

Virtual Backdrop

Edward Jezisek, Nichole Dugan, Brian Schulte

CprE 575

3/10/16

Project Concept

The team's Virtual Backdrop project will eliminate the need to spend hours at the Driver's License Station waiting for other driver's license staff to take photos. It does this by creating a professional grade driver's license photo using a picture from a cell phone. Benefits to this project include reduced costs in DMV staff, decreased time spent at the DMV, and a license photo you are proud to have. Virtual Backdrop will let you take advantage of this by taking your photo, removing the background and allowing you to choose your favorite photo before you even step into the Driver's License Station. Several attempts have been made at this concept in the past, but the main drawback to many of the alternatives is that hair is often cropped out of the picture.

Virtual Backdrop will specifically be used to separate the background from pictures/videos. This will allow it to be used in applications other than providing a backdrop for Driver's license photos. However, the primary purpose of this application is to allow users to use their own photos as their driver's license photo.

Target Audience

Self-service kiosks are becoming popular in states across the US to give citizens a convenient choice in renewing or replacing their driver's licenses. This project will focus only on kiosks being used by the state of Iowa. Currently, the following services are offered at kiosk locations:

- Renew/Replace Iowa Driver's License or identification card
- Update Iowa Driver's License (removing "under 18" or "under 21", updating address)

As a result of these select options, kiosk users must meet the following requirements:

- Possess a valid Iowa Driver's License or ID card (not a Commercial Driver's License)
- Be between the ages of 18 and 70
- Reside in Iowa with US citizenship
- Not be required to provide a medical or vision report during the renewal process

Although anyone between 18 and 70 can use the kiosks, the system is designed to target drivers between the ages of 18 and 30 who routinely download applications and tend to choose self-service over in person interactions. The typical kiosk user also likes to avoid time wasted at the Driver's License Station. We are specifically targeting users with Android based smart phones and anyone who currently uses a kiosk to replace their driver's license photo. At the current time, kiosks need to have a standing our mounted backdrop behind the photos and Virtual Backdrop could help eliminate this. (Iowa Department of Transportation)



Figure 1

Need for Application

Virtual Backdrop is needed to enhance the experience of getting a Driver's license for residents of the state of Iowa. This application can further be used to populate a database for facial recognition. The pictures we create will meet the standards for a driver's license and in doing so, will meet the standards for facial recognition.

With the introduction of new technologies that require a photo with a high enough quality to be included in biometric databases used for facial recognition, photos will need to remove artifacts from their backgrounds to create consistent images of faces. Previously, colored backdrops were used to ensure noise was reduced from the photos. But this places significant limitations on the photos that can be used in the databases and the process to take said photos (i.e. cellphone cameras cannot be used). Ideally, software could be created to identify faces, remove backgrounds and produce a uniform virtual backdrop for the photo regardless of the composition of the original photo. One specific hurdle to overcome is that current algorithms can sometimes get confused with a person's hair and fill in the hair with the backdrop color instead of leaving the alone.

Previous Approaches

In researching current algorithms for this process, we found 3 major ways of trying to solve this problem. Each of these algorithms appeared to have their own drawback from either having problems identifying hair properly, not removing the whole background, or other issues that arise from doing color matching on pixels.

The first way is an algorithm by the identity solution vendor, MorphoTrust USA. MorphoTrust currently provides physical credentials for 80% of the jurisdictions in the United States. MorphoTrust is looking to expand their business by providing this virtual backdrop solution. In a patent filed in 2015, the algorithm for the virtual backdrop solution is to try to estimate the foreground and background of the image. To determine the difference between the two, Alpha matting is applied to the image. Their algorithm also tries to estimate hair pixels, facial pixels, and other foreground pixels using a formula that utilizes a convex hull algorithm. It also uses a cross algorithm to figure out where the chin is located on the subject and again trying to figure out where the face exists in the foreground. The algorithm also tries to determine the background by using Eigenvectors and pixel intensity. It also uses neighborhood matching to try to determine similar pixels in a region. Their approach seems to be a good starting point, but as of yet, the company does not appear to have a working prototype. (MorphoTrust USA)

The second algorithm the team found was a paper from Incheon National University in Korean. The authors Jeong-In Park and Jin-Tak Choi used an algorithm that utilized a code book to try to determine smaller portions of the image to try to determine background for each smaller image, in a similar fashion to solving our Sudoku problem. Each of the smaller units is a vector, and uses vector rotation about a center of rotation for the vector. Reviewing the author's work, the background does not appear to be completely removed from the images, so the algorithms appear that they could use more work.

The authors also noted that their algorithms appear to be slower than they would optimally like. The paper also notes limitations of dealing with clusters of regions next to each other. Ideally, the smaller vectors would need to know something about the blocks around them in order to make better guesses about the background of the image as a whole. Splitting up the image processing into blocks would allow for the algorithm to be multi-threaded, which would give this approach an advantage over other algorithms. (Jeong-In Park)

The third algorithm we found was a paper written for Microsoft Research written by authors Jian Sun, Weiwei Zhang, Xiaoou Tang, and Heung-Yeung Shum. This paper deals with removing the background image in videos. Their algorithm again seeks to determine the difference between the foreground and the background of the image. The authors note that the background subtraction is relatively easy to do with videos if the algorithm has two cameras trained on an object, but note that most people do not have two webcams hooked up to their computers. The algorithm uses a Gaussian mixture model to try to determine if a pixel belongs to the foreground or the background. The authors note that sudden changes in lighting is the biggest hurdle to overcome with their algorithm. (Jian Sun)

Our team believes the shortcomings of these algorithms give us opportunities to improve on existing techniques. Our team will use our knowledge from this class in C++ and OpenCV to try to determine the subject by using video techniques in OpenCV. Our team also has a strong mathematical and programming background to give us a fresh perspective on this challenge.

Previous Experience

Our team is comprised of three people: Edward Jezisek, Nichole Dugan and Brian Schulte. Edward Jezisek has experience with mobile development and a strong interest in developing consumer based applications. He is currently a programmer for Verizon Wireless, and is incredibly interested in algorithms, data structures, overall systems architecture, and learning more. Some of his Android applications have been featured on Fox news and XDA-Developers.

Nichole Dugan is a C# developer working for the state of Iowa in the Department of Transportation. She has been working with driver's license issuance since 2007 and has helped with the facial recognition program in the state of Iowa. She has also worked with vendors for the state of Iowa implementing self-service kiosks for the state, and recently has worked with a vendor to interface with the state of Iowa's system of record for the mobile driver's license project.

Brian Schulte is a software engineer for General Dynamics Mission Systems. He has experience in C++ and JAVA. In addition to his programming skills, Brian has a minor in Math and strong skills and experience in developing algorithms. Brian's experience in systems engineering will also help with managing the project, project requirements and system testing.

Approach

Our proposed approach to this project is two-pronged. First, we want to research current algorithms

being used, as well determine if these algorithms could be modified to accomplish our goal. Secondly, we want to design a high-level process that is streamlined, scalable and secure for users to import pictures taken with many types of cameras in many environments, manipulate them easily (like clicking a button rather than manually altering), and save a picture that meets the necessary specifications into the database.

To accomplish this, we have divided our project into 4 phases.

- I. Phase 1 will involve developing a way for users to upload images that will be used for their license. Users will be able to use a camera to capture video or upload a previously saved photo. Once the picture is loaded into the application we will be able to manipulate it to conform to the Iowa DOT standards.
- II. Phase 2 of the project will be to create the algorithms used to alter the photographs. More specifically, during this phase algorithms will be developed to find the background, determine the edge of the face without cropping off hair, and replace the background of the photo from Phase I with a blue background.
- III. Once the algorithms are in place, Phase 3 will integrate Phase 1 and Phase 2 into an Android application. Users will be able to take photos using an android application and then submit them to be used for their driver's license.
- IV. The photos used for driver's license and identification cards must meet certain criteria. These criteria are defined by the AAMVA and the ICAO (International Civil Aviation Organization). These requirements include rules such as the following, no sun glasses, well-lit with minimum shadows, no hats, and many more. Figure 2 shows the photo requirements for passport photos imposed by the ICAO. During Phase 4, we will attempt to check the uploaded photo for several of these requirements. This phase is a stretch goal for our team. Because of this, some ICAO requirements may not be imposed. (American Association of Motor Vehicle Administrators)

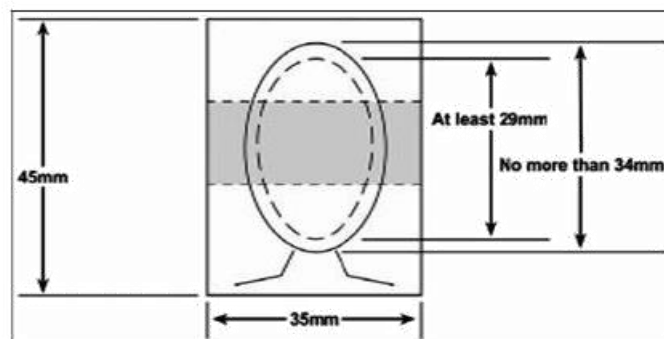


Figure 2

Overall Requirements

The AAMVA (American Association of Motor Vehicle Administrators) is the governing body regulating

driver's license documents. According to AAMVA standards, the background for driver's license photos is stated as, "Background. A uniform light blue color or white background shall be used to provide a contrast to the face and hair. Note: Preference is for uniform light blue color, such as Pantone 277 (though the specific Pantone color is not a requirement – a uniform light blue color or white background is a requirement". (American Association of Motor Vehicle Administrators)

An example photo from a driver's license is shown in Figure 3.



Figure 3

A typical driver's license station with the photo backdrop is shown in Figure 4. You can see from the image that the blue backdrop is placed behind the camera in order to ensure that the photo in Figure 3 contains the required background coloring.



Figure 4

When kiosks were introduced as a solution for citizens of the state of Iowa to be able to renew or replace their driver's licenses without visiting a driver's license station, certain limitations were required because of the backdrop functionality. The kiosk needs to have a standing or mounted backdrop behind the photo area to allow for the photo to have the required blue backdrop. News station KCCI reported on the state's installation of the kiosks in the state of Iowa, with their news story focusing on the kiosk in the West Des Moines library. In the video, you can see how the kiosk needs to have a particular

configuration in its location. (KCCI)

In addition to the kiosks, the state of Iowa has recently introduced a pilot program to allow for mobile driver's licenses. One of the suggested features has been allowing the user to take a photo from their device and use that as their driver's license photo.

Hardware Requirements

For this project a camera will be needed. Photos on driver's licenses need to be at least 72 dpi. Although dpi is a printer specification, the camera needs to be able to produce pictures with a ppi equivalent to the 72 dpi requirement. To produce this level of quality, the camera will need to support at least 1 megapixel. Since most cameras commercially available today are over 1 MP, almost any camera should be acceptable. These pictures will be approximately 4MB in size (or smaller). For our project to be successful; we need to be able to process all of the drivers in Iowa in a one month time period. As there are 2.3 million drivers we need to be able to process 8 Terabytes worth of pictures in a one month time period. In addition, if this is to scale to multiple states; we need to make sure that our solution is distributive and able to scale/be run on multiple machines. Our initial product will run on the Ubuntu O/S, but future releases may need to use Redhat to ensure scalability.

Software Requirements

Our application will be created using C++ and the gcc (Ubuntu 4.8.4-2ubuntu1~14.04) 4.8.4 compiler, the OpenCV 3.1.0 library and Boost. This will allow us to use the library based functions inside of OpenCV and develop intermediary tests for our application using the Boost::Test framework. Finally, we will provide several videos/pictures that can be used in our application to ensure that the provided photos meet necessary standards. Some of the processing will need to be done on the device as to prevent the server from receiving and processing too much of the required load.

This will require us to include parts of our code in Android. The specific version of Android that we will use is: 5.0.1. This will specifically run on an HTC One. It may run on other phones as well, but due to the scope of the project we aim to ensure that it at least works on an HTC One. Other applications may be developed to ensure access by as many users as possible. An Android phone was chosen for the convenience of having a device we already own that takes pictures/video and has a connection to the Internet. We felt that many of our customers would have a similar device and that they would be able to use the product.

Evaluation

To successfully complete our project, testing will need to be performed. This testing will initially consist of a smoke test which will test the primary features of the application. This will be followed by a regression test and finally a load test.

Test Plan

The smoke test will consist of an image being processed and displaying just the face. To do this automatically we will obtain an image that has been processed and its original input. We will compare the output to the result and have a threshold for differences. If there are too many differences, the smoke test will have failed. This will run for several small subsets of images. Furthermore, we will have tests that should fail. These tests will include pictures without anyone included. We will perform these tests with random objects entering the screen, and more than one individual in the picture. These tests will give a fast turnaround regarding whether or not a feature was successfully included. If the pictures are different enough a comparison image will be displayed to the developer. This will allow the developer to either accept or reject the new image.

Following the smoke test a regression test will be performed. This regression test will consist of the user manually checking many of the photos included to prevent inaccuracies. This test will highlight any key differences in pixels between the previous and current picture. This will allow the developer to easily notice any changes in the application; whether or not they are positive. The regression test will also take in at least one new video that the user performs and add this video to the list of tests to be performed in the full regression. This video must be taken with webcam to be determined as a success. The reason for this is to ensure that our application is able to extract an image from a new video as well as the old tests.

Finally a load test will be performed. This load test is needed to ensure that our server can handle the load of the inserted pictures. For this project we aim to be able to upload 100 pictures per hour for the server. If this is inserted in a kiosk, the kiosk may need a small computer or system to perform the processing. Hopefully our system runs faster; but approximately one photo processed per minute should be acceptable for our application.

Determination of Success

If all of the previously defined tests run successfully our application will be a success. The previous test cases will prove that our application is successful with video/pictures from a webcam/mobile phone and that our application is able to withstand a specific amount of load. The tests will prove accuracy as well as speed and will automatically deduce how accurate the application is. Finally, complete success would include being included in the DMVs kiosk based product. Our attempt should be able to fix many of the issues in current approaches.

Finally, if our solution's pictures meet the requirement for a driver's license in 90% of the pictures taken, our application will be a success. This will require continuous analysis after the completion of the project to ensure quality; but it will hopefully remove many of the issues apparent in the current approach. These pictures must be accessible to facial recognition and ensure that a user is the same as their previous photo.

Test Conditions

Our application will consist of a minimum of three videos/photos in the regression test. These videos/photos will be of each team member. The reason for this is to ensure the quality of our solution. We will likely add more videos/photos as test evidence, but at this point three photos/videos will be required for the successful completion of our application.

Test Subjects

As our application will consist of a minimum of three videos/photos in the regression test; we will require a separate subject for each of these pictures/videos. This will ensure there are at least three subjects used as a test subject. This will ensure our application works for more than one person and will be convenient as we are a group of three people. More people will be added and more videos will be taken with respect to the necessary regression test.

Evaluation of Results

Our results will be evaluated by comparing what we have created with a standard driver's license photo. If our application falls short of creating acceptable photos with a blue backdrop, the problem will need to be investigated. Furthermore, if it's necessary, we will also perform facial recognition on past photos to ensure that users are unable to replace their photo with a picture of someone else. Our results will be evaluated based on how accurately we can create a driver's license photo with a background. This background will be of varying colors. Some backgrounds will be unacceptable due to them being close in color to hair or skin. If the display does not change, it may be impossible to recognize a face correctly.

The initial results of our application will be compared manually as a final evaluation metric. However, in the future we will automatically compare previous successful results with the most recent result. This will allow us to potentially speed up development and create an acceptable solution to this problem.

Improving the Application

After our initial goals are met, we will try to improve our application. We will ensure that our results are comparable with the previous results or manually input them as the new base image. This will allow for fast development and it will ensure the quality of our application. After our application is successful with a large majority of faces, the application will no longer need to be "improved" and will be ready for production implementation.

If we are able to meet the first phases of our project early, we will look into additional features. Specifically, we would like to include having the application running on multiple Android devices and OSes. Also, as we want to improve the perception a user has over their picture; we will potentially allow them to do some image editing on the final picture. This editing naturally must be compliant with the aforementioned requirements.

We will also maintain compliance to the ICAO standards. This will ensure that our product is usable and acceptable to the state of Iowa. If performance becomes an issue, we will measure various aspects of our application and make modifications to provide a quality project. This will provide for a solution that meets federal requirements and will ensure that our project is a success. These standards include the size of a picture needed for a driver's license and the background image that needs to be included. We hope to print at least one imaginary driver's license to prove the concept we have created. We plan to display video of us taking the picture, inputting it into a driver's license format, and displaying the resulting image.

Preliminary Results

Coding has already begun on the Virtual Backdrop program. The coding is using the MOG2 background subtraction algorithm available in OpenCV. In addition, the project is using a facial detecting algorithm using Haar Cascades. These combined approaches try to determine the background of the image from the video by focusing on the non-moving aspects in the background versus the moving aspect in the foreground. The algorithm secondarily uses facial recognition and determines that everything within the bounding ellipse of the face is not in the background.

Preliminary results of testing with a background subtraction yielded the results shown in Figure 5. As you can see, the algorithm while doing a fairly decent job, still improperly identifies some of the subjects facial features as being part of the background. It also incorrectly determines pieces of the background are in the foreground.



Figure 5

Additional Tools

In addition to the software requirements and tools listed above, our team will utilize a few other tools to assist with project management and development.

GitHub will be used for version control. All documents, presentations, and source code will be maintained by version control. This ensures collaboration between team members is done efficiently as possible and that all members of the team can participate by adding their changes easily.

Microsoft Office will be used for developing project papers and presentations. All members of the team had access to the Microsoft Office suite by virtue of being Iowa State students. Word is also a standard for writing of documents in most environments.

Project deadlines and issues will be tracked in Yodiz, a web-based project management tool. This allows the team to assign specific tasks to each member of the group and ensure that deadlines are met.

Conclusion

This project will have useful applications not only for the citizens of the state of Iowa, but also for all citizens if the application can have widespread adoption. We believe that our team's combination of now only computer science and engineering techniques give us an advantage in figuring out a solution to this problem, but researching how other teams have tried to solve this problem unsuccessfully and improving on their techniques. Also by using industry standard tools such as Git, Microsoft Office, and issue tracking software, we believe we will be able to manage our project effectively and use our resources to their full potential. Our team looks forward to using the techniques learned in this class to apply to a real world problem and achieve a working solution.

References

- American Association of Motor Vehicle Administrators. "AAMVA DL/ID Card Design Standard." 10 February 2016. *AAMVA*. <<http://www.aamva.org/WorkArea/DownloadAsset.aspx?id=4435>>.
- Iowa Department of Transportation. *Kiosks*. 20 February 2016. Web site. 20 February 2016. <<http://www.iowadot.gov/mvd/kiosks.html>>.
- Jeong-In Park, Jin-Tak Choi. "A Background Removal Algorithm using the Variable Order nxm Dimensional Vector." 19 February 2016. *SERCS*. <http://www.sersc.org/journals/IJSEIA/vol7_no6_2013/22.pdf>.
- Jian Sun, Weiwei Zhang, Xiaou Tang, Heung-Yeung Shum. "Background Cut." 19 February 2016. *Microsoft*. <http://research.microsoft.com/en-us/um/people/jiansun/papers/BgCut_Camera.pdf>.
- KCCI. *List of locations released for new driver's license kiosks*. 18 February 2016. <<http://www.kcci.com/news/new-kiosks-let-you-renew-your-iowa-drivers-license/33374300>>.
- MorphoTrust USA. "System and Method for Creating a Virtual Backdrop." 19 February 2016. *Free Patents Online*. <<http://www.freepatentsonline.com/20150030213.pdf>>.
- Wikipedia. *Background subtraction*. 19 February 2016. <https://en.wikipedia.org/wiki/Background_subtraction>.