Musing: Interactive Didactics for Art Museums and Galleries via Image Processing and Augmented Reality

Providing Contextual Information for Artworks via Consumer-Level Mobile Devices

Gentry Atkinson, Kevin Whiteside, Dan Tamir
Department of Computer Science
Texas State University
San Marcos, TX, USA
{gma23, kjw52, dt19}@txstatet.edu

Grayson Lawrence, Mary Mikel Stump
School of Art and Design
Texas State University
San Marcos, TX, USA
{gl16, mr14}@txstate.edu

***Abstract*—Textual didactics, used in museums and galleries, provide access to historical, socio-political, technical, and biographic information about exhibited artworks and artists. These types of didactics are considered to be cost-effective. However, they do not enable the use of audio, video, and Web interface that allow for multiple forms of usage for the museum visitors. We have developed a smartphone application, called *Musing*, for interaction of museum visitors with informational content and enhancement of their museum experience. *Musing* is an augmented reality (AR) application that enables the visitor to capture an artwork with a smartphone camera. Using image processing, the application recognizes the artwork and places graphical user interface objects in the form of Points Of Interest (POIs) onto the image of the artwork displayed on-screen. These POIs provide the visitor with additional didactic information in the form of text overlays, audio, video, and/or Web sites. The *Musing* application, described in this paper, is designed with several performance and efficiency goals, including high reliability and recognition rate, high usability, and significant flexibility. The application is designed to be adaptable to a variety of museums and galleries without requiring special hardware or software. Furthermore, an administrative interface enables museum staff to provide content for the didactic purposes without requiring software development skills.**

Keywords—interactive didactic; museum didactic; virtual meseum; image recognition; augmented reality

I. INTRODUCTION

Museums have historically been tasked with providing access to and educating visitors about artworks. Museum didactics attempt to clarify artworks' meanings by addressing concepts of art, history, geography, politics, and artistic medium/techniques, as well as the lives of artists. For many visitors, however, museum and gallery exhibitions may lack the proper context to allow access points for exhibited works and can leave the "uninitiated viewer" feeling intimidated, "particularly when it comes to interpretation" [1].

In many ways, mobile technologies, such as responsive Web sites and Augmented Reality (AR), present an ideal opportunity to make those personal connections with the visitor, as well as help the visitor make connections to the exhibited objects and/or works of art. As such, the context for the artwork is broadened via interviews, videos, Web sites, source material, art historical influences, and other artworks with shared conceptual frameworks, all of which can be integrated into a mobile application for the museum. Such a personalization of experience through narrative is a highly effective way to expand the context for the work and deepen viewers' connections as they process and integrate the information into their existing world-view [2].

Nevertheless, under the current paradigm, in order to add audio and video to exhibits, museums must rely on proprietary hardware and software. The hardware must be provided by the institution at significant cost both in capital investment and in maintenance. The software used on these devices is often proprietary for the exhibition, reliant on external hardware installed in the gallery, and must be reprogrammed for new exhibitions. While large museums may have the resources to purchase and maintain these systems, smaller community-based museums often do not.

Pedagogical shifts away from passive to active participation are occurring in higher education, as well as in museological practices, and reflect the changing needs of the visitor [3]. An enriched learning environment requires incorporating diverse learning styles including visual/print, visual/picture, auditory, kinesthetic, and verbal/kinesthetic modalities [3].

A. Problem Statement

In order for museums and galleries to fully meet the needs of their visitors, they must incorporate didactic information that embraces diverse learning styles and present multiple types of didactic information.

In order to reach the highest number of museums and their visitors, an interactive didactic system should be designed, which does not rely on proprietary hardware, the installation of external devices in the gallery, or the need to reprogram the system when exhibits are modified or added.

In order to create a system that does not require proprietary hardware, the system should be developed on mobile hardware that many of the museum visitors already possess. This hardware would include classes of smartphones and tablets running on iOS or Android operating systems.

In order to minimize the technical burden on museums, the system should not rely on extra hardware such as Bluetooth or Near Field Communication (NFC) devices.

Finally, image processing and image recognition (IR) algorithms should be used in order to provide the opportunity for the viewers to deepen their connections to artworks and remove the need for external tokens such as Quick Response codes (QR) or number codes to be entered by users.

B. Hypothesis

By using a combination of off-the-shelf image recognition algorithms and unmodified consumer-level hardware, the research team will be able to create a system that is fast and accurate enough to be usable in a museum, without the need for proprietary hardware or external tokens. In addition, retrieving exhibition data via a database will allow for a client program that is sufficiently flexible and does not require reprogramming when exhibitions are added or modified.

The proposed interactive didactic system will be designed with a client-server architecture. A database will provide the client application with access to didactic information without the need to permanently store that information on the device. The client application will be programmed for current popular hardware such as a smartphone or a tablet, either owned by the museum visitor or provided in the form of a loaned device.

In order to test the relative success of the application and its acceptance by museum visitors, *Musing* will be deployed in an exhibition at The University Galleries at Texas State University, a three thousand square foot, university-based, contemporary art exhibition venue. Benchmark testing of the application will be conducted in order to determine IR accuracy rate and speed. An exit questionnaire will be given to visitors in order to determine their acceptance of the system and perceptions of system performance and usability.

C. Proposed Solution

Musing, a mobile, image recognition and AR application, runs on consumer-based mobile hardware, requires no external tokens or hardware, and does not require reprogramming between exhibits. The application has passed the Apple approval process and is available at [4].

The main contributions of this research is the design, development, and deployment of an end-to-end reliable, usable, and effective AR system that provides a museum visitor with virtual information and provides museum staff with adaptable, cost effective, and easy to maintain virtual museum utility. To date and to the best of our knowledge, this is the only fully functional system that integrates hardware agnostic and software agnostic virtual museum content delivery, and administrative support.

This paper is organized in the following way: Section II provides background in the form of relevant past research performed by this team, with Section III containing a Literature review. The application deployment of *Musing* is outlined in Section IV, followed by deployment results showcased in Section V. Section VI explains the evaluation of results from both benchmark testing and exit questionnaires given to the museum visitors. Lastly, Section VII outlines the conclusions and future research for *Musing*.

II. BACKGROUND

A. Previous Research

In 2012, the research team developed a series of responsive Web pages triggered by QR codes used in an exhibition at The University Galleries at the Texas State University [5].

In this pilot program, a QR code was included in the tombstone wall label placed next to artworks in the gallery. These codes, when scanned with reader software on the user’s smartphone, presented the visitor with a custom-built Web page for each artwork. These pages provided supplemental didactic information via news articles that pertained to the artwork’s subject matter, full artist biographies, video interviews with the artist, photos of the artist’s workspace, and links to external Web sites.

During the pilot exhibition, the gallery Web site recorded 23 unique visitors per day with an average time on-page of 3 minutes and 37 seconds. The Web pages that were only accessible by the QR codes were responsible for 16 of the 23 unique daily visitors (69%) and the majority of the time on-page (3 minutes and 33 seconds). For comparison, exhibits installed after the pilot test did not include QR codes. The subsequent exhibit showed a decline in both the number of online visitors (-26%) and the amount of time visitors spent on the gallery Web site (-42.5%). This data indicates that when QR codes are included with the artworks in the gallery, there is an increase in both online traffic and online interaction with the visitor.

The experiment with QR codes in the gallery indicated that visitors would use interactive technologies in the gallery and that they would spend the time necessary to consume the extra content. However, a major drawback of the QR codes was the inability for the museologist to contextually place information within artworks’ representation. This ability would allow the administrator to place content exactly where it would be most pertinent to the visitor’s view of the artwork. For example, a POI could be visibly placed relative to a specific element of an artwork to provide information about that element’s significance. Lastly, QR code reader software is not created specifically for the needs of museums and galleries, as they are designed to work for a wide variety of applications, from advertising to stock keeping.

Following the successful response to the QR code project, it was decided that the next step in the research should be to create an AR system that would allow for information placed within an artwork, designed specifically for the needs of museums and galleries.

III. LITERATURE REVIEW

A literature review showed a number of teams researching the possibility of using AR to augment the information provided by museum didactics. In most of the cases, however, these didactics rely on proprietary hardware, require reprogramming between exhibitions, or installation of external tokens (e.g., Bluetooth, RFID, and

QR) within the museum space. Some work has been done with respect to the challenges of image recognition, but little attention has been paid with regard to integrating hardware/software agnostic image based picture recognition with content delivery.

Bimber et al. have developed a mobile system, named, *PhoneGuide* allowing museum visitors to use mobile phones to detect artworks in a physical museum space [6]. Their method includes image recognition, using the phone's camera, as well as pervasive tracking techniques using a grid of Bluetooth emitters distributed in the space [6]. The reliance on external tokens (e.g., Bluetooth) to assist in the object recognition would require the museum to install new hardware and provide for updates in each gallery space.

Hatala et al. describe a prototype system, called Ec(h)o, developed to provide "spatialized soundscapes" for museum visitors [7]. That is, specialized audio is played for the listeners depending on their position within the museum. The supplied audio is meant to improve the overall experience of the exhibit rather than providing information specific to each artwork.

Jing et al. have developed a mobile augmented reality prototype system which uses image recognition running on specialized hardware to provide additional information on physical images displayed in museums for Personal Museum Tour Guide Applications [8]. The system uses the SIFT recognition algorithm that employs "coarse to fine" recognition to improve the speed of the process [9]. Nevertheless, some users complained of slow processing speed [8].

Blockner et al. developed a prototype system which allows users to create virtual museum tours on a mobile app. The mobile device uses NFC to transmit these tours to projectors positioned within the gallery which display the desired information [10].

Miyashitat et al. have developed an interactive device at the Dai Nippon Printing (DNP) Museum Lab at the Louvre Museum (Paris) for use with an exhibition on Islamic Art. This device used a neural network based system to map content of exhibits and was able to recognize three dimensional objects from a single viewpoint, but also relies on purpose specific hardware which is not available outside the Louvre and requires that Bluetooth enabled hardware be installed in the gallery [11].

Klopfer et al. proposed a "location aware field guide" which operated in a manner similar to *Musing* but it was not adapted to use in a museum [12].

Lee et al. used an ultra-mobile PC, inertia tracker and camera for object recognition [13]. This system did not rely on external devices; instead, it relied on template matching. In this case, a translucent image of the next artwork is placed on the screen, guiding the user to the next artwork to be matched and used to locate the user within the museum space, attempting to estimate the user's location by the last artwork scanned. However, this approach does not provide for an accurate location estimate. Furthermore, this project relied on proprietary hardware supplied by the institution.

Another system that used specialized hardware to provide an augmented reality experience is described in [14]. The system overlays the picture of a physical image displayed on a custom hardware with pertinent information in real-time. The detection of the artwork is accomplished using ultrasound sensors and gyros for pose tracking. The information is then matched to the image using an edge-detection algorithm.

IV. Application Development and Deployment

Musing, developed by an interdisciplinary team that included researchers within Computer Science, Communication Design, and Museology backgrounds, was deployed from October 8th, 2013 through November 14th, 2013, in The University Galleries at Texas State University, for the exhibition, *Eric Zimmerman: West of the Hudson* (example images, scanable by *Musing*, are available in [15]). During the 38-day run of the exhibit, 242 visitors downloaded *Musing*. In addition, 11 visitors borrowed iPod Touch devices provided by the galleries, indicating a high number of visitors used their personal devices. Gallery guest book logs showed that a minimum of 962 visitors attended the exhibit, resulting in 25% of visitors choosing to use *Musing*. This indicates a relatively strong initial acceptance rate of the concept. However, these figures do not account for repeat visitors, visitors who did not sign-in at the front desk, or visitors who shared devices.

A. Pedagogical Design

At the heart of the ideal 21st century museum/gallery experience is what educator and innovator John Dewey referred to over a century ago when he spoke of the importance of interactivity to provide for an enriched learning environment [3]. Such interactivity, and the resulting enrichment, requires providing for diverse learning styles by including visual/print, visual/picture, auditory, and verbal/kinesthetic modalities. These enriched learning environments are comprised of seeing, hearing, and interaction by moving beyond the traditional linear model of communication that provides didactic information via textual labels and gallery talks, to a non-linear model of communication through the provision of individual POI associated with each scanned artwork. Through the visitor's ability to access the POIs contained within *Musing*, the application allows for the creation of an enriched environment in which the visitors can participate in creating context for the works exhibited. The provision of additional information about each work via POIs, positions the visitor as a collaborator in the process of making meaning and serves to engage the visitor with the provided information which solidifies the content knowledge [3]. Meaning is made in a variety of ways and looking at art can begin by seeing the work through several different filters. The individual POI provides an opportunity to show the viewer

 ISBN: 978-1-61208-342-1

the works within an art historical, biographical, conceptual, or technical framework. As museums and galleries continue to seek ways in which the visitor's experience can be augmented and expanded, these POIs are an easy way to provide access for visitors to more contextual information for the exhibited works, broadening the exhibitions' theses for the novice viewer, as well as augmenting the meaning for the more initiated viewer. This extends the application's ability to meet the needs of a variety of visitors who access works on a multitude of levels. As such, the broadening of the exhibited works' context via interviews, videos, Web sites, source material, art historical influences, and other art with shared conceptual frameworks allows for a personalization for the visitor through the implied narratives [2]. This is thought to be the most effective way to expand the context for the work and deepen viewers' connections through the exercise and action of gathering the information, resulting in the visitors' "[integration of] the information into their existing world view" [2].

For the novice viewer, whose frame of reference may be lacking in depth to fully make these associations, the POI format is ideal to expand reference points. As these associations and connections deepen, the experience begins to look more familiar, something that can also make looking at art more comfortable. As museologist Marjorie Schwarzer writes, "Today, when the meaning of art is more contested than ever, [technologies] offer visitors the possibility of diverse interpretations" [16]. Schwarzer adds, "The branches of information available on these devices are close in spirit to the multiple ways in which we engage art" [16]. The ability to allow for different levels and a wide range of information, as well as a seemingly endless number of interpretive applications, reflects the diversity of the museum audience, itself [16]. Ultimately, the knowledge and deepened understanding that the POIs facilitate are filtered through the learning and innovation skills of the 21st Century—that of creativity and innovation, communication and collaboration, and cross-disciplinary thinking [2][3][16]. The resulting associations within the gallery setting, moving into the viewers' world, are essential to deepening the understanding of subject matter—a result of the user transferring what he or she already knows and reflecting upon it [3].

Musing's effectiveness comes from the immediacy with which the user can access the POIs content and making information available on demand allows for visitors to move freely within the space, not having to rely upon the preconceived schedule of their guide or any predetermined path.

B. User Interface Design

Musing was designed to employ a client-server architecture that allows museum administrators to upload, remove, and alter content, post-deployment. This is accomplished through an administrative Web interface which feeds the shared database. The application retrieves this content as requested by the user. This approach allows the material provided to the user to be as current as possible. Hence, the application is flexible and not limited to "on board" data, allowing any museum to more closely serve the needs of its visitors. The application relies on an open source library called *OpenCV* for the processing and recognition of images which have been captured by the user.

The User Interface was designed in such a way as to adhere to the Apple Human Interface Guidelines for a tab-bar navigation style application. The application consists of the Exhibitions Screen, Scan Artwork Screen, Artwork View Screen, and Favorites Screen.

C. The Exhibitions Screen and the Artwork View Screen

The Exhibitions Screen, depicted in Figure 1a, consists of a list-view of exhibits that a visitor may visit. The list is organized by "Permanent Exhibits" and "Augmented Reality Exhibits". The Permanent Exhibits are previews of the experience that visitors can expect when using the application in-gallery. They contain artworks that can be viewed outside of the gallery setting (e.g., residence, dorm, etc.). This type of exhibit is included to advertise the application's features, to familiarize the user with how the application works, and encourage users to attend a live exhibition. The AR exhibition section includes exhibits that must be attended in person to gain access to the didactic information for the artworks. This view provides information such as the name of the exhibit, the museum in which the exhibit is located (provided more than one organization uses *Musing*), and a representative image to advertise the exhibition. Figure 1b shows a portion of the "Art View" screen: a captured and identified image along with the overlaid POIs.

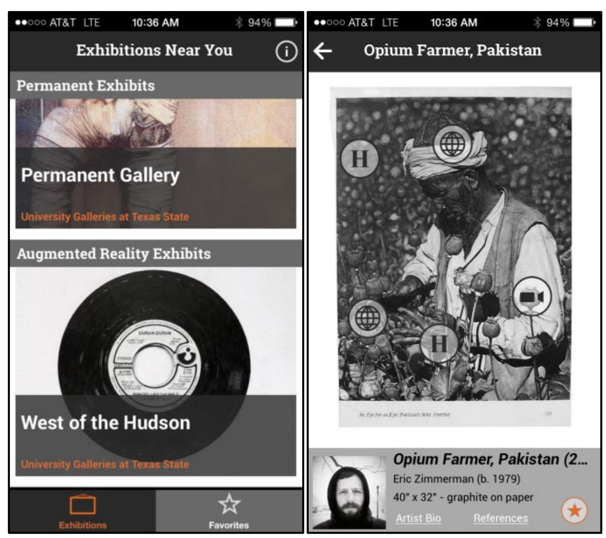

Figure 1: (a-left) Exhibitions Screen, including exhibition selection, and primary navigation; (b-right) A captured and identified image along with the overlaid POIs.

D. *Hardware/Software Architecture*

Currently, *Musing* runs on iOS-based hardware, such as iPhone, iPod Touch, and iPad. An Android version is under design.

1) *Back-end Processing*

The back-end (server) application provides two main functionalities. First, it supplies information in the form of reference images and relevant didactic information to the user, enabling its operation inside the gallery or remotely for a permanent exhibition. Second, the back-end is designed to provide an administrator (e.g., a museum staff member) with the capability to edit the contents of an exhibition's didactics within the system. The server, which is shared by the application and the administrative support back-end utility, is used by the gallery administrators to load content into *Musing*.

The back-end was written in PHP and uses standard web-technologies (including HTML, CSS, JavaScript, AJAX, jQuery, and several Open-Source JavaScript libraries) to deliver a user-centric experience. It is designed to allow users unfamiliar with database systems to create, read, update, and delete entries for exhibits from a database stored within the web application's framework. The entries include artworks contained within a chosen exhibit, the associated artists, and curated POIs.

Musing was developed with the intention of packaging within the application as little data as possible. When the user activates *Musing*, it requests an XML document containing a list of available exhibits from the back-end data server. The application parses the XML document and extracts the information into an Exhibit object within the application. Along with the XML document, which contains the names of the exhibits, locations, and id values which the application can use to retrieve data about specific exhibits, the application retrieves a “banner image” for each exhibit, which is displayed in a list for the user to browse.

When the user selects an exhibit from the list, the application passes its id value to a PHP script hosted on the data server. This process is referred to as ‘synching’. During synching, the server compiles the pertinent information and returns information in the form of XML file and a set of JPEG images of the gallery artworks to the app. The XML document contains information about each artwork, along with the set of POIs related to the information. The user can tap on POIs to display additional information about the artwork or artist. The images retrieved along with this document are used both for displaying POIs on the Artwork View screen and as references by the image recognition.

As in the case of the exhibit list, the XML document provided by the data server when the application is synched to a particular exhibit is parsed. The extracted information is used to populate painting and POIs within the application for each painting and POIs listed in the database. The images are also incorporated into these objects. Testing has shown that this process of synchronization typically takes approximately 20 seconds, during which time the user is shown a modal progress graphic.

The second functionality of the back-end is to support museum staff in modifying existing exhibition didactics within *Musing* and generating didactics for new exhibitions. This module is still under development. Nevertheless, the following is a description of current and planned functionality.

Artworks and information are added to the *Musing* database using a Web application that can only be accessed by specific museum/gallery staff members and by *Musing* developers.

After selecting an exhibit, the authenticated user is presented with a thumbnail for all of the artworks currently associated with that exhibit. This user is also given the option of adding a new artwork image to the exhibit within the systems. When a new work is added, the user selects an image of the art from local storage on their machine. The image is expected to be cropped such that only the artwork itself and its frame are shown. This greatly improves the recognition performance of *Musing* and creates a more fluid experience for users of the application.

When an image has been selected for a new artwork, the user is directed to a page where information regarding the particular artwork can be entered or edited. This same screen is reached when an existing work of art is selected from the exhibit listing. The user can enter the artwork's title, size, year of creation, medium, and the artist's name. Artists' information is stored and catalogued by the site and details such as year of birth, year of death if applicable, and a link to a biography, can be entered and saved and the user does not need to reenter this information.

Next, the administrative support utility enables the administrator to define and edit POIs for an artwork. This is done using a graphical interface designed with JQuery. The user selects a position on a displayed image of the artwork, chooses what media type that the POI references—along with its associated icon—and the text or URL as appropriate. Users can also alter the position of existing POIs by dragging and dropping them. The user can add and modify exhibits, as well as artists in a manner similar to that described for artworks.

2) *Front-end Processing*

As noted, *Musing* supports two types of exhibits—permanent and AR. The synching process is the same for both. If the database indicates that an exhibit is permanent, the user is shown a list of artworks available in an exhibit and each may be selected by tapping. This displays the artwork's image with the proper set of overlaid POIs. The second type of exhibit is the AR variety. In this case, the user is given an image detection view rather than a list, which displays a real-time feed from the devices camera over which is laid a graphic of an empty painting frame, along with a button which the user can use to capture a photograph.

During image detection, the users are instructed to position themselves so that a *Musing* enabled artwork fully fills the frame displayed (this is not mandatory, yet it can improve the recognition rate) on the device's screen and to take a picture of the artwork. When this is done and an image is captured, the application compares the captured image to each reference image currently synchronized for the exhibit. If a match can be made, the application proceeds to the Artwork View screen, exactly as it does when the user selects an image in a permanent exhibit. Otherwise, an error message is displayed in a modal dialog. To save in space, the captured image is discarded after being matched or rejected.

From the Artwork View screen, the user has the option of capturing the artwork and its information by making the artwork one of their "Favorites." This is the only condition under which *Musing* locally stores the artwork and its information. This is done by passing the image, POIs data, and artist information to a Favorites Database object that incorporates those values into an array of artwork objects. The data is then written into *Musing's* internal database. The information stored in the favorites array is accessible by the user regardless of whether or not the device is connected to the internet.

Image Processing and Recognition

Musing relies on the Oriented FAST and Rotated BRIEF (ORB) image detection algorithm [9]. The ORB procedure combines the "FAST" key-point detection and "BRIEF" determination of descriptors. Key-points are clusters of pixels within an image which are unusual enough to stand out and to help distinguish a particular image from other images. After identifying a set of key-points within an image, a set of descriptors is calculated for each key-point using BRIEF [17]. This functionality is provided by the *OpenCV* open source computer vision library which is available for use in iOS and Android devices.

Key-point detectors frequently rely on finding "corners" and "edges" within images since image boundaries often create distinguishable pairings of shade and color [17]. By definition, ORB is translation invariant. Additional operations are performed to compensate for rotation and scaling [9].

In the training stage, BRIEF employs binary comparisons between pixels in a smoothed image [17]. This algorithm takes a relatively large set of key-points—often as many as 500—and builds a classification tree for the set. The tree serves as an image "signature" used to measure similarities between images. Alternatively, under the approach used in this research, one can employ the results of the BRIEF stage using the *k* nearest neighbors (kNN) and one-to-one and onto mapping (bijection) test approach.

Following the synching process, users can point their device at an artwork in the gallery and capture its image. This image is processed using ORB and then compared to each of the reference images which were downloaded at sync time. Each reference image is processed to determine its key-points / descriptors at the time of comparison and this information is recalculated for each comparison. *Musing* employs the kNN and bijection approach to the key-points. Each key-point in a captured image is compared to each other in the reference image. A small set of matching key-points in the reference image is found for each key-point in the captured image. The goal is to find a maximal, high reliability, bijection between a subset of the key-points in a reference image and a subset of the key-points in the captured image. Hence, if any key-point in the reference image matches more than one key-point in the captured image with equal reliability, then *Musing* dismisses that match. The literature has suggested 0.65 as a reliability threshold and as the best threshold ratio for selecting one match as superior to the other [18]. The kNN is done twice, creating a set of directional matches that compares the reference image to the photograph taken and vice-versa. Then both sets are compared, dismissing any match that is not bidirectional. If a significant number of bidirectional matches is identified, the images are considered a match. *Musing* currently uses a threshold of 4 bidirectional matches as the minimum subset size.

When *Musing* has determined that a captured image matches a reference image, the reference image is displayed on screen along with an overlay of POIs.

The following is a description of the applied image recognition algorithm, starting with the captured image and the first reference image.

Step One: Captured Image Key-point Calculation - Find the key-points for the captured image using the FAST method [9]. This method checks a ring around each pixel and compares their intensities. It returns the point as a key-point if the gray level of a number of pixels within the ring is sufficiently higher or lower than the nucleus pixel itself.

Step Two: Captured Image Descriptor Calculation - BRIEF is used to take a patch of pixels surrounding a key-point and uses binary intensity thresholds to create a 256-bit binary vector describing the area around the key-point [9].

Steps Three & Four: Reference Key-points and Descriptors - Steps one and two are repeated for the reference image.

Step Five-A: Descriptor Matching (Captured to Reference) - A kNN matching of the Hamming Distances of each descriptor in the captured image to its K nearest neighbors in the reference image is performed. The two best matches for each key-point are retained.

Step Five-B: Descriptor Matching (Reference to Captured) - Step Five-A is applied with the roles of the captured and reference image reversed.

Step Six-A: Ratio-Test (Captured to Reference) - This step discards every match identified for the captured image where the best match and second-best match have similar Hamming distances. This produces a one-to-one match.

Step Six-B: Ratio-Test (Reference to Captured) - Weeding, using the same criteria as in step Six-A is performed on any match from the set of matches identified for the reference image.

Step Seven: Symmetry Cross-Check Test - The Symmetry cross-check test returns only the pairs of matches that are found from the captured image to the reference image and from the reference image to the captured image. This process enables keeping only the strongest symmetric correspondences and maintaining a bijection.

Step Eight: Output if Found - If four or more matches remain after the weeding performed by the ratio tests and symmetry test, the procedure retains the identity of the reference image and returns to step three for the next reference image (if such an image is available). The procedure keeps track of the identity of the image that produced the largest number of matches and outputs its id. If all reference images have been tested and no match has been found, then a message "Image Not Found" along with instructions to the user on how to improve the possibility of match are displayed.

Figure 2 illustrates the process performed in steps 5 to 7.

E. *Design of Testing Instruments*

Testing instruments consisted of quantitative benchmark testing and a qualitative user perception exit questionnaire.

As a part of the quantitative testing, each reference and captured image has been processed to generate 500 identifying key-points in each of 60 total images. The 60 images consist of: ten reference images ($R_1 - R_{10}$) and ten images that served as captured images($P_1 - P_{10}$). Each of the captured images was captured four additional times for a total of five capturing per image. The first time was with maximum alignment to the reference images the rest of the four where taken with increasing rotation translation and scaling (due to different distance). The maximal rotation was 40 degrees.

The procedure described above was applied to the ten reference images and fifty captured images. A threshold of 0.3% over the percent of matching key-points, which was empirically identified as the most suitable threshold was used by the program and applied to the matching results.

For the qualitative testing, we have used a 23-question exit questionnaire designed to capture feedback from in-gallery users. The questions were written to determine the user's acceptance of the application, their perceptions of application performance, enjoyment of the application, as well as pedagogical concerns.

V. DEPLOYMENT RESULTS

A. *Technical Results (Internal Testing)*

Figure 3 shows a heat-map of the results of this experiment. The figure shows a recognition rate of 96.4% with 0% error of type-1 (false positive) and 3.3% error of type-2 (false negative) obtained with $P_{(1,4)}$ and $P_{(1,5)}$. We have found however, that with rotation of more than 45 degrees there were numerous false negatives; but, still 0% of false positive error.

The testing has shown that *Musing* recognizes images with near perfect reliability under ideal conditions, that is, when a user is directly in front of the artwork, has positioned the artwork correctly within the image capture frame, and is not holding the device at an angle. Nevertheless, excessive rotation of the camera while capturing an image diminishes reliability. Our testing indicates that *Musing* recognizes images at a 45 degree rotation with 90% reliability and a 90 degree rotation with 84% reliability. The application performance degrades when the user stands off of the center line when photographing a piece of art, producing a skewed image. A slight deviation from the center (approximately 15 degrees) produced no noticeable change in testing but at greater values (approximately 45 degrees) the system produces 40% true positives and 60% false negatives. As far as can be determined, in the field-deployment testing, the system did not generate false positive results. Furthermore, the user surveys have indicated that the application did not produce a false positive error in use. Additionally, if the user stands too far from the artwork to properly fill the capture frame the reliability has suffered as well, with the reliability rate dropping to 48% at approximately twice the recommended distance. User surveys indicate that the application's reliability was sufficient to produce a positive experience for most users.

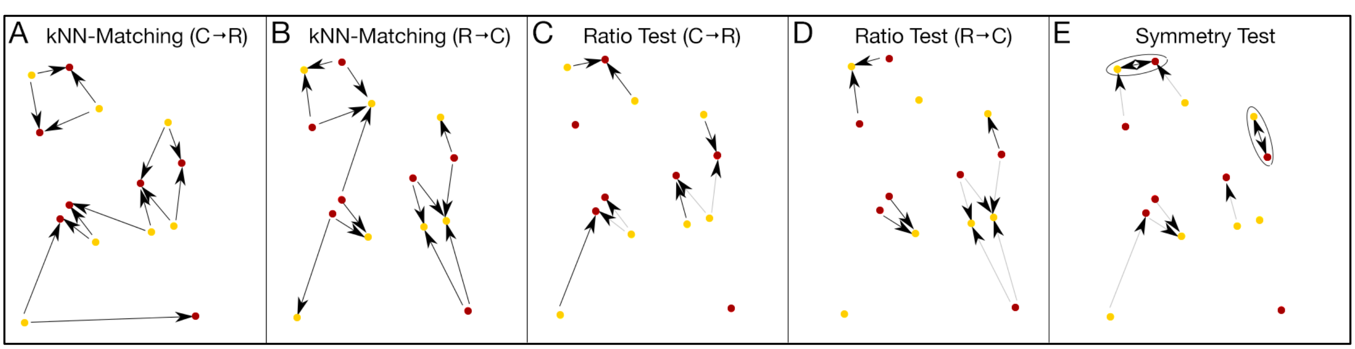

Figure 2: (A) and (B) kNN matching ($k = 2$); (C) and (D) Descriptor matching - the process discards matches with similar quality (Hamming distance) and retains the best match for distinctive matches; (E) Symmetry cross checking – only bidirectional matches are retained.

Testing performed to evaluate the processing time revealed that with 10 reference images, the application was able to compare and either display or reject an image in approximately 3.3 seconds on a stock iPod Touch-5. Again, user surveys indicate that this was sufficient to produce a positive experience for most users.

B. Exit Questionnaire Results with Live Users

Of the pertinent questions, 83.6% responded that *Musing* was able to recognize the artwork "every time" or "most of the time." 77.5% considered *Musing* to be quick and responsive. 87.7% considered *Musing* enjoyable to use and 93.8% wishing to see *Musing* in a future exhibit.

VI. RESULTS EVALUATION

The deployment results show high recognition accuracy and relatively short synching/recognition delay time, therefore the functionality of the entire system has been verified. The application has passed the Apple approval process and is available for download [4].

Formal user feedback obtained via questionnaire was consistent with our evaluation of the system and with informal feedback. Visitor responses to *Musing* characterized the application as informative and usable. Their perception of precision and timing was favorable and overall they have commended the system and expressed interest in its further use. Informal feedback from users, including staff members associated with other museums and galleries, was overwhelmingly positive.

VII. CONCLUSIONS AND FUTURE RESEARCH

We have designed, implemented, and deployed a usable mobile application that facilitates an enriched museum visitor experience via AR using interactive didactics. Per our assessment, the application has achieved its stated goals and has shown that the research hypothesis is valid.

The field testing via the exhibition shows that *Musing* can be used on non-proprietary smartphone hardware and provide visitors with didactic information, without the need for external tokens and reprogramming for information changes. This enables reduced reliance on loaner hardware. The implication of such is that the ease of in-gallery application of the technology may allow for higher levels of adoption by individual institutions.

A. Future Research

The University Galleries will be hosting another exhibition deploying *Musing* in the first quarter of 2014. This will provide an opportunity to further asses the capabilities of *Musing*—in specific, several capabilities that have been designed after the first deployment, including the administrative support part of the back-end of the system. This administrative site will allow the application to be deployed in independent galleries and museums by middle to late 2014.

Future enhancements to the *Musing* smartphone application (client) will include abilities for users to share images and didactics via social media such as *FaceBook* and Twitter, as well as the ability to comment on artworks within the application. Additionally, there are plans to complete a port of the current iOS-based implementation to the Android environment.

Other plans for future activities include expanding the image processing capabilities by further improving recognition accuracy, resilience, and time performance. Lastly, we plan to investigate the integration of algorithms for recognition of 3-D objects using the smartphone/tablet camera.

ACKNOWLEDGEMENT

The research team would like to thank Texas State University's Research Enhancement Grant program for providing the initial funding for this research. In addition, continued project funding was provided by the office of the Vice President of Research, the Dean of the College of Fine Arts and Communication, the Dean of the College of Science and Engineering, the Director of the School of Art and Design, and the Chair of the Department of Computer Science of Texas State University.

REFERENCES

[1] S. Sayre, "Assuring the Successful Integration of Multimedia Technology in an Art Museum Environment," in S. Thomas and A. Mintz (Eds.), The Virtual And The Real: Media In The Museum, Washington, D.C., 1998, pp. 1-10.

[2] K. Morrissey and D. Worts, "A Place For The Muses? Negotiating The Role Of Technology In Museums," in S. Thomas and A. Mintz (Eds.), The Virtual And The Real: Media In The Museum, Washington, D.C., 1998, pp. 147-171.

[3] T. C. Clapper, "The Enriched Environment: Making Multiple Connections" in The Academic Leadership Journal, 8(4), 2010, pp. 1-2.

[4] *Musing*, a photo recognition application that allows users to scan artwork at participating museums and art galleries to learn more about the work, Apple Store, https://itunes.apple.com/us/app/musing/id694382407?ls=1&mt=8, [retrieved March 2014.]

[5] G. Lawrence and M. Stump, "Connecting Physical and Digital Worlds. A Case Study of Quick Response Codes and Social Media in a Gallery Setting," The International Journal of Design in Society, 6(3), 2013, pp. 79-95.

[6] O. Bimber and E. Bruns, "PhoneGuide: Adaptive Image Classification for Mobile Museum Guidance," IEEE International Symposium on Ubiquitous Virtual Reality, Jeju, South Korea, 2011, pp.1-4.

[7] M. Hatala, L. Kalantari, R. Wakkary, and K. Newby, "Ontology And Rule Based Retrieval Of Sound Objects In Augmented Audio Reality System For Museum Visitors," *ACM symposium on Applied computing*, New York, NY, 2004, pp. 1045-1050.

[8] C. Jing, G. Junwei, and W. Yongtian, "Mobile Augmented REality System For Personal Museum Tour Guide Applications", IET Wireless and Mobile Computing, Shanghai, China, 2011, pp. 262 – 265.

[9] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: an Efficient Alternative to SIFT or SURF" IEEE

International Conference on Computer Vision, Barcelona, Spain, 2011, pp. 2564-2571.

[10] M. Blöckner, S. Danti, J. Forrai, G. Broll, and A. De Luca, “Please Touch the Exhibits!: Using NFC-based Interaction for Exploring a Museum,” *International Conference on Human-Computer Interaction with Mobile Devices and Services*, New York, 2011, Article 71, pp. 1-2.

[11] T. Miyashitat, et al., “An Augmented REality Museum Guide, in IEEE International Symposium on Mixed and Augmented Reality, Cambridge, 2008, pp. 103 – 106.

[12] E. Klopfer and K. Squire, “Environmental Detectives—The Development of an Augmented Reality Platform for Environmental Simulations,” in Educational Technology Research and Development, 56(2), 2008, pp.203-228.

[13] D. Lee, and J. Park, “Augmented Reality based Museum Guidance System for Selective Viewings,” IEEE Workshop on Digital Media and its Application in Museum & Heritage, 2007, Chongign, China, pp. 379-382.

[14] J. Oh et al., “Efficient Mobile Museum Guidance System Using Augmented Reality,” IEEE International Symposium on Consumer Electronics, Vilamoura, Portugal, 2008, pp.1,4, 14-16.

[15] *Eric Zimmerman: West of the Hudson*, example images, scanable by *Musing*, http://www.musingapp.com/test_images/, [retrieved March 2014].

[16] M. Schwarzer, “Art & Gadgetry: The Future of the Museum Visit”, Museum News. http://www.aam-us.org/pubs/mn/MN_JA01_ArtGadgetry.cfm, [retrieved, March, 2014].

[17] M. Calonder et al., “BRIEF: Computing a Local Binary Descriptor Very Fast,” IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(7), 2012, pp. 1281-1298.

[18] E. Rosten, R. Porter, and T. Drummond, “Faster and Better: A Machine Learning Approach to Corner Detection,” Ieee Transactions On Pattern Analysis And Machine Intelligence, 32(1), 2010, pp. 105-119.

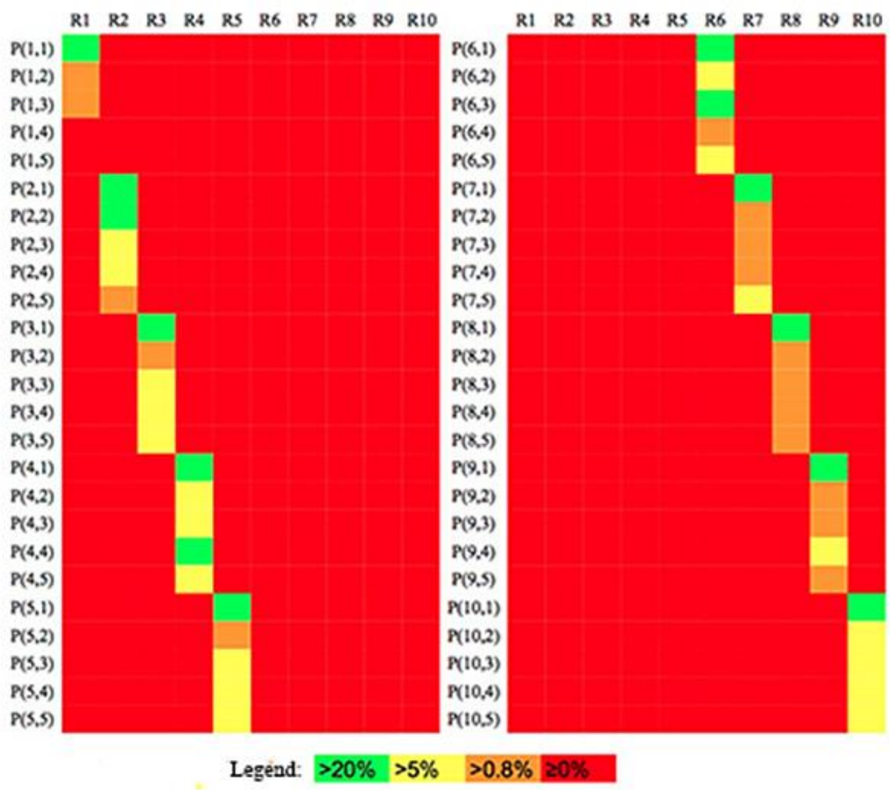

Figure 3: A heat-map of the results of the image matching experiment.