedia. ACM, pp 471–474

BinSubaih A, Maddock S, Romano D (2005) Game logic portability. In: Proceedings of the 2005 ACM SIGCHI international conference on advances in computer entertainment technology. ACM, pp 458–461

Borst WN (1997) Construction of engineering ontologies for knowledge sharing and reuse, Universiteit Twente

Chang PHM, Chien YH, Kao ECC, Soo VW (2005) A knowledge based scenario framework to support intelligent planning characters. In: Intelligent virtual agents. Springer, Berlin, Heidelberg, pp 134–145

Chen F (2015) Visions for the masses: Chinese shadow plays from shaanxi and shanxi. Pacific Science 201

De Boeck J, Raymaekers C, Coninx K (2006) Comparing NiMMiT and data-driven notations for describing multimodal interaction. In: Task models and diagrams for users interface design. Springer, Berlin, Heidelberg, pp 217–229

Dubin D, Jett J (2015) An ontological framework for describing games. In: Proceedings of the 15th ACM/IEEE-CE on joint conference on digital libraries. ACM, pp 165–168

Falcidieno B (2004) Aim@ shape project presentation. In: Shape modeling applications, proceedings. IEEE, p 329

Falcidieno B, Spagnuolo M, Alliez P, Quak E, Vavalis M, Houstis C (2004) Towards the semantics of digital shapes: the AIM@ SHAPE approach. In: EWIMT

Flotyński J, Walczak K (2013a) Conceptual semantic representation of 3d content. In: Business information systems workshops. Springer, Berlin, Heidelberg, pp 244–257

Flotyński J, Walczak K (2013b) Semantic modelling of interactive 3d content. In: Proceedings of the 5th joint virtual reality conference. Eurographics association, pp 41–48

Flotyński J, Walczak K (2013c) Semantic multi-layered design of interactive 3d presentations. In: 2013 federated conference on computer science and information systems (FedCSIS). IEEE, pp 541–548

Flotyński J, Walczak K (2014) Multi-platform semantic representation of interactive 3d content. In: Technological innovation for collective awareness systems. Springer, Berlin, Heidelberg, pp 63–72

Funkhouser T, Min P, Kazhdan M, Chen J, Halderman A, Dobkin D et al (2003) A search engine for 3D models. ACM Trans Graph (TOG) 22(1):83–105

Ghannem A (2014) Characterization of serious games guided by the educational objectives. In: Proceedings of the second international conference on technological ecosystems for enhancing multiculturality. ACM, pp. 227–233

Google 3D Warehouse. https://3dwarehouse.sketchup.com/. Retrieved 21 March 2016

Gruber TR (1993) A translation approach to portable ontology specifications. Knowl Acquis 5(2):199–220

Grüninger M, Fox MS (1995) Methodology for the design and evaluation of ontologies, Workshop on Basic Ontological Issues in Knowledge Sharing

Gutierrez M, Thalmann D, Vexo F (2005) Semantic virtual environments with adaptive multimodal interfaces. In: Multimedia modelling conference, 2005. MMM 2005. Proceedings of the 11th international. IEEE, pp 277–283

Gutiérrez M, García-Rojas A, Thalmann D, Vexo F, Moccozet L, Magnenat-Thalmann N, Mortara M, Spagnuolo M (2007) An ontology of virtual humans: incorporating semantics into human shapes. Vis Comput 23(3):207–218

Kazhdan M, Funkhouser T, Rusinkiewicz S (2003) Rotation invariant spherical harmonics representation of 3D shape descriptors. In: Symposium on geometry processing, vol 6, pp 156–164

Leino OT (2010) What erotic tetris has to teach serious games about being serious? Design implications of an experiential ontology of game content. In: Meaningful play 2010 conference proceedings, USA

Li Z, Karthik K (2007) Ontology-based design information extraction and retrieval. AI EDAM Artif Intell Eng Des Anal Manuf 21(2):137–154

Liang H, Chang J, Kazmi IK, Zhang JJ, Jiao P (2015a) Puppet narrator: utilizing motion sensing technology in storytelling for young children. In: 2015 7th international conference on games and virtual worlds for serious applications (VS-Games). IEEE, pp 1–8

Liang H, Chang J, Deng S, Chen C, Tong R, Zhang JJ (2015b) Exploitation of novel multiplayer gesture-based interaction and virtual puppetry for digital storytelling to develop children's narrative skills. In: Proceeding VRCAI '15 proceedings of the 14th ACM SIGGRAPH international conference on virtual reality continuum and its applications in industry. ACM, pp 63–72

Mansouri H (2005) Using semantic descriptions for building and querying virtual environments. Licentiate's thesis, Vrije Universiteit Brussel

Mirbakhsh N, Didandeh A, Afsharchi M (2010) Concept learning games: the game of query and response. In: 2010 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology (WI-IAT), vol 2. IEEE, pp 234–238

Ohbuchi R, Yamamoto A, Kobayashi J (2007) Learning semantic categories for 3D model retrieval. In: Proceedings of the international workshop on workshop on multimedia information retrieval. ACM, pp 31–40

Ontology for Media Resources 1.0. http://www.w3.org/TR/mediaont-10/. Retrieved 18 March 2016

Otto KA (2005) The semantics of multi-user virtual environments. In: Proceedings of the workshop towards semantic virtual environments

Parkkila J, Hynninen T, Ikonen J, Porras J, Radulovic F (2015) Towards interoperability in video games. In: Proceedings of the 11th biannual conference on Italian SIGCHI chapter. ACM, pp 26–29

Ramaprasad A, Papagari SS (2009) Ontological design. In: Proceedings of the 4th international conference on design science research in information systems and technology. ACM, p 5

Ruminski D, Walczak K (2014) Semantic contextual augmented reality environments. In: 2014 IEEE international symposium on mixed and augmented reality (ISMAR). IEEE, pp 401–404

Shadow play. http://www.chinatraveldesigner.com/travel-guide/culture/chinese-opera/shadow-play.htm. Retrieved 21 March 2016

SPARQL Query Language for RDF. https://www.w3.org/TR/rdf-sparql-query//. Retrieved 21 March 2016

Studer R, Benjamins VR, Fensel D (1998) Knowledge engineering: principles and methods. Data Knowl Eng 25(1):161–197

SWRL: a semantic web rule language. https://www.w3.org/Submission/SWRL/. Retrieved 21 March 2016

Tangelder JW, Veltkamp RC (2008) A survey of content based 3D shape retrieval methods. Multimed Tools Appl 39(3):441–471

Teixeira JSF, Sá EDJV, Fernandes CT (2008) A taxonomy of educational games compatible with the LOM-IEEE data model. In: Proceedings of interdisciplinary studies in computer science SCIENTIA, pp 44–59

Thalmann D, Farenc N, Boulic R (1999) Virtual human life simulation and database: why and how. In: 1999 International symposium on database applications in non-traditional environments, 1999. (DANTE'99) Proceedings. IEEE, pp 471–479

Tutenel T, Smelik RM, Bidarra R, de Kraker KJ (2009) Using semantics to improve the design of game worlds. In: AIIDE

World Wide Web Consortium (2012) OWL 2 web ontology language document overview