Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 6, Issue, 05, pp.6576-6579, May, 2014

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

REVIEW ARTICLE

SURVEY ON BUILDING INEXPENSIVE FOG SCREEN, AN IMMATERIAL DISPLAY SYSTEM WITH 3D NATURAL USER INTERFACE

Dadasaheb Mhaske, *Mayur Kshirsagar, Namrata Singh, Samreen Khan and Prof. Rahul Kulkarni

Department of Computer Engineering, Sinhgad Institute of Technology, Lonavala, Maharashtra, India

ARTICLE INFO

Article History:
Received 07th February, 2014
Received in revised form
26th March, 2014
Accepted 19th April, 2014
Published online 20th May, 2014

Key words:
Gestures,
Gesture Recognition,
Kinect.

ABSTRACT

Fog screen is patented technology that creates a seamless, thin and flat layer of screen floating in the air. Fog screen is a miracle in itself – here fog is generated using normal fog machine, so it won't make you wet. Fog screen captures the light from the regular projector, so anything you can see on computer screen you can also project in the air. "In this paper we are presenting a new and fascinating way of human computer interaction, because we are developing a fog screen for computer screen projection with 3D gesture implementation. Developing solution for 3D gesture implementation and building an inexpensive low budget fog screen is important part of this paper."

INTRODUCTION

The purpose of this paper is to provide an even more interesting point of view of handling and working with computers. This is advancement in the field of computers. It is always interesting to work with a floating window, rather than manually working on actually providing every command using hardware. With the help of 3D Sensors like Microsoft Kinect and a fog screen we can develop 3D images. The FogScreen is an optimal, method for creating a physically penetrable particle display that allows projection of interactive content to appear floating in free space. The basic principle is the use of a large non-turbulent airflow to protect a flow of particles, such as dry fog, inside it from turbulence. The resulting thin, stable sheet of fog enables projections on a screen that is dry and feels like moderately cool air. The light from a standard LCD projector is scattered through this sheet of fog screen, creating a rear-projection image.

Existing System

A projection screen is an installation consisting of a surface and a support structure used for displaying a projected image for the view of an audience. Projection screens may be permanently installed as in a movie theatre, painted on the wall, semi-permanent or mobile, as in a conference room or other non-dedicated viewing space such as an outdoor movie screening (open air cinema). Uniformly white or grey screens are used almost exclusively as to avoid any discoloration to the image, while the most desired brightness of the screen depends on a number of variables, such as the ambient light level and the luminous power of the image source. Flat or curved screens may be used depending on the optics used to project the image and the desired geometrical accuracy of the image production, flat screens being the more common of the two. Screens can be further designed for front or back projection, the more common front projection systems having the image source situated on the same side of the screen as the audience. Different markets exist for screens targeted for use with digital projectors, movie projectors, overhead projectors and slide projectors, although the basic idea for each of them is very much the same: front projection screens work on diffusely reflecting the light projected on to them, whereas back projection screens work by diffusely transmitting the light through them. Alex Olwal1, Stephen DiVerdi, Nicola Candussi Ismo Rakkolainen Tobias Höllerer, presented An Immaterial, Dual-sided Display System with 3D Interaction (Figure 2). Present interactive, dual-sided, wall-sized system that allows multiple unencumbered users to naturally manipulate objects floating in midair, as shown in Figure 1. Two individual, but coordinated, images are projected onto opposite sides of a thin film of dry fog and an integrated 3D tracking system allows users on both sides to interact with the content on display, while the non-intrusive and immaterial display makes it possible to freely pass physical objects between users or move through the fog Screen.

***Corresponding author: Mayur Kshirsagar**
Department of Computer Engineering, Sinhgad Institute of Technology, Lonavala, Maharashtra, India.

We are developing a Fog screen which is in small size as compared to the system given the paper. We are also implementing 3D gestures captured through Asus Xtion PRO to interact with the system, so that the system will become more interactive and attractive. We are going to use OpenNI framework, an open-source framework to communicate with Asus Xtion PRO. The figure 3 shows the proposed system.

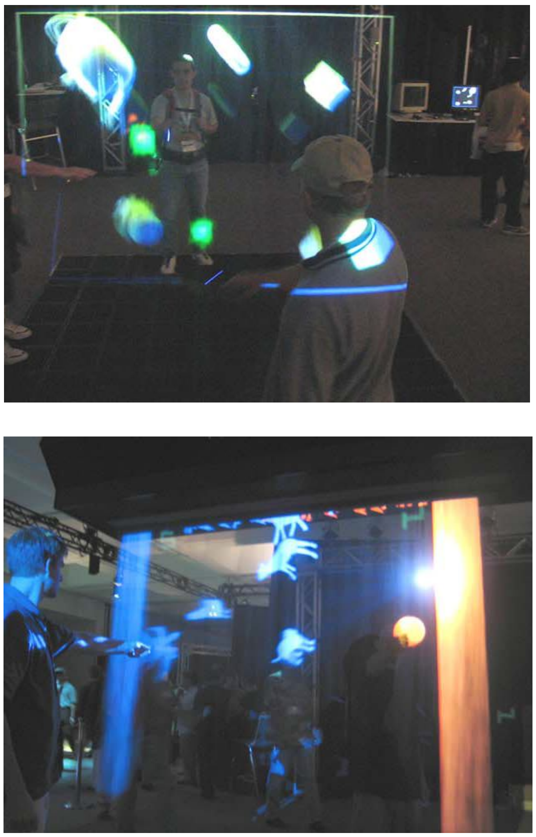


Figure 2. Existing wall sized FogScreen

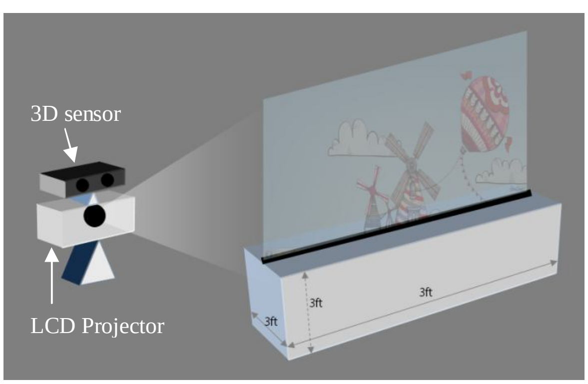


Figure 3. Proposed System

Gesture Recognition

Gesture is the way to express feelings or information through human body. Gesture recognition is the process of detecting patterns of hand movement to act upon any event. Gestures are expressive, meaningful body motions involving physical movements of the fingers, hands, arms, head, face, or body with the intent of Conveying meaningful information or Interacting with the environment.

Kinect

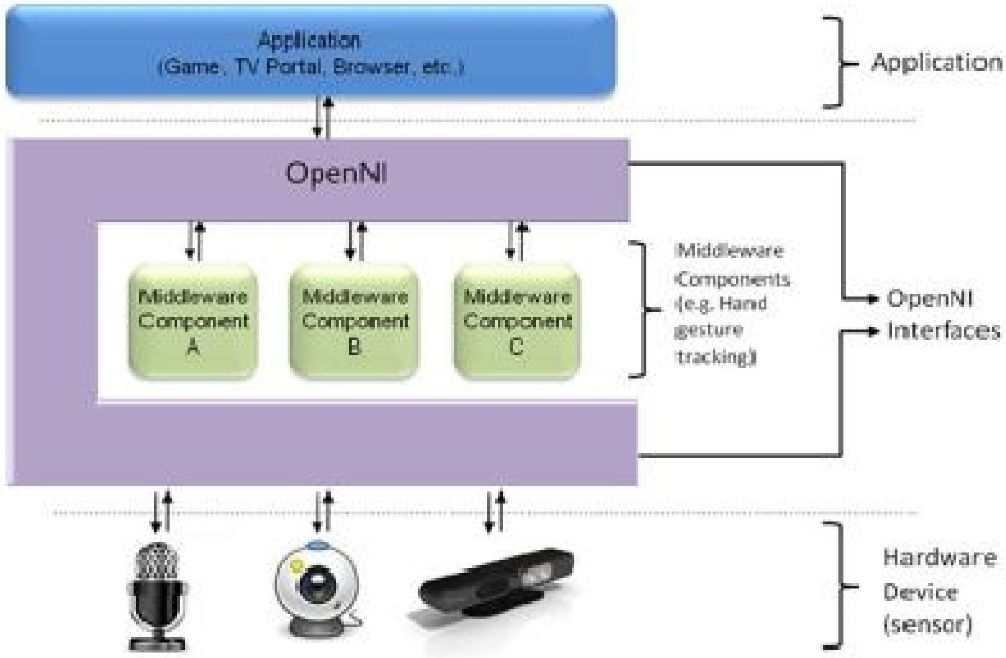
Kinect is a motion sensing input devices by Microsoft for Xbox 360 and Xbox One video game consoles and Windows PCs. Kinect allows users to control and interact with their computer, through a natural user interface using gestures and spoken commands. The Kinect sensor has a practical ranging limit of 1.2–3.5 m (3.9–11.5 ft) distance. Kinect sensor is a horizontal bar connected to a small base with a motorized pivot and is designed to be positioned lengthwise top or bottom of the video display. The device features an "RGB camera, depth sensor and multi-array microphone running proprietary software", which provide full-body 3D motion capture, facial recognition and voice recognition capabilities. The default RGB video stream uses 8-bit VGA resolution (640 × 480 pixels).The depth sensor camera consists of an infrared laser projector combined with a monochrome CMOS sensor, which captures video data in 3D .Asus Xtion PRO is alternative to Kinect for gesture recognition

Asus Xtion Pro

This is the motion sensor. It is almost exactly the same as the Microsoft Kinect sensor but is far smaller in size. It is also the only PCexclusive motion sensor and is therefore aimed at the creation of a wide variety of applications rather than just for use as a gaming peripheral. This particular sensor contains an RGB camera, a depth camera that works via the use of infrared and two microphones so that speech input may also be used in conjunction with libraries such as CMUSphinx. The depth camera is capable of producing a VGA image size of 640x480 at a rate of 30fps and can also produce a QVGA image size of 320x240 at a rate of 60fps. The RGB camera can produce an SXGA image size of 1280x1024. The gesture generators and tracking objects that are used in OpenNI and NITE can be used at a minimum of 0.8m and a maximum of 3.5m when using this camera.

Openni

OpenNI stands for Open Natural Interaction and it is a multilanguage, 2 cross platform framework for applications utilizing natural interaction. OpenNI’s main purpose is to form a standard API for communication with both motion sensing input devices and their middleware which means that it provides a set of API’s for use by the sensor device and a set of API’s for use by the middleware components. This breaks dependency between the sensor and the middleware to give a write once, run anywhere effect so that different middleware modules can be used. OpenNI also allows developers to track reallife scenes through data types that are calculated from the input of a sensor. These are all provided in the xn namespace and provide types such as XnUserID which is used to hold the unique identifiers of the users in the scene, XnCallbackHandle which is used to hold all the callback functions for the



generator nodes and XnStatus which is used many times throughout most applications to check that there have been no errors. We can visualise the OpenNI architecture in a layered view. This starts with the application utilising Open NI at the top in its own layer. Following this the middleware components such as the world in three dimensions using a Prime Sensor's depth images. It translates this perception into meaningful data in the same way that people do. An analogy of this is to think of the PrimeSensor (ASUS Xtion Pro) as a human's eyes and the NITE middleware component as a human's brain which is the perception engine for comprehending interaction within the surroundings. NITE is a set of computer vision algorithms and a framework API for implementing a natural interaction based user interface. The design of the NITE middleware is based on two paradigms, hand control and full body control. This project will use part of the full body paradigm focusing only on the top half of the body. This is because a hand can be brought straight into use once the top half of the body is calibrated and tracked unlike the hand control paradigm which requires a focus gesture to start a session. This focus gesture makes the mouse harder to properly control at the beginning of the systems' use and thus lowers its efficiency as an alternate means of interaction. The full body control paradigm generates a simple point skeleton. The points created by this can then be assigned individual controls for an interface. This method means that applications are not exposed to infinite amounts of depth data which can overwhelm it. This leads to a computationally inexpensive extraction of user intentions because all that is extracted from an image is a set of several joint points which outline a user's skeleton.

Algorithms for Gesture Recognition

Model Based Approach

The 3D model approach can use volumetric or skeletal models, or even a combination of the two. Volumetric approaches have been heavily used in computer animation industry and for computer vision purposes. The models are generally created of complicated 3D surfaces, like NURBS or polygon meshes. The drawback of this method is that is very computational intensive and systems for live analysis are still to be developed. For the moment, a more interesting approach would be to map simple primitive objects to the person's most important body parts (for example cylinders for the arms and neck, sphere for the head) and analyse the way these interact with each other. Furthermore, some abstract structures like super-quadrics and generalised cylinders may be even more suitable for approximating the body parts. The very exciting about this approach is that the parameters for these objects are quite simple. In order to better model the relation between these, we make use of constraints and hierarchies between our objects.

Appearance Based Approach

These models don't use a spatial representation of the body anymore, because they derive the parameters directly from the images or videos using a template database. Some are based on the deformable 2D templates of the human parts of the body, particularly hands. Deformable templates are sets of points on the outline of an object, used as interpolation nodes for the object's outline approximation. One of the simplest interpolation function is linear, which performs an average shape from point sets, point variability parameters and external deformators. These template-based models are mostly used for hand-tracking, but could also be of use for simple gesture classification. A second approach in gesture detecting using appearance-based models uses image sequences as gesture templates. Parameters for this method are either the images themselves, or certain features derived from these. Most of the time, only one (monoscopic) or two (stereoscopic) views are used.

View-independent hand poses recognition

A view-independent hand pose recognition system the recognition of a limited set of pre-defined postures from single, low resolution depth images in real time on standard hardware in unconstrained environments. The system consists of three modules: hand segmentation and pose compensation, feature extraction and processing, and hand pose recognition. In this paper they used principal component analysis to estimate the hand orientation in space and Flusser moment invariants as image features for visual recognition. The implementation details, classification accuracy and performance measures of the recognition system are reported and discussed. The experimental results show that the system can recognize the pose of two hands at full frame rate with an average total latency lower than 5 ms.

Future Scope

Future research on this topic will consider environment friendly and more cheap alternative for fog generation ,Improved Image Quality, Stereoscopic rendering. Research on complex primitives and fog display quality. Future research will also consider new 3D technology algorithms, complex gestures.

Acknowledgement

We would like to thank our reviewers for their thoughtful comments and suggestions.

REFERENCES

Akira Kataoka, Yasuhiro Kasahara "Method for Fog Screen and Image-forming Method using same" US patent 270752

Alex Olwal, Stephen DiVerdi, Nicola Candussi, Ismo Rakkolainen, Tobias Höllerer "An Immaterial, Dual-sided Display System with 3D Interaction" IEEE Virtual Reality 2006 March 25 - 29, Alexandria, Virginia, USA

http://www.openni.org

Losan Oikonomidis , Nikolaos kyriazis , Antonis A. Argyros "Efficient model based 3D tracking of hand articulations using kinect" Institute of Computer Science, FORTH

Luigi Gallo, Alessio Pierluigi Placitelli, "View-independent Hand Posture Recognition from Single Depth Images using PCA and Flusser moments" 2012 Eighth International Conference on Signal Image Technology and Internet Based Systems.

****