Paper and Pencil Interactive Storytelling Based on Generic Hand-drawn Sketches

Thiago Cler Franco Edirlei Soares de Lima *

Rio de Janeiro State University (UERJ), Department of Computational Modeling, Brazil

Abstract

Storytelling has been playing an important role in the culture of many societies. The advent of new technologies and the emergence of the first interactive narratives transformed the traditional stories into a new form of entertainment. In this context, Lima et al [9] proposed recently a new form of Interactive Storytelling called Paper and Pencil Interactive Storytelling, where stories are presented in augmented reality over conventional sheets of paper and users can interact with the virtual characters by sketching objects on the paper. In the present work, we aim to expand the Paper and Pencil Interactive Storytelling system by improving its repertory of identifiable drawings. We propose a new type of user interaction based on scenario changes, which involves two new classes of recognizable drawings: (1) environmental objects, which are used to directly modify the environment of the virtual world; and (2) climate objects, which are used to cause weather changes in the environment. The main objective of this paper is to present a new method for the recognition of generic hand-drawn sketches and their subsequent conversion into 3D objects in the augmented reality environment of the Paper and Pencil Interactive Storytelling system.

Keywords: interactive storytelling, augmented reality, sketch recognition.

1 Introduction

Since prehistoric times, storytelling has been playing an important role in the culture of many societies as a way of teaching, inspiring and entertaining. The emergence of new technologies and digital media favors the narrative process, providing new forms of storytelling. Recently, Lima et al. [9] proposed a new form of Interactive Storytelling called Paper and Pencil Interactive Storytelling, where the narrative is presented in an augmented reality over conventional sheets of paper, which creates the illusion that the paper is a virtual world populated by virtual characters. Users can interact with the characters by sketching objects on the paper (Figure 1). For example, in a narrative starring a hero who must face a dragon to save a princess, the user could draw a sword close to the hero's position in the paper. The sword will be recognized, transferred to the virtual world and taken by the hero, which will increase the hero's strength and his chance of success.

Although the system proposed by Lima et al. [9] presents good results, it has two main limitations: (1) it can only identify a limited set of predefined drawings that are represented by specific 3D models, which reduces the user creativity; and (2) users can only interact with the story by giving objects to the virtual characters, which reduces the user freedom. In this paper, we aim to expand and enhance the Paper and Pencil Interactive Storytelling system by improving the repertory of identifiable drawings, as well as the 3D objects used to represent them in the augmented reality environment. More specifically, we propose a new type of user interaction based on scenario changes, which involves two new classes of recognizable drawings: (1) environmental objects, which are used to directly modify the environment of the virtual world (e.g. rivers, walls, pits); and (2) climate objects, which are used to cause weather changes in the environment (e.g. rain, snow, fog).

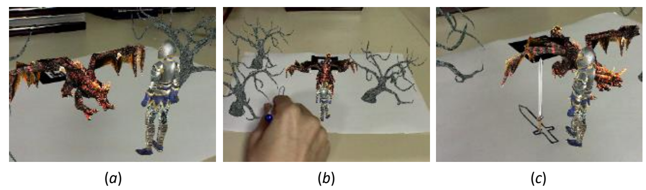


Figure 1: User interaction process: (a) virtual environment; (b) user sketching an object; (c) object (a sword) recognized and transferred to the virtual world.

Different from the previous objects used by the Paper and Pencil Interactive Storytelling system, the environmental objects created in the virtual world must reflect the characteristics of their respective drawings, such as shape, size and position. For instance, if the user draws a river with a specific path, the virtual representation of the river must follow the same path and have a similar shape. The inclusion of such objects requires from the system a more robust sketch recognition algorithm capable of recognizing generic drawings. In addition, the system must be capable of extracting all the important features of the sketches in order to create a 3D representation of them. Therefore, the main objective of this paper is to present a new method for the recognition of generic hand-drawn sketches and their subsequent conversion into 3D objects.

The rest of this paper is organized as follows. Section 2 reviews related work. Section 3 presents an overview of the system architecture. Section 4 describes the proposed method to recognize generic hand-drawn sketches. Section 5 presents the method to convert sketches into 3D objects. Section 6 contains some concluding remarks.

2 Related Work

There are several works contemplating the use of hand-drawn sketches to generate 3D models. One of the approaches is presented by Pu and Ramani [12], where the features extracted from sketches are compared to the features of a variety of preexisting 3D models in order to identify the model that matches more features. Another approach is presented by Rossa et al. [14], where algorithms to detect corners are used to generate walls for a 3D labyrinth. However, none of the aforementioned works are applied to augmented reality. The method presented by Bergig et al. [1] is the one that comes close to our interest. In their system, 3D models representing mechanical objects are created in an augmented reality environment according to hand-drawn sketches. Their method was also extended to augmented reality games as a content-authoring tool for drawing scenarios [5].

The use of hand-drawn sketches was also explored previously in the context of Interactive Storytelling systems. Vogelsang and Signer [16] and Kuka et al. [8] propose to use hand-drawn

*e-mail: edirlei.lima@uerj.br

sketches to interact with digital narratives creating virtual objects. However, their systems depend on special pen and paper to record and send the user's strokes to a computer.

Most of the previous work on sketch recognition presents methods capable of recognizing and representing only predefined objects. Here, we propose a method capable of: (1) recognizing different groups of generic hand-drawn sketches; and (2) creating unique 3D objects in the augmented reality environment based on characteristics extracted from the sketch.

3 System Overview

The Paper and Pencil Interactive Storytelling system is composed of a computer equipped with a webcam, an ordinary sheet of paper with a fiducial marker printed on it, and a common pencil (Figure 2). The system uses the camera to capture the image of the real world and apply several image processing and machine learning techniques to recognize the sketches drawn by the user.

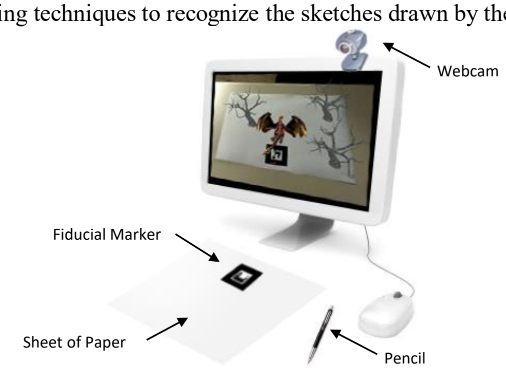


Figure 2: The Paper and Pencil Interactive Storytelling system.

As illustrated in Figure 3, the architecture of the proposed system is composed of four main modules. The Autonomous Characters Story Planner handles the actions of several virtual characters, each one introduced with predefined goals, whose behavior may be redirected via user interactions. The sketches drawn by the user on the sheet of paper are recognized by the Generic Sketch Recognition module, which uses two Artificial Neural Networks [6] trained to recognize groups of generic objects. When a sketch is recognized, a 3D model with the characteristics of the drawing is generated by the Generic Object Generation module and sent to the AR Dramatization module. The AR Dramatization controls and renders the virtual world superimposed over the real world objects, creating a mixed reality environment. If a fiducial marker is found on the image, the system renders the virtual world objects and characters according to the virtual location identified by the marker. The AR Dramatization module uses the ARToolKit Library[1].

The system is based on a parallel architecture that separates in different threads the augmented reality dramatization process from the sketch recognition process, which is important to guarantee that there will be no noticeable delays in the rendering process – which is currently limited to 30 frames per second, due to the camera capture speed. Since the recognition of user sketches is the most expensive process in the system, it must be executed in a separate thread, so that the system is able to efficiently render the output images in real-time.

While the main modules of the system are implemented in C++, the Neural Network Classifiers are implemented in Java using the Weka Library[2]. The communication between the Generic Sketch Recognition module and Neural Network Classifiers is done through a local TCP/IP network connection.

[1] https://github.com/artoolkit
[2] http://www.cs.waikato.ac.nz/ml/weka/

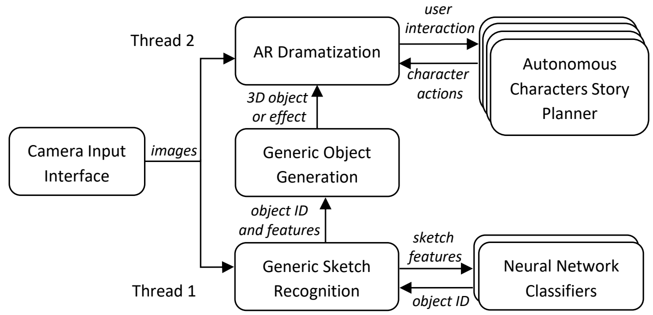


Figure 3: System architecture.

4 Generic Sketches Recognition

Our approach to recognize generic hand-drawn sketches consists in transforming the problem of classifying the drawing as a specific object into the problem of identifying to which generic group of objects the drawing belongs. These generic groups are characterized by generic features that describe all the objects that belong to the same group and are somehow related. In this way, we can infer the meaning of the sketch according to its group and then create a 3D object in the augmented reality environment according to characteristics extracted directly from the sketch.

The generic groups of objects are divided into two different types: environmental objects and climate objects. While environmental objects are used to directly modify the physical environment of the virtual world, the climate objects are used to cause weather changes in the environment. Table 1 describes the generic groups of objects used in our system.

The process of recognizing the different groups of sketches is divided into three phases: (1) a pre-processing phase, in which the contours of the sketches are identified and extracted; (2) a features extraction phase, where all the important features of the drawings are extracted from the contours; and (3) a recognition phase, in which the drawings are classified into their respective groups.

4.1 Pre-processing Phase

Assuming that drawings are always produced on a blank paper and the strokes contrast with the paper background, the sketches can be clearly distinguished by their contours. The process to extract the contours is implemented using the OpenCV Library [11]. As illustrated in Figure 4, the process is divided into three steps: (1) edges detection; (2) edges enhancement; and (3) contours extraction. In the first step, the image captured by the camera is pre-processed and the Canny Edge detection algorithm [2] is applied in order to identify the edges of the drawings (Figure 4b). Then, the edges are enhanced with the application of morphological filters (erosion and dilation), which reduces noise and segmented edges (Figure 4c). In the final step, the contours of the sketches are extracted from the edges using the algorithm proposed by Suzuki and Abe [15] (Figure 4d).

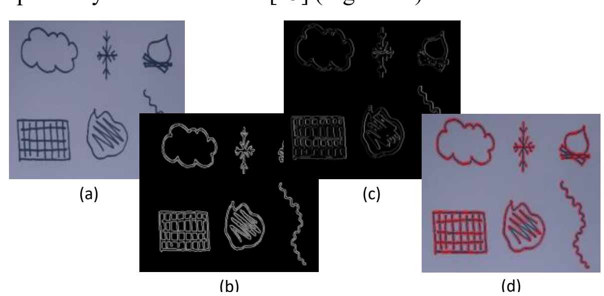
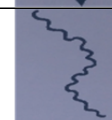


Figure 4: Contour extraction process: (a) original image; (b) detected edges; (c) enhanced edges; (d) original image with the extracted contours in red.

Table 1: Generic groups of objects.

Group	Type	Effect	Characteristics	Example
Wall	Environmental	A 3D wall will be placed over the sketch, preventing characters from passing by the location.	The rectangular form of the drawing will be used as base for the 3D wall, which will have the same shape, size, rotation, and position.	
Pit	Environmental	A 3D pit will be placed over the sketch. Characters will be able to pass by the location only if they could jump over the pit.	The shape of the 3D pit will be stablished according to the sketch contour, size, rotation, and position.	
Fire	Environmental	A 3D bonfire will be added to the environment over the sketch. Characters will be able to get warm with the fire.	The size of the sketch will affect the intensity and the size of the fire effect in the 3D environment.	
River	Environmental	A 3D course of a river will be placed over the contours of the sketch, forcing characters to find a way to cross the river.	The course of the 3D river will be stablished according to the sketch contour, position and rotation.	
Cloud	Climate	Drawing clouds on the paper causes rain.	The amount, size, form and position of the clouds' sketches affect the intensity and direction of the rain.	
Ice	Climate	Drawing snowflakes on the paper triggers a snowstorm.	The amount, form, size and position of the snowflakes' sketches affect the intensity and direction of the snowstorm.	

4.2 Feature Extraction Phase

After identifying the contours of the drawings, the next phase of the sketch recognition process involves the extraction of features from the contours. These features express common characteristics between the sketches of each group of generic objects, which is useful to identify the group to which a specific sketch belongs. The features used in our system are presented in Table 2.

Table 2: Features extracted from the sketches contours.

Feature	Description
Area/perimeter proportion	The area of the contour divided by the contour perimeter.
Vertices	The number of vertices of a polygon that approximates the contour.
Defects	The number of convexity defects present in the contour.
Defects proportion	Number of convexity defect points divided by the total number of points in the contour.
Smallest circle proportion	The area of the contour divided by the area of the smallest circle that circumscribes the contour.
Inner contours	The number of contours situated inside of the main contour.
Inner area proportion	The area of the main contour divided by the total area of the inner contours.
Inner perimeter proportion	The perimeter of the main contour divided by the total perimeter of the inner contours.
Inner contours proportion	The number of vertices of a polygon that approximates the main contour divided by the total number of vertices of a polygon that approximates each of the inner contours.
Moments	The Hu's moment invariants [7] on the contour image.

4.3 Recognition Phase

The correct recognition of a sketch can be seen as a classification problem, wherein the system knows a set of generic drawings (Table 1) and must recognize a new drawing based on its similarity to some member of the known set. In order to perform this classification, we tested different Machine Learning algorithms, such as Decision Trees, Artificial Neural Networks, Support Vector Machines, and Logistic Regression. To test the algorithms, we created a training dataset with 40 examples of sketches for each of the generic groups of objects (total of 240 examples) and performed a 10-fold cross validation test using the most common classification algorithms implemented in the Weka Library. Among the tested algorithms, the Artificial Neural Network classifier was the one that obtained the best recognition rate results (accuracy of 92.18%). After training the Neural Network with the training dataset, we also validated the classifier with a new testing dataset composed of 255 new examples of sketches. For this testing dataset, the classified obtained a recognition rate of 87.05%.

By analyzing the initial results of the Neural Network classifier, we noticed that most of the incorrect classifications occur between the groups of "Cloud" and "River". In order to deal with this problem, we created a secondary Artificial Neural Network classifier trained specifically to identify whether a given sketch belongs to the group of "Cloud" or "River". This secondary classifier obtained a recognition rate of 97.46% in a 10-fold cross validation test. In this way, when the first Neural Network classifies a sketch as "Cloud" or "River", the secondary Neural Network is activated in order to define the final class of the sketch. By using the two Artificial Neural Networks combined, the accuracy of our generic sketch recognition system for the testing dataset (with 255 new examples of sketches) is 89.80%.

5 GENERATING 3D OBJECTS FROM SKETCHES

In order to handle the new classes of generic objects, we developed a new module for the Paper and Pencil Interactive

Storytelling system that is responsible for creating parameterized 3D models of the generic groups of objects. When a new sketch that belongs to one of the defined groups of objects is identified, the system automatically calculates the position, form, size and rotation of the object according to the contours of the sketch. The process to generate parameterized 3D models for the user sketches is implemented directly in OpenGL.

For each group of generic objects, a different object generation approach is adopted. For the generation of walls, we use the polygon that better approximates the points of the drawing contour to form a polygonal base (lower base) for the 3D object (Figure 5a). Then, we create a second base (upper base) as a translated copy of the lower base (Figure 5b) and a set of faces joining the corresponding sides of the two bases (Figure 5c). Figure 6 shows an example of wall created in the augmented reality environment according to the hand-drawn sketch.

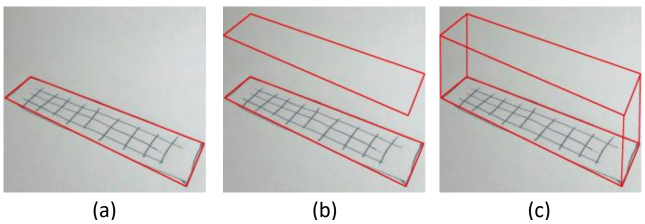


Figure 5: Wall generation process.

For the generation of pits and rivers, we use the set of points of the sketch contour to generate a polygon matching the points in a 3D space. In order to generate this polygon (often a concave polygon), we use an ear clipping algorithm [4][3] so that the contour of the sketch can be represented in terms of OpenGL triangle primitives. Figure 6 shows examples of a pit and a river generated in the augmented reality environment.

The generation of fire, rain and snow involves the simulation of particle systems [13]. Once a drawing that represents one of these elements is identified, a particle system is instantiated in the augmented reality environment and its parameters are adjusted according to some characteristics extracted from the sketch contour. For the rain and snowstorm effects, the emission rate of the particle system is adjusted according to the amount and size of clouds and snowflakes drawn by the user, and the direction of the particles is modified according to the rotation of the drawings. While rain and snowstorms are simulated over the entire environment, the particle system for the fire effect is instantiated on the exact position of the user sketch. For this latter case, the size of drawing is used to adjust the emission area and the size of the particles of the system that simulates the fire effect.

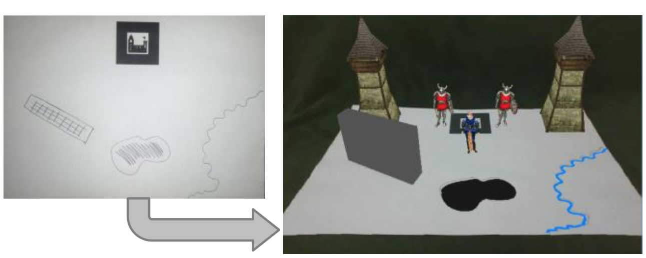


Figure 6: 3D objects generated in the augmented reality environment according to the hand-drawn sketches.

6 Concluding Remarks and Future Works

In this paper we present a new user interaction method for the Paper and Pencil Interactive Storytelling system using objects that directly modify the environment of the virtual world through physical and climate changes. The method is supported by a generic sketch recognition system and a generic object generation system, which are capable of identifying hand-drawn sketches and converting them into 3D objects in the augmented reality environment of the story.

Although the results produced by the proposed methods achieved the primary objectives of this paper, further research still need to be done in order to evaluate them in more complex scenarios. As further research, we intend to validate our methods through the development of a new narrative entirely based on the new interaction mechanisms. Furthermore, we also plan to conduct user studies in order to evaluate the proposed interaction methods from the user's perspective.

References

[1] O. Bergig, N. Hagbi, J. El-Sana, and M. Billinghurs. In-Place 3D Sketching for Authoring and Augmenting Mechanical System. In *Proceedings of the 8th International Symposium on Mixed and Augmented Reality*, pages 87-94. IEEE, 2009.

[2] J. Canny. A Computational Approach To Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 8 (6), pages 679-714. IEEE, 1986

[3] J. A. de Loera, J. Rambau, and F. Santos. Triangulations: Structures for Algorithms and Applications. Springer, 2010.

[4] H. ElGindy, H. Everett, and G. T. Toussaint. Slicing an ear using prune-and-search. *Pattern Recognition Letters*, volume 14 (9), pages 719-722. Elsevier, 1993.

[5] N. Hagbi, R. Grasset, O. Bergig, M. Billinghurst, and J. El-Sana, J. In-place sketching for augmented reality games. *Computers in Entertainment*, volume 12 (3), article 3. ACM, 2014.

[6] S. Haykin. Neural Networks and Learning Machines. Prentice Hall, 2008.

[7] M. K. Hu. Visual Pattern Recognition by Moment Invariants. *IRE Transactions on Information Theory*, volume 8, pages 179-187, 1962.

[8] D. Kuka, O. Elias, R. Martins, C. Lindinger, A. Pramböck, A. Jalsovec, P. Maresch, H. Hörtner, and P. Brandl. Deep space: high resolution VR platform for multi-user interactive narratives. In *Proceedings of the 2nd Joint International Conference on Interactive Digital Storytelling*, pages 185-196. Springer, 2009

[9] E. S. Lima, B. Feijo, S. D. J. Barbosa, A. L. Furtado, A. E. M., Ciarlini, and C. T. Pozzer. Draw Your Own Story: Paper and Pencil Interactive Storytelling. *Entertainment Computing*, volume 5 (1), p. 33-41. Elsevier, 2014.

[10] J. Meehan. TALE-SPIN, an interactive program that writes stories. In *Proceedings of the 5th International Joint Conference on Artificial Intelligence*, pages 91-98, 1977.

[11] OpenCV. OpenCV (Open Source Computer Vision Library. Available at: http://opencv.org/.

[12] J. Pu, and K Ramani. A 3D Model Retrieval Method Using 2D Freehand Sketches. In *Proceedings of the 5th international conference on Computational Science*, pp. 343-346. Springer, 2005.

[13] W. T. Reeves. Particle Systems - a Technique for Modeling a Class of Fuzzy Objects. *ACM Transactions on Graphics*, volume 2 (2), pages 91-108. ACM, 1983.

[14] P. Rossa, D. Camozzato, F. Marson, and R. Hocevar. 3D Model Generation from Freehand Drawings. In *Proceedings of XV Brazilian Symposium on Computer Games and Digital Entertainment*, pages 226-229, 2016.

[15] S. Suzuki, and K. Abe. Topological Structural Analysis of Digitized Binary Images by Border Following. *Computer Vision, Graphics, and Image Processing*, volume 30, pages 32-46. Elsevier, 1985.

[16] A. Vogelsang, and B. Signer. The lost cosmonaut: an interactive narrative environment on the basis of digitally enhanced paper. In *Proceedings of 3rd International Conference on Virtual Storytelling*, pages 270-279. Springer, 2005.