Taylor Moon

Art 185GL Research Project

Intro

StoryCam has two main functions which are both contingent upon tracking in virtual reality and image processing. Firstly, this camera allows anyone to capture their environment and convert it to a virtual reality landscape that can be replayed in a 360 degree video format. Secondly, this same tracking system also allows the camera to capture the environment according to cinematic shots for book publication and to extract colors from the images to use elsewhere in the book for graphic design coherence. The 360º element of the camera lens is optimal for both functions due to the variety of angles that one would want to obtain for clearer narration in the book storyboard. This allows the distinction between book imagery and movies to become more blurred as these books begin to increasingly appropriate cinematic shots. Both virtual reality films and the books generated by *StoryCam* would center their narratives around subjects found through face-and-object detect tracking systems. <u>StoryCam utilizes the intersections of cinema, virtual reality, architecture, and robotics by looking towards technological precedents of object and face detection, optic tracking, and color detection; StoryCam democratizes storytelling, making the creation of film and books accessible to anyone through tracking in virtual reality.</u>
1 minute 18 sec

***StoryCam* - book generating**

StoryCam is a camera that presents the user with multiple computer-generated narratives from the surrounding environment. It essentially captures different types of photos based upon predesigned storyboard templates from which the user can select their favorites to create a book. These templates are inspired by cinematic storyboard layouts and will auto-determine the camera angle, content (action or sedentary, number of people, etc), zoom, cropping, and focus in order to create different moments in the narrative (Golda). The camera can either operate auto-pilot, in which it takes the entirety of the narrative series, or one could take each individual photo through the shooting mode dial. This dial is semi-automated because it is preset with zoom shot, action shot, find face, & other modes that the camera would otherwise find and take for you. The advantage of switching the camera off of complete auto-piloting to the semi-automated dial is that the user can then take more of one particular shot by leaving it on that particular dial mode rather than having to circulate through taking the entire narrative sequence to get a different version of one particular shot. Rather than upload an image to a web browser or a vector graphics editor, such as Adobe Illustrator, for the color picker and eyedropper tools, *StoryCam* is already equipped with image processing, thus, eliminating third-party, corporate, software-giants. Using the arrow keys, one could select anywhere within an image and 9 x 9 pixels will be magnified within a small box ("Get Colors.."). This type of technology is already available through HTML5 color codes.
1 minute 39 sec

***StoryCam* - VR films**

StoryCam will change the way in which the viewer navigates through virtual space and engages with mediated interactivity. *StoryCam* is able to translate the real world coordinate system into a local object coordinate system (LOCS) through methods of data collection (Goldfeather 28). This process requires that the camera locate real-world objects through optic tracking and extract their coordinates for conversion in order that virtual reality be made possible (Goldfeather 28). With this technology available, *Pearl*, which is a 360-degree animation, one of Google's Spotlight Stories, and an Oscar nomination, has become a landmark innovation in storytelling through virtual reality (D'Zurilla; "Oscar Nominees"). It simultaneously represents the types of work that can be done through *StoryCam. Pearl* uses ___ with which *StoryCam* is equipped, such as the amalgamation of stereoscopic and monoscopic rendering and the full 6-degrees of freedom (Curtis). The stereoscopic footage proffers a more three-dimensional appearance for objects in the foreground by shooting from two different angles, whereas, monoscopic portrayals are used for distant content in VR films (Palandri). This mimics human visual experience because stereoscopic shooting mimics two eyes' slightly different views that inform depth perception (Rogers-Ramachandran) and the monoscopic effect eyes see when distant mountains appear to look like fake, cardboard facades. How does the camera know which areas should be assigned to stereoscopic rather than monoscopic and vice versa? Limits define that a pixel after a certain depth will not be applied with a stereoscopic projection but will engender a monoscopic at that limit (Palandri). Moreover, this programming pixels already rendered ensures that a pixel already rendered stereoscopically will not be rendered monoscopic (Palandri). The intent with *StoryCam*, like that of *Pearl*, would be for smooth operation multiple platforms (Google Spotlight Stories). Meaning, a viewer is able to easily interface with the VR video after it is created and uploaded.

Face and Object Detection in Both Functions

Face detection is integral for both of the camera's narrative functions (i.e. virtual reality film and image generation for books). Face detection technology witnesses a variety of uses, seen in human computer interaction (HCI), surveillance systems, and biometric systems (Candido 179). In a related webcam project, face recognition is processed through a geometric face model that anticipates the eyes' expected position in order that user's with Parkinson's and other disabilities are able to optically control a computer mouse (Candido 180). This same type of logic and technology is used in *StoryCam*, which estimates a subject's ensuing motion and captures it for optimal action-shots. For the camera to even initially recognize a face, it is programmed with color algorithms that define what parameters of color classify as skin (Vezhnevets). A metric calculates distance of the pixel color to skin tone and the relative size of the cluster so that a stray color pixel in the midst of a color field does not trigger a false face-alert (Vezhnevets). These bounds would appear as an if-statement, delineated as, "if R>95 and G >40 and B > 20 and max{R,G,B} - min {R,G,B} > 15 and |R-G| > 15 and R > G and R > B" then a pixel is labeled as "skin color" (Vezhnevets).