eHeritage of Shadow Puppetry: Creation and Manipulation

Min Lin1, Zhenzhen Hu2, Si Li1, Meng Wang1, Richang Hong2, Shuicheng Yan1
1ECE Department, National University of Singapore
2School of Computer and Information, Hefei University of Technology
{mavenlin, huzhen.ice, eric.mengwang, hongrc.hfut}@gmail.com,
{dcslius, eleyans}@nus.edu.sg

ABSTRACT
To preserve the precious traditional heritage Chinese shadow puppetry, we propose the puppetry eHeritage, including a creator module and a manipulator module. The creator module accepts a front view face image and a profile face image of the user as input, and automatically generates the corresponding puppet, which looks like the original person and meanwhile has some typical characteristics of traditional Chinese shadow puppetry. In order to create the puppet, we first extract the central profile curve and warp the reference puppet eye and eyebrow to the shape of the frontal view eye and eyebrow. Then we transfer the puppet texture to the real face area. The manipulator module can accept the script provided by the user as input and automatically generate the motion sequences. Technically, we first learn atomic motions from a set of shadow puppetry videos. A scripting system converts the user’s input to atomic motions, and finally synthesizes the animation based on the atomic motion instances. For better visual effects, we propose the sparsity optimization over simplex formulation to automatically assemble weighted instances of different atomic actions into a smooth shadow puppetry animation sequence. We evaluate the performance of the creator module and the manipulator module sequentially. Extensive experimental results on the creation of puppetry characters and puppetry plays well demonstrate the effectiveness of the proposed system.

Categories and Subject Descriptors
J.5 [Arts and Humanities]: Arts, fine and performing

Keywords
Chinese shadow puppetry; face rendering; sparsity optimization over simplex; animation

1. INTRODUCTION
Shadow puppetry has a long history in China, Indonesia, India, Greece, etc. as a form of entertainment for both children and adults, and is also popular in many other countries around the world. We focus on Chinese shadow puppetry in this work. Chinese shadow puppetry, as shown in Figure 1(a), is a traditional artistic form of theater performance with colorful silhouette figures. These figures are produced with puppets made of leather or paper. A traditional Chinese shadow puppet is composed of a head part shown in Figure 1(b) and a body part shown in Figure 1(c). As an ancient form of storytelling, sticks and flat puppets are manipulated behind an illuminated background to create moving pictures. By moving both the puppets and the light source, various effects can be achieved. But in the 21st century, shadow puppetry in China is on a steep and fast decline. Audiences and apprentices are evaporating at an alarming rate. To protect this ancient artistic heritage, China’s State Council put Chinese shadow puppetry in the first list of National Intangible Cultural Heritage of China in 2006, and United Nations Educational, Scientific and Cultural Organization (UNESCO) listed the artistic form in Intangible Cultural Heritage in 2011.

Nowadays, the preservation of culture heritage attracts growing attentions of the world. Our artistic motivation is to develop a system by applying the latest multimedia technologies to aid preservation, interpretation, and dissemination of this ancient cultural heritage. To attract more people to be interested in shadow puppetry, we design two puppetry modules, including a creator module and a manipulator module that people all over the world can experience the fun of the puppet creating and performing by themselves.

The input of the creator module is two face images. One of the most important advantages of our creator module is personalization. The making process of a puppet includes seven steps which are all complex and ingenious. Our creator module can automatically generate anyone’s personalized puppet, and the process is almost immediate. We pay special attention to keeping the characteristics of the traditional puppet during the automatic creator process. For example, the puppet’s face has long narrow eyes, a small mouth and a straight bridge of nose as shown in Figure 1(b). Technically, we extract the central profile curve from the profile face image and warp the reference puppet eye into.
the frontal view eye simultaneously. Then we transfer the puppet texture as the texture of the profile curve. One male and one female puppet generated by our creator module are shown in Figure 1(d) and Figure 1(e).

We also design a manipulator module. For this module, we aim to preserve the performance pattern of Chinese shadow puppetry during the shadow puppetry manipulation process. There are several basic puppet motion patterns (denoted as atomic actions afterwards), such as walk, dance, fight, nod, laugh, etc. Besides, as shown in Figure 1(c), the real puppet is controlled with three sticks which are fixed on the puppet’s neck and two hands separately, and the motion pattern of other puppet parts is affected by gravity. Therefore, we need to simulate all the atomic actions in the puppet style. There are some existing research works on shadow puppets focusing on the user’s body interaction with the virtual puppet, i.e., the puppet’s motion imitates the user’s motion 2,3,4. But all of them cannot preserve the puppet’s specific motion style. To the contrary, our module can convey this traditional artistic charm completely. Our manipulator module can directly accept text scripts as input and display the specified motions accordingly. The interface is shown in Figure 1(f). For manipulation, we identify atomic motions for the animation and collect instances from a set of shadow puppetry videos. A scripting system converts the user input into atomic motions, and finally synthesizes the animation using the collected instances.

The organization of the rest of this paper is as follows. In Section 2 we provide a review of related work. Section 3 makes an overview of our puppetry module: creator and manipulator. Next, in Section 4 and 5, more detailed step by step introduction of the modules is presented. The experimental results are shown later in Section 6. Finally, we conclude the paper in Section 7.

2. RELATED WORK

Our shadow puppetry system is built on several areas of related work.

2.1 eHeritage

Recently, for the propose of preserving cultural heritage through the application of advanced computing technologies, eHeritage becomes a hot research topic. The related literature is mainly divided into two categories: tangible cultural heritages and intangible cultural heritages. Tangible cultural heritages include buildings and historic places, monuments, artifacts, etc. Anna Paviotti et al. 5 dealt with the problem of estimating the lighting field of the multispectral acquisition of frescoes by a variational method. Lior Wolf et al. 6 studied the task of finding the joins of the Cairo Genizah which is a precious collection of mainly Jewish texts. Tao Luo et al. 7 presented a multi-scale framework to generate 3D line drawing for archaeological illustration. Intangible culture heritage includes traditional festivals, oral traditions, oral epics, customs, ways of life, traditional crafts, etc. Markus Seidl et al. 8 proposed a detection of gradual transitions in historic material. Anupama Mallik et al. 9 studied on the preservation of Indian classical dance. The shadow puppetry belongs to intangible cultural heritages.

2.2 Face Rendering

Digital image processing provides a solid foundation for building artistic rendering algorithms. All image-based artistic rendering approaches utilize image processing operations in some forms to extract information or synthesize results 10. Non-photorealistic rendering focuses on enabling a wide variety of expressive styles of digital arts. Among these arts, cartoon and paper cutting are the most related to our work.

Cartoon: A cartoon is a form of two-dimensional illustrated visual art with a typically non-realistic or semi-realistic drawing or painting. CharToon system 11 provides special skeleton-driven components, an extensive set of building blocks to design faces and the support to reuse components and pieces of animations. Chen et al. 12 explored a Pictoon system that allows users to create a personalized cartoon and animation from an input face image based on sketch generation and cartoon stroke rendering.

Paper Cutting: Artistic paper cutting is also a traditional and popular Chinese decorative art which, usually

\[ \text{http://www.unesco.org/new/en/cairo/culture/tangible-cultural-heritage/} \]

\[ \text{http://www.unesco.org/new/en/cairo/culture/intangible-cultural-heritage/} \]
in a very concise two-tone form, has its unique beauty of expressive abstraction. Xu et al. \cite{19} generated arrangements of shapes via a multilayer thresholding operation to compose digital paper cutting designs. M. Meng et al. \cite{14} rendered paper cutting images from human portraits. They localized facial components and used pre-collected representative paper cutting templates. Then they obtained a synthesized paper cutting image by matching templates with the bottom-up proposals.

Our work differs from cartoon and paper cutting in that we only focus on the face rendering of shadow puppetry, which has its unique characteristics. Thus, very tailored image processing techniques are required.

2.3 Puppetry Animation

Developing an interactive and user-friendly interface for people of all skill levels, to create animation is a long-standing problem in computer graphics. Recently, a few works have been conducted on digital puppetry. As a visualization tool for traditional cinematic animation, digital puppetry transforms the movements of a performer to the actions of an animated character to provide live performance such as \cite{15}. Shadow puppets are also used to produce animated films, but producing an animated film with shadow puppets, frame by frame, is laborious and time-consuming. The solution of animation performed by two-dimensional puppets appears only recently. Hsu et al. \cite{16} introduced a motion planning technique which automatically generates the animation of 2D puppets. Barnes et al. \cite{17} created a video-based animation interface. Users first create a cast of physical puppets and move these puppets to tell a story. They tracked the motions and rendered an animate video. Tan et al. \cite{18} presented a method for interactive animation of 2D shadow play puppets by real-time visual simulating using texture mapping, blending techniques, and lighting and blurring effects. Kim et al. \cite{19} controlled 3D avatars to create user-designed peculiar motions of avatars in real-time using general interfaces. ShadowStory \cite{20} was a project created by Fei et al. for digital storytelling inspired by traditional Chinese shadow puppetry. The system allows children to design their own puppets and animate them with a tablet PC. Pan et al. \cite{21} presented a 2D shape deformation of the triangulated cartoon which is driven by its skeleton and the animation can be obtained by retargeting the skeleton joints to the shape.

There are some websites where users can entertain with the puppetry, such as Puppet Parade\footnote{http://design-io.com/projects/PuppetParadeCinekid/} an interactive puppetry installation that animates puppets by tracking the arms of the puppeteers, or We Be Monsters\footnote{http://wearemonsters.net/} a collaborative puppet art installation that tracks multiple skeletons to animate a puppet. Although shadow puppetry is not a focus on these websites, the body motion is used to control the puppets. All these puppetry methods do not very accurately consider the unique motion style of shadow puppetry.

With the development and increasing popularity of Microsoft Kinect Camera, many researchers begin to explore how to achieve human-puppetry interaction via Kinect. Robert Held et al. \cite{22} presented a 3D puppetry system that allows users to quickly create 3D animations by performing the motions with their own familiar, rigid toys, props, and puppets. During a performance, the puppeteer physically manipulates these puppets in front of a Kinect depth sensor. Leite et al. \cite{23} presented an anim-actor technique. Anim-actor is a real-time interactive puppets control system using low-cost motion capture based on body movements of non-expert artists. Zhang et al. \cite{4} proposed a general framework for controlling two shadow puppets, a human model and an animal model. Cooper S. Yoo et al. \cite{24} created a tangible stage interface that can be used to control marionette without wires based on Kinect. Kinect camera can provide very accurate depth estimation which can greatly facilitate the animation process, but also limits the system.

3. OVERVIEW

In this section, we give an overview of the shadow puppetry system. As shown in Figure 2, the whole system contains two modules: creator module and manipulator module.

The first part is creator module. Since the real puppet face usually contains a profile (including forehead, nose, mouth, etc) and a frontal view eye, we require users to input one frontal view image and another profile view face image. First, we extract the eye and eyebrow from the frontal view image, and the puppet eye and eyebrow are deformed into the frontal view eye and eyebrow shape. Then the profile curve of the real profile face is extracted and deformed. Finally, the texture is transferred from a certain sample puppet to our generated puppet.

The second part is manipulator module. The objective of this part is to create a system that takes in a user script and outputs an animation sequence. To build this system, we collect puppetry show videos and manually define atomic motions as building blocks of the animation. Instances of the atomic motions are extracted from these videos. The script input by the user is first interpreted into an atomic motion sequence, which is then synthesized by weighted mean of the instances as later introduced.

4. MODULE I: PUPPETRY CREATOR

The design of Chinese shadow puppetry figures follows traditional aesthetics. There are mainly four kinds of puppetry faces: Sheng (male roles), Dan (female roles), Jing (roles with painted faces) and Chou (comic characters). These roles are mostly from classical plays and appearances are for-

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{system_overview.png}
\caption{Illustration of the whole system framework. The upper panel is the creator module and the lower panel is the manipulator module.}
\end{figure}

\begin{itemize}
\item \cite{18} presented a method for interactive animation of 2D shadow play puppets by real-time visual simulating using texture mapping, blending techniques, and lighting and blurring effects.
\item \cite{19} controlled 3D avatars to create user-designed peculiar motions of avatars in real-time using general interfaces.
\item \cite{20} presented a 2D shape deformation of the triangulated cartoon which is driven by its skeleton and the animation can be obtained by retargeting the skeleton joints to the shape.
\item \cite{21} presented a 2D shape deformation of the triangulated cartoon which is driven by its skeleton and the animation can be obtained by retargeting the skeleton joints to the shape.
\end{itemize}
4.1 Face Alignment

Given a frontal view face image and a profile face image of the same person, we utilize the frontal view eye & eyebrow and the profile curve to create the head of a user-specific puppetry figure head.

Face Alignment: We process the face alignment of the frontal face and profile face separately. For a frontal view face, it is aligned by a commercial frontal face alignment algorithm\(^9\). The alignment of a profile face is performed based on the unified model proposed by Zhu et al.\(^9\) for face detection, pose estimation, and landmark estimation in real-world, cluttered images. We show the aligned results of frontal and profile view respectively in Figure 3. After face alignment, we get the locations of a set of keypoints on face. Then based on these points, we extract the profile curve from the profile face and rotate the curve for direction adjustment. Similarly, eye and eyebrow areas are extracted from the frontal image.

4.2 Eye & Eyebrow Warping

In Chinese shadow puppetry, most of the figures’ faces are in profile view. For describing features of characters, eyes and eyebrows are abstracted and exaggerated in an artistic way. For example, the eye of a puppet is stretched and the eyebrow is bent. The topology of the puppet eye is similar to the frontal view eye. To guarantee the similarity of the output figure face, given a target frontal view face image and a puppet face, we warp the puppet eye and eyebrow into the real eye and eyebrow shape.

Most state-of-the-art warping methods\(^26,27\) assume that the user provides explicit correspondences between the anchor points of the source and the target shape, and other points are mapped according to their relative positions with respect to the anchor points. Barycentric coordinates of triangles provide a convenient way to linearly interpolate the points within a triangle. Given a planar triangle \([v_1, v_2, v_3]\), any point \(v\) inside it has the unique Barycentric coordinates \([w_1, w_2, w_3]\) and:

\[
\frac{w_1v_1 + w_2v_2 + w_3v_3}{w_1 + w_2 + w_3} = v. \tag{1}
\]

The Barycentric coordinates are used as invariants to build the correspondences between two topologically equivalent polygons, characterized by the vertexes\(^27\).

In our case, we first extract the contour points of puppet eye \(C = \{c_1, \ldots, c_n\}\) by annotating landmarks and detect the contour points of real eye \(P = \{p_1, \ldots, p_m\}\) automatically. In our setting, \(n = m = 10\).

We consider puppet eye and real eye as two topologically equivalent polygons, characterized by \(C\) and \(P\), and then

\(^9\)http://www.omron.com/r_d/coretech/vision/okao.html

Figure 3: Aligned frontal and profile view faces.

4.3 Profile Curve Generation

Central profile curve is an important geometric feature. To generate an appropriate puppet profile, we will consider two aspects: 1) this profile should look like the input face and 2) this profile should have puppetry style. We keep the most unique parts of the profile curve and transform them into one with puppetry style.

4.3.1 Profile Abstraction

The central profile curve is highly discriminative. Some 3D facial research works use central profile curve as a significant feature for face matching\(^28\) and face recognition\(^29\). To obtain a discriminative and puppet-like personalized puppet, we generate a profile curve which has the uniqueness of the input face and the characteristics of the reference puppet face. In order to estimate the uniqueness of
a real profile face, we collect a 90 degree profile face dataset including 500 images from the well-known MultiPIE benchmark [30].

Given a 90 degree profile face, we extract the edge of a profile face by using extended Difference of Gaussian (DOG) algorithm [31] and based on the face alignment result, we obtain the profile curve. A profile curve is separated into four parts: forehead, nose, mouth and jaw. We aim to preserve unique parts and replace others with the corresponding parts of the reference puppet, which are manually annotated offline.

First, we process this curve area by binary morphology tools and extract the binary skeleton line with only one pixel width [22], and then for each part, we extract histogram of oriented gradients (HOG) [33] feature as a vector. For each part, based on the MultiPIE dataset, we calculate the average distance $d$, and the importance of the part from the input face is measured by the ratio of $d$, which is the mean of its distances to the K-Nearest-neighbors ($k = 3$ in our implement), over $d$. The two most unique parts are kept and the other two parts are replaced by the corresponding parts of shadow puppetry profile.

From the profile image as shown in Figure 5(a), we get the direction adjusted profile curve in Figure 5(b). The parts in green circles are the two most unique parts and retained as shown in Figure 5(c). The parts in gray circles are replaced with the parts in red circles in Figure 5(e) extracted from Figure 5(d). Finally, we recombine these selected parts and generate a puppet profile curve in skeleton form.

4.3.2 Profile Texture Transfer

After previous steps, we obtain a profile skeleton which is like both the real face and the puppet face. Then we dilate this line till the same profile width of the reference puppet as shown in Figure 6(a). To be more like a puppet, we transfer the leather texture to the profile area. Efros et al. [34] presented a simple image quilting algorithm of generating novel visual appearance in which a new image is synthesized by stitching together small patches of existing images. Given a piece of leather sample in Figure 6(b), the texture for the profile curve can be synthesized by the image quilting method [34]. In Figure 6(c), we show an example of texture transfer result.

4.4 Post Processing

In shadow puppetry play, there are a lot of vivid props, including architecture, furniture, plants and animals, designed with distinct dynamic features. Some representative examples are shown in Figure 7(a). To make the roles more artistically charming, puppets are decorated with exquisite headdresses, which are classified into male style and female style. We can match the generated puppet with the proper headdresses, and some results are shown in Figure 7(b).

5. MODULE II: PUPPETRY MANIPULATOR

Traditional Chinese shadow puppet is created by hinging together nine parts: head, upper body, left & right arm,
The conformation of the puppet is controlled by assigning values to the 12 DOFs. Animation can thus be synthesized by feeding a sequence of DOFs. The method to generate DOF sequence is described in the next sections.

5.2 Motion System

5.2.1 Identification of Atomic Motions

Based on five videos of popular real world puppet show, i.e. “Tale of the White Snake Madam”, “Emperor Xuanzong of Tang”, “Yanfei Wonder Woman”, “Yang Saga”, “Green Dragon Sword”, it can be observed that the animation of puppets can be decomposed into smaller building blocks, which we call atomic motions. Eight atomic motions identified from the videos are listed in Table 1.

5.2.2 Data Preparation and Processing

After the atomic motions are identified, we manually extract video chunks that contain atomic motions from the puppet shows. For each of the atomic motions, we extract 50 video chunks which we call instances of that atomic motion. Each frame in the video chunks are labeled with 12 key points that describe the conformation of the puppet as shown in Figure 8(b). From the key points, we can easily calculate the DOF’s information for each frame. Thus each instance of the atomic motion is finally converted to an instance matrix, with each column being the DOFs of the corresponding frame.

However, shaking defect is observed when the DOFs are applied directly to our digital puppets, because the available videos are generally of low quality and the point labeled by hand will introduce a high frequency shaking defect, which compromises the annotation accuracy. To overcome this defect, we apply locally weighted scatterplot smoothing to the rows of the DOF matrix. Later experiments show that the smoothing improves the visual experience significantly.

After smoothing, the instances of the same atomic motion are then temporally realigned by linear interpolation. For example, the atomic motion \( A \) has \( n \) instances, each converted to a matrix \( A_i \). The \( A_i \) are linearly interpolated to have the same column number as the matrix with the most columns, denoted by \( A' \). We consider the weighted mean of the aligned instances as eligible variants of that atomic motion:

\[
A_{new} = A_1 w_1 + A_2 w_2 + \cdots + A_n w_n
\]

\[
\sum_{i=1}^{n} w_i = 1, \quad w_i > 0.
\]

Unlimited variants of the atomic motion can be generated which form a rich set of ingredients for building complex animations.

5.2.3 Recombination of Atomic Motions

Atomic motions captured from the videos have limited types. One way to increase their number is through recombination. Each atomic motion can be decomposed into independent parts. For example, the upper body motion is independent of the lower body motion. We can thus combine upper body motion of “bow” with lower body motion of “kneel” to form a new atomic motion “kowtow”.

5.2.4 Composite Motions

While atomic motion provides high flexibility for animation synthesis, composite motion offers convenience. Composite motion consists of sequential atomic motions. It is defined for easier usage of the scripting interface. For instance, composite motion “talk for 10 minutes” means “breathe” continuously and “point” sometimes for 10 minutes. In this case, it will take much more efforts if we input atomic motions one by one. The scripting system will handle the conversion of composite motions to atomic motions. Composite motion is an addition to atomic motion which gives users higher level control of the puppet.

5.3 Animation Generation

The animation of puppets can be decomposed to smallest building blocks called atomic motions. Based on atomic motions, composite motions are synthesized as larger building blocks. In this work we render the animation using browser based webgl for ease of sharing with others, each puppet has a corresponding javascript object. Through the scripting interface, manipulators of the digital puppet invokes different functions on the object. For each atomic or composite motion, we define a corresponding function. The manipulator can easily control the puppet by invoking the function with parameters like time duration and spatial distance. For example, puppet.walk(10,15) tells puppet to walk 10 units in the defined 2D world in 15 frames’ time. Through script, the manipulator can also place props like chair to specific locations, the locations of the props can then be used in the script. The manipulator can either do scripting on-line by inputing one motion after the other or off-line by writing done all the scripts and feed to the puppet. The on-line system is mainly for the user to experience the individual motions. For an entire animation sequence generation, off-line is preferred because it offers better control of timing. In the rest part of this section, we will present the method for off-line animation generation.

The off-line animation generation system takes in a sequence of atomic & composite motions as input from a scripting interface. A preprocessing step on the input motion sequence decomposes the composite motions into atomic motions, which makes the motion sequence monotonously atomic. The DOF sequence is finally created by converting each atomic motion into DOFs. As mentioned in Section 5.2.2, atomic motions are generated by weighted means of their instances. Reusing a limited number of the instances makes the animation monotonous; by taking different weight

<table>
<thead>
<tr>
<th>Atomic Motion</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>jump</td>
<td>the whole body move forward for some distance, which represents walking in shadow puppet show</td>
</tr>
<tr>
<td>bow</td>
<td>To lower the head or upper body as a social gesture</td>
</tr>
<tr>
<td>sit</td>
<td>Sitting on the chairs in the scene</td>
</tr>
<tr>
<td>kneel</td>
<td>Kneel down on the ground</td>
</tr>
<tr>
<td>point</td>
<td>A hand gesture frequently appearing during speech, indicates the puppet is talking.</td>
</tr>
<tr>
<td>swing</td>
<td>The forward and backward motion of hands</td>
</tr>
<tr>
<td>breathe</td>
<td>Undulation of upper body, in puppet show, this motion indicates the puppet is talking</td>
</tr>
<tr>
<td>turn</td>
<td>Flipping the puppet about the y axis of the scene, changing its facing direction</td>
</tr>
</tbody>
</table>

Table 1: Atomic motions
vector, infinite variants of one atomic motion can be generated from its limited instances. The aesthetic of the animation is greatly influenced by the smoothness of the animation. The visual experience is discounted if two adjacent frames differ a lot. Thus, minimizing the position gap between atomic motions is the criterion in selecting the weight vectors.

### 5.3.1 Weights Selection

The smoothness of the animation sequence can be measured by the position gap between consecutive atomic motions, namely the distance from the last frame of one atomic motion to the first frame of the next. Smoothing the whole sequence can be achieved by optimizing the following problem:

\[
\min_{\{w_i\}} \sum_{i=1}^{N-1} \|F_i w_{i+1} - L_i w_i\|_2^2 \tag{3}
\]

subject to \( \sum_j w_{ij} = 1, \; w_{ij} \geq 0, \; i \in \{1, 2, \ldots, N\} \).

For better explanation, we introduce the instance matrix \( M^k_j \), which means the \( j \)th instance of the atomic motion \( k \). \( k \) is in the set of atomic motions listed in Table 1; \( k(i) \) denotes the atomic motion type of the \( i \)th atomic motion in the motion sequence. \( F_i \) is formed by taking the first columns of \( M^{k(i)}_j \) for each \( j \). \( L_i \) is formed by taking the last columns. \( w_i \) is the weight vector for the \( i \)th atomic motion. For an given atomic motion, the weights are constrained to be non-negative, and sum to one.

Due to the imperfectness of instances realignment, weighted recombination of the instances has shaking defects. We impose a sparsity constraint on the weight vectors, so that a minimum number of instances are selected. The following optimization problem is formed:

\[
\min_{\{w_i\}} s\{w_i\} + \lambda \sum_i \|w_i\|_0 \tag{4}
\]

subject to \( \sum_j w_{ij} = 1, \; w_{ij} \geq 0, \; i \in \{1, 2, \ldots, N\} \).

Generally, the \( \ell_0 \)-norm is relaxed into \( \ell_1 \)-norm for computational efficiency. However, the \( \ell_1 \)-norm heuristic does not work here because the constraint makes \( \ell_1 \)-norm of the weight vector equal to 1, namely constant.

We utilize another relaxation method to handle this sparsity on probability simplex problem proposed in [35]. The \( \ell_0 \)-norm can be lower bounded by reciprocal of infinity norm when the corresponding \( \ell_1 \)-norm is a constant. For any \( x \) constraint to probability simplex:

\[
\|x\|_1 = \sum_{i=1}^n |x_i| \leq \|x\|_0 \max_i |x_i| \leq \|x\|_0 \|x\|_{\infty} \tag{5}
\]

Because \( \|x\|_1 = 1 \),

\[
\frac{1}{\|x\|_{\infty}} \leq \|x\|_0. \tag{6}
\]

Then we resort to solve the following problem:

\[
\min_{\{w_i\}} s\{w_i\} + \lambda \sum_i \frac{1}{\|w_i\|_{\infty}} \tag{7}
\]

### 6. EXPERIMENTS

In this section, we first evaluate the performance of the creator module and manipulator module sequentially. Then we comprehensively show the results of the our system in a so-called “love story with puppetry” scenario to show the combination effectiveness of creator results and manipulator results.

For the user studies in this section, 30 subjects (7 females and 23 males) ranged from 22 to 40 years old (\( \mu=27.3 \), \( \sigma=3.9 \)) are invited to participate in our experiments.

#### 6.1 Results on Creator Module

##### 6.1.1 Quantitative Results: User Study

Here we discuss the properties of the creator puppet in the following three aspects: aesthetic, puppetry-alike and discrimination.

**Aesthetic:** Whether the result looks elegant?
Table 2: ANOVA analysis of comparing smoothed and original atomic motions

<table>
<thead>
<tr>
<th>F-statistics</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1025.49</td>
<td>5.02281e-121</td>
</tr>
</tbody>
</table>

**Puppetry-alike**: Does the generated face keep the characteristics of puppet?

In these two questions, we set five values for subjects to choose: excellent, good, ordinary, weak and poor. We present 10 creator puppet images to subjects and the results are shown in Figure 10. We can see that our puppet module has achieved very high score in both evaluation metrics.

**Discrimination**: Could you recognize the people from the generated face? We evaluate this index by several single-choice questions. For example, we show a puppetry of Emma Watston to the subjects, in the same time, we show four face images: one is Emma Waston and the other three are random choices. An example of user study interface is shown in Figure 10. The subjects choose which face image among the four candidates most looks like the puppet. We test 10 creator results and summarize the subjects’ feedback. The average precision is 71.1%. The precision is not quite high because the generated puppet images are all in profile view, but the candidates are in frontal view. However, the precision of 71.1% is much higher than random guess (25%).

6.1.2 Qualitative Results

Here we show some result examples of the creator module. As shown in Figure 11, the first column are input frontal view face images, the second column are input profile face images of the same person. The creator results are shown in the third column and results decorated with headdress are shown in the forth column.

6.2 Results on Manipulator Module

6.2.1 Atomic Motion

Here we evaluate the result of applying locally weighted scatterplot smoothing to the instance matrices of atomic motions. For each type of atomic motion, one instance is randomly selected and smoothed. The subjects are presented both the original and smoothed version of the atomic motion instances. For each original and smoothed pair of the instances, the subjects are required to evaluate their smoothness on the 5-point Likert scale without prior knowledge which one is smoothed. One randomly selected instance for all eight types of atomic motions are evaluated. We show the notched box plot of the result in Figure 12(a). One way ANOVA analysis is also performed on the data shown in Table 2. The results show that locally weighted scatterplot can significantly improve the smoothness of the atomic motions.

6.2.2 Animation Smoothness Evaluation

The smoothness of generated animation is not only dependent on the smoothness of atomic motion itself, but also on the transition of adjacent atomic motions. To evaluate the weight selection method, we generated five animation sequences each about 30s in length consisting of random atomic motions. Two versions of the animations are generated, one with weights selected using our method described in Section 5.3 and the other with randomly selected weights. Subjects are given the same options as in the evaluation of atomic motion smoothness. Figure 13(a) and Figure 13(b) compare the transitions selected from the animation sequences generated with smoothness optimized weights and random weights. It can be observed the position gap between consecutive atomic motions is smaller when smoothness optimized weight is used.

The notched box plot of the result is shown in Figure 12(b), and the one way ANOVA analysis is shown in Table 3. The results show that optimized weights using our method is significantly better in transition smoothness than randomly selected weights.

6.2.3 Short Video Evaluation

Two video chunks about one minute in length is selected from the puppetries “Tale of the White Snake Madam” and “Green Dragon Sword”. Using the scripting interface, we reproduced those two videos. The subjects in this study are required to watch the original and reproduced videos, then answer questions regarding three aspects of the synthesized videos. The questions are answered on a 5-point Likert scale from “very bad” (1) to “very good” (5). The mean and standard deviation of the scores are calculated and displayed in Table 4. From the results, conclusion can be made that the puppetries generated from our system preserve the motion pattern of real puppetry, and successfully convey the information to the subjects. The subjects are satisfied with the resolution of the generated puppetry, which is expected because videos captured from real puppetries are limited in the capturing condition and are usually low in quality.

6.3 Love Story with Puppetry

To demonstrate the usage of both the creator and manipulator module, we make a short video that tell a love story.

Table 3: ANOVA analysis of comparing animation sequence generated using weights from smoothness optimization and randomly selection

<table>
<thead>
<tr>
<th>F-statistics</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>327.53</td>
<td>6.60827e-50</td>
</tr>
</tbody>
</table>
between a man and a female puppet. The animations of the puppets in this video are generated using our manipulator module. The man in this story finally becomes a puppet, which is generated by our creator module.

Story Synopsis: A boy named Ryan was a PhD student. He was sad because he was still single. One day, a miracle happened, a very beautiful puppet girl appeared and said to him: “Would you come to my world?” So he was changed into a puppet and entered the girl’s world. They got married and lived happily ever since. Some example frames are displayed in Figure 11. The full video can be watched online and also included in the supplementary materials.

We designed 8 questions and asked all the subjects to watch and evaluate the video on a 5-point Likert scale: very good, good, normal, bad and very bad from score 5 to score 1. Besides watching the video, the subjects also tried the manipulator module by typing the script themselves and experiencing the play creation. Then we count the average score of each question, and the results are shown in Table 5. Q1 and Q2 are related to the user friendliness of our system. Q3 and Q4 are related to the performance of the creator module while Q5 and Q6 are related to the performance of the manipulator module. Q7 and Q8 are related to our initial motivation.

From the results we can conclude that most subjects agree that both the generated puppet (mean score 4.37) and its action pattern (mean score 4.31) can well preserve the traditional puppet style. And most subjects think our work can help preserve the traditional Chinese shadow puppet culture (mean score 4.27), and want to see similar product in the puppet museum (quite high score 4.60). Due to the limitation of the profile face, people cannot always recognize the puppet generated by our creator module (score 4.01). We leave refining this part as our future work. The subjects also provided some suggestions. One subject commented, “The idea is quite good, and I look forward to more wonderful puppetry videos”. Some of the subjects thought that more atomic actions should be added. One subject thought we should enhance the soundtracks of the puppet.

7. CONCLUSIONS AND FUTURE WORK
In this paper, we proposed the eHeritage of shadow puppetry, including a creator module and a manipulator module. The creator module can generate a puppet for one person based on his/her frontal view face image and profile face image. The manipulator module can automatically generate the motion sequences based on the script provided by the user. We conduct extensive experiments on puppetry creator and manipulator module, and the results show the effectiveness of the proposed system. Currently our system mainly focuses on the visual effect of the puppetry. In future, we would like to take more consideration to the audio parts, such as automatic vocal style transfer from human singing to puppetry style singing.

8. ACKNOWLEDGMENT
This research is partially supported by the Singapore National Research Foundation under its International Research
Puppet.walk(distance=10,duration=15)

# Script

……

Puppet.kneel(duration=20)

Puppet.point(duration=18)

……

Puppet.sit(duration=20)

Puppet.point(duration=15)

Puppet.talk(duration=10)

……


9. REFERENCES


