Exhibit recognition through image processing in museum tours

Guțanu Nicola-Carina

Abstract:

Museum tour guides are one of the most important parts of the experience of going to a museum, but they are rather an afterthought to the management. The same old system has been in place for some years now. This paper takes a look into the usefulness of a system using image recognition to discern exhibits in a museum, more specifically art pieces. In order to assess the public's openness to a new form of museum tours, we have conducted a series of surveys. Having received positive reviews, we are hopeful this approach will soon be implemented and distributed in all museums around the globe.

Classification:

- ***ACM:***

I. Computing Methodologies:

I.4 IMAGE PROCESSING AND COMPUTER VISION:

I.4.8 Scene Analysis: *Object recognition.*

- ***ASM:***

54-XX General topology: 54Hxx Connections of general topology with other structures, applications:

54H30 Applications of general topology to computer science (eg. digital topology, image processing)

Introduction:

With the new advancements that mobile phones have seen in the past few years, phone cameras are only getting better and better. We should take advantage of that in every way possible. Nowadays, most if not all people, have some sort of phone. It's the 21st century's sickness, but is it really? For all the downsides of having a phone, we need to take into consideration that they help keep us in contact with people, create all sorts of content, keep us connected to the world and, most importantly, they can be a tool for education. That is why we came up with the idea of introducing image recognition on phone cameras for recognizing exhibits in museums.

It seems only natural that with the most recent developments in technology, all things around us are in need of an upgrade. We would like to look at guided tours in museums and

how inconvenient they've become when we have these tools that could make it so much easier. Considering that phones are a must for the general population, it would be so easy to integrate a guided tour system for phones instead of using the traditional approaches found in museums at this moment.

In the context of the current pandemic, this would be even more practical as we would no longer need a person to hold the tour or give out equipment that would later need sanitizing. We are looking to replace all types of tours (person guided tours, sheets of information in front of exhibits, devices that talk about the exhibit, interactive computers with the museum's layout) with this new approach. This would then, most likely, reduce costs for museums since there would be no need for as much staff or for devices that could present the tours. It could also possibly increase interest for guided tours since it would be so easy to start one, as easy as one click away on your phone.

This is a rather new approach to guided tours, so there hasn't been much research done on this specific topic. What research there is already has had problems with the software available at the time of the research being conducted or they have taken a different route other than image recognition, also because of technology advancements at the time. We will be working with the best knowledge so far discovered and the newest hardware on the market.

We are looking to see how people would interact with such a system and how it would be received by the general population by using a mock system with existing software. This way, we can gage the reaction, come up with solutions for possible complaints and generally make a plan for the app on how it's supposed to work and what its goals would be without having invested or coded much yet. We are also interested in finding how relevant such a system is for disabled people and how it could be adjusted to be more inclusive, as well as how children would find navigating it without the help of an adult.

Research questions:

1. Do the users find the application intuitive?
2. Do users prefer using an app compared to normal guided tours?

Related work:

This part of the computer science domain is still rather untried, mostly because it is very expensive and needs a lot of adjusting. But some attempts similar to the one we proposed have been made in the past few years.

A study made by Sarah LeMire, Stephanie J. Graves, Michael Hawkins & Shweta Kailani called "*Libr-AR-y Tours: Increasing engagement and scalability of library tours using augmented reality*" said "*When scanning an object, multiple vantage points and perspectives were captured by walking all the way around each item to ensure that the stop could be triggered from all possible directions.*", which would be a great solution to no longer having to stand at a specified angle in front of an art piece in order for the software to recognise it. The study also showed that people were more likely to negatively review the app because they had to backtrack (and therefore walk more) when wanting to go back to a previous object. So that is another thing to be taken into consideration when preparing the tour so as to not make it sctrily liniar.

Another study made by Benjamin B. Bederson called "*Audio Augmented Reality: A Prototype Automated Tour Guide*" showed that through augmented reality, we can make a guided museum visit more pleasant by making it more social. The study emphasized that taped audio guides take away from the social interactions we would be having if not restricted by the guides "*Taped tour guides conflict with these goals because the tapes are linear, preplanned, and go at their own pace. This makes it hard to stay with friends because if one person turns off their tape temporarily, it is very difficult to get synchronized again*".

A study made by Costas Boletsis and Dimitra Chassanidou called "*Audio Augmented Reality in Public Transport for Exploring Tourist Sites*" made it obvious that having structure in a guided tour should be a priority when developing our app "*However, study results show that the possible combinations of familiarity and novelty sought by different tourists need further addressing; thus, along with the free unstructured exploration of urban destinations that audio AR can support, tourists may also require some kind of structure in their city tours.*" As a result, we need to take into consideration that art pieces in a museum are, usually, placed in a certain way to tell a story, whatever that may be. Therefore, we will give the option of taking the guide's suggested path through the room/museum or allow the tourist to choose his/her own way.

Another study conducted by V. Vlahakis, T. Pliakas, A. Demiris and N. Ioannidis called "*Design and Application of an Augmented Reality System for continuous, context-sensitive guided tours of indoor and outdoor cultural sites and museums*" proposed a close approach to what our app would like to eventually be. Our app would be an upgrade for their approach since their hardware is extremely old and not reliable in terms of battery "*The integrated mobile device offers autonomy of over 1 hour between battery recharges and weighs less than 5 kg. It is carried in a small backpack for easier transportation and protection against damage.".* The study also proposes timing how long a tourist is staying at a certain exhibit and presenting it with more and more detailed information "*The more time spent, the more additional information was given in the form of narration in his language of preference.*"

Another study closely related to our project is "*IMAGE-GUIDED TOURS: FAST-APPROXIMATED SIFT WITH U-SURF FEATURES*" written by Eric Chu, Erin Hsu, Sandy Yu. They proposed an algorithm for image recognition similar to both SIFT and SURF algorithms, but that ends up being different from both in multiple areas "*Several algorithms have been established for feature detection, including Scale-Invariant Feature Detection (SIFT) [4][5][6] and Speeded Up Robust Features (SURF) [7]. Both algorithms create an image pyramid, filter the pyramids appropriately to detect points of interest, quantify the features, then compare with features of another image to find matches. The algorithm described is based on this same structure but deviates from both SIFT and SURF in different areas.*"

And multiple articles and studies, such as "*Recognizing Art Pieces in Subway Using Computer Vision* ", "*Design and Application of an Augmented Reality System for continuous, context-sensitive guided tours of indoor and outdoor cultural sites and museums* ", "*Audio Augmented Reality in Public Transport for Exploring Tourist Sites*" to name a few, discussed the importance of having user tracking mechanisms in order to be able to recommend the next object to be viewed by them. This would work by tracking how, where and when the viewer moves in order to best recommend the next exhibit.

Original approach:

This study is aimed at finding the public openness to a new application used for museum tours. The website we would like to create would present as a global database containing all exhibits in the museum and information about them. Using a mobile phone, the user would log into this website and receive information about the exhibits based on the phone camera and location. In other words, the user points the camera at the exhibit, the application recognizes and finds that exhibit in the database and displays all needed information on the screen and also plays a recording of that information.

Compared to other similar applications, we have decided to firstly conduct the survey in order to assess the public's needs. Using other similar papers for reference and gathering information directly, we were able to identify a few functionalities that would drastically improve the user's experience: the possibility to pause and rewind the recordings, the possibility to go back to a previous exhibit without having to finish the tour or start from the beginning, the possibility to take multiple paths throughout the museum (in order the view the exhibit in a logical, story like progression).

Experiments:

The needed equipment:

In order to not waste money and resources into a project that might not be feasible, we can use existing software to conduct our first experiments. Most phones nowadays have camera software that recognizes QR codes, so we can use that in order to check how easy it would be for people to try and recognize said "art pieces in the museum". No other equipment needed other than the participant's personal phone.

Setting the scene:

In order to be able to conduct any experiment, we first need to set the scene. As the main idea for this project is image recognition, we'd need to set up some things to be recognized.

In order to make this as real as possible, we would need a space with multiple rooms in which we would set up QR codes around the rooms to be recognized by the participants of this experiment.

We also want to check the angle at which an image can be easily recognized, so each "art piece" will have placed before it a 180 degree scale. Participants at the experiment will be asked to remember what degree they were on if the camera could not identify the image.

Because we also want to check how easy it would be for people to log onto the "museum's" website, at the entrance of the space, we would place a register that would hand out tickets with the website QR code and with the password. With this, the scene is set.

The participants:

We would need a group of people of all ages possible that are not physically or mentally impaired in any way and from different countries. We would be excluding kids under the age of 12, because they will be a part of another group of people needed for experiments.

The next group of participants would be made up of people with different disabilities, mental or physical, aged above 20 years old, also from different countries.

And our last group of participants would be kids between the ages of 6 and 12, also from different countries.

Actual experiments:

- For the first experiment, we will be testing the ease of accessing the website with the QR code and password provided on the ticket. We will be making this experiment, separately, with every group of participants.
- For the second experiment, all groups of participants, also separately, will be going through the mock up “museum” and trying to recognize as many “art pieces” (QR codes) as possible. They will be advised to try sitting closer and farther from the codes and at different angles.

Original approaches (term of comparison):

- The closest approach would be a hand held device, so while walking through the museum, you’re essentially holding a device similar to a phone to your ear the whole time and listening to someone explain what’s in front of you. What’s more, this device needs you to personally dial the attraction number and then listen.
- Other guides one could come across are those that are placed physically next to the attraction so you can read off of it .
- Another approach would be guides that look like a computer game that allow one to move through the museum and read about the attractions, also static and placed in a few spots, not all over the museum
- And the last found approach is the guided tour held by a person.

Expectations:

- The first experiment conducted with the first and second group of participants will reveal how many people own phones that are advanced enough to be able to take part in the guided tour since there were people invited from all groups of ages.
- The first experiment conducted with the first group of participants will also reveal the ease with which day-to-day people found accessing the website in order to be able to use the guided tour. We are looking to find how easily elderly people would find using this method of logging since they might not know how to use the camera software.
- The first experiment conducted with the second group of participants will reveal if the software is a viable option for disabled people since not all of them might be able to use their arms. We will also consider how many of them are accompanied and if they are accompanied everywhere all the time.

- The first experiment conducted with the last group of participants will reveal the general age under which a child does not own a phone of his own or knows how to use one or does not own a phone advanced enough. This will later be used by the museum in order to decide the age under which a child can enter for free (and, as a result, the guided tour will be presented by their parents/teachers).
- The first experiment conducted with all the groups of participants will also reveal if English is a language known widely enough (which means there's no need for other translations) and what are the most demanded languages for the guide.
- The second experiment conducted with the first group of participants will reveal multiple parameters for closeness and angle at which the software will perform best, as well as ease of use.
- The second experiment conducted with the second group of participants will reveal the ease with which the software can be used by disable people, as well as any changes needed to be made in order to be useful for them.
- The second experiment conducted with the last group of participants will reveal if it's easy enough to navigate for a child and decide whether the age limit should be lifted if it is too hard for them to use.

Other results:

- In these experiments, we will also be checking if the participants of any of the groups would have preferred the old approach to guided tours (described above).
- And whether they find this approach more comfortable since they are using their own property, especially taking into consideration the current global pandemic.

Case study:

Initial data set:

Since we will need participants from different countries, we will consider all tourists from the year 2019 in Romania: 13.37 million. This is the total number of people we are trying to study.

Since it is impossible to ask so many people to take part in the experiments, we can definitely expect some margin of error, which we will need to calculate later.

We also need to take into consideration the confidence level, which will tell us if the results obtained from the sample size are true for the whole population, taking into consideration our margin of error.

We'd like to reach a 90%-95% significance level and if we choose a margin of about 3%, we would need about 1000-1200 people in each of our groups for the experiments. As a result, we will have 1000 people in each of the groups.

Experiment:

In order to find all the data needed for the research, we will compose multiple surveys for each experiment conducted and each group. The surveys would have a common part, but may

deviate from one another in order to be inclusive to all participants and find all the results in order to reach a conclusion. Kids' surveys will be handed to parents/teachers, which will then decide if they would be the ones to complete the survey or if the child is old enough to complete it on their own.

Questions:

1. How old are you?
2. Sex? (Male/Female/Rather not say)
3. Do you speak English? (Yes/No)
4. What is your main language?
5. Do you own a phone? (Yes/No)
6. Do you own a phone with a camera? (Yes/No)
7. Does the camera have software that recognizes QR codes? (Yes/No)
8. Have you ever used your phone camera before? (Yes/No)
9. If not, how easy was it for you to find the camera? (Easy/Had some problems, but managed/Needed help)
10. Have you used the QR identifying software before? (Yes/No)
11. If not, was it easy to use at first try? (Easy/Had some problems, but managed/Needed help)
12. How did you find accessing the website through the QR code? (Easy/Had some problems, but managed/Needed help)
13. Did you need any help logging onto the website with the given password? (Yes/No)
14. If the camera did not manage to recognize the image, could you estimate (in meters) your distance from the image?
15. If the camera did not manage to recognize the image, please write down the angle at which you were on the scale.
16. Did you find the whole process of recognizing an image easy (intuitive)? (Easy/Had some problems, but managed/Needed help)
17. If disabled, were you able to log in without anyone's help? (Yes/No)
18. If disabled, are you accompanied by anyone right now (that is there just to help balance your impairment)? (Yes/No)
19. If disabled, are you always accompanied by someone (that is there just to help balance your impairment)? (Yes/No)
20. If disabled, was it easy for you to use the software to recognize the images? (Easy/Had some problems, but managed/Needed help)
21. If disabled, any recommendations for us for future upgrades? (Free answer)
22. If you are a child, what age were you when you received your first phone?
23. If you are a child, what age did you learn to use a phone?
24. If you are a child, do you know how to read? (Yes/No)
25. Any feedback/recommendations/complaints for us? (Free answer)
26. Compared to other guided tours you had in the past, would you consider this more comfortable for you (taking into consideration the lack of a guide

that would dictate the pace of the tour and any other factors you've come across before)? (Yes/No)

27. Is this approach more to your liking considering the fact that you are using your own property and know its status of cleanliness? (Covid 19 related question) (Yes/No)

Results:

With the answers gathered from the 3000 participants (1000 for each group) from the survey, we have found the following:

- Out of the total of 3000 participants, 2186 (72%) speak English while 814 (28%) do not.
- Out of those 28% that do not speak English, 193 speak Romanian, 178 speak Hungarian, 111 speak Bulgarian, 102 speak Italian, 74 speak German, 60 speak Serbian, 37 speak Hebrew, 23 speak Polish, 20 speak Ukrainian and 16 speak Turkish.
- Of all the participants, 2670 (~89%) own phones while the rest of 330 (~11%) do not.
- And out of the 2670 that owned a phone, only 2394 (~90%) had a camera, the rest of 276 (~10%) did not.
- And out of those 2394, 2027 (~85%) had qr code recognizing software, while the rest of 367(~15%) did not. The last 15% were mostly made up of elderly people (people over 65).

- Out of all the participants, 2190 (~73%) have used a camera before while 810 (~27%), mostly elderly, have never used one.
- Out of those 810 that have never used a camera before, 200 (~23%) found it easy to find, 521 (~59%) managed in the end while 89 (~18%) of them needed assistance.

- Out of the total of 3000, 1272 (~42%) people have used the qr code identifying software in the past, while the rest of 1728 (~58%) have never used it, the majority of them being elderly once again.
- And out of the 1728 that have never used the software before, 363 (~29%) found it easy to use at first try, 718 (~42%) of them managed in the end while the rest of 647 (~29%) needed help.

- Accessing the website came easily to 1383 (~40%) out of the 3000 people interviewed, while 836 (~28%) managed on their own in the end and 781 (~32%) ended up needing help in order to access it, mostly elderly people.
- After accessing the website, most of the participants had on trouble logging onto it (only about 8% needed help and they were mostly elderly).

- As for the positioning of the participants, 163 (~5%) found that they were standing too close to the image, 361 (~12%) were standing too far, while the rest of the participants, 2476 of them, reported no problems.
- As for the angle at which they were positioned, 593 (~20%) were sitting too much to the right, 472 (~16%) were positioned too much to the left, while 1935 had no problem at all.

- The survey also shows that about 64% of the participants found this an intuitive process with only about 23% needing any help.

- Out of the 1000 participants invited to this experiment, 209 (~21%) of them required help (special help for disability) for logging onto the website.
- Out of those 209, only 75 (~35%) were accompanied by someone at the moment of the experiment.
- And out of those 75, only 12 (16%) are always accompanied by someone.
- Also out of the 1000 disabled people we interviewed, 646 (~65%) found the software easy to use, while 127 (~13%) managed to use it in the end and 227 (~22%) needed assistance (because of their impairment).

- Out of the 1000 kids we have interviewed, we found that the age at which a child first received a phone is pretty well balanced between ages 6-11. Age 6 was the one in which most kids received their first phone, 182 kids, but the rest of the ages are also around that number.
- We also found that most kids learn to use a phone at age 9 (191 kids), but all ages are pretty balanced, just as above.
- And we also found that out of the 1000 kids, 780 (78%) know how to read, while the rest 220 (22%) do not.

- Out of the 3000 people, we also discovered that 2067 (~69%) would consider this a more comfortable approach.
- And we also discovered that 2654 (~88%) liked this approach to guided tours. In the context of the global pandemic.

Conclusions:

The survey questions answered by all the participants after going through the experiments revealed the following conclusions:

- While most of the participants spoke and understood English, there were some that did not. This is not that big of an impediment seeing as the website opens up in an online browser that usually would translate pages to whatever language needed. But this is good to be taken into consideration for future updates in order to try and be more inclusive to all people of the globe. The survey also revealed the most on demand languages for the future upgrade.

- Another thing that was revealed and needs to be taken into consideration when building this software are the number of participants that own a phone with a working camera and that has the software needed. It was shown that around 85% of the participants had access to the needed device in order to fully enjoy the tour.
- One other important thing to be taken into consideration would be the people that have never used their camera or image recognition software before. They could find this whole experience disorienting and confusing and hard to adjust to in order to fully enjoy the tour. Although, looking at the survey answers, this would not seem to be a problem since most of the people that have never worked with a camera have managed to use it in the end, with only about 18% of them needing to ask for help. The same goes for the people that have never used image recognition software before, only about 29% needed help.
- The process of accessing the site and using the password was troubling to few people and most of them were elderly, meaning they most likely have never needed to work with something like this before. In the end, about 30% of the people needed help accessing the website, while a lot less needed help inputting the password.
- The survey also shows that more than half of the participants found this whole experience pretty intuitive once they got used to the software and the way it works.
- Moving on to people with disabilities, more than half found the software easy to use, taking into consideration their disability, while only about 20% of the 1000 interviewed ended up needing assistance because of their impairment. The survey also showed that of those participants that needed help, less than half were accompanied by someone at the moment of the experiment and out of those, about 15% are accompanied at all times. This is something to be taken into consideration by the museum staff for when the software would be introduced in the museum since they might need to hire people in order to help those that are disabled. Therefore, we have found the answer to the first research question: both disabled and able bodied users find the interface to be intuitive.
- As for the kids, it would seem that while most of them know how to read, maybe making the age limit 10 years old for kids would be a good idea. Handling a phone and the software could become confusing for a child, while also trying to concentrate on what he/she is seeing in the museum. This is also something to be taken into consideration by the museum staff: a ticket can be offered to parents/teachers that specifically ask for it for their child.
- We also wanted to know if people preferred this approach to guided tours more than any other they have encountered before, in order to answer our second research question. Results show that 69% prefer using the app, which allows us to conclude that integrating such a tool into museums would be well received by the general public. Moreover, this approach is even higher regarded due to the ongoing pandemic situation.

Bibliography:

1. **Sarah LeMire, Stephanie J. Graves, Michael Hawkins and Shweta Kailani: "***Libr-AR-y Tours: Increasing engagement and scalability of library tours using augmented reality*" published in College & Undergraduate Libraries, Volume 25, 2018 - Issue 3
2. **Benjamin B. Bederson:** "*Audio Augmented Reality: A Prototype Automated Tour Guide*" published in CHI'95 MOSAIC OF CREATIVITY, May 7-11 1995
3. **Costas Boletsis, Dimitra Chasanidou:** "*Audio Augmented Reality in Public Transport for Exploring Tourist Sites*" published in NordiCHI '18: Proceedings of the 10th Nordic Conference on Human-Computer Interaction, September 2018
4. **V. Vlahakis, T. Pliakas, A. Demiris, N. Ioannidis:** "*Design and Application of an Augmented Reality System for continuous, context-sensitive guided tours of indoor and outdoor cultural sites and museums*" published in 4th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage VAST, 2003
5. **Eric Chu, Erin Hsu, Sandy Yu:** "*IMAGE-GUIDED TOURS: FAST-APPROXIMATED SIFT WITH U-SURF FEATURES*" published in January 2007
6. **Tengjiao Cai:** "*Recognizing Art Pieces in Subway Using Computer Vision*" published in Examensarbete 30 hp, June 2011