Exploiting Koinophilia-Phenomena to Classify Geographic Origin

Austin Lubetkin
Florida Polytechnic
University Lakeland, FL,
USA
bocaaust@gmail.com

Abstract—This paper primarily explores a new method that pairs evolutionary psychology and computer vision to classify the geographic origins of an unknown subject. This work first outlines existing research into koinophilia, an evolutionary psychology phenomenon describing phenotypical attraction to averages. An associated hypothesis to this paper is the geographic isolation of this phenomena. This research explores the potential of a simulation where an experimental group of known geography interacts with a selection of averaged faces, their relative koinophilia phenomena can be used to infer a geographic classification of the unknown group.

Keywords—computer vision, koinophilia, geotagging, evolutionary psychology

I. INTRODUCTION

Multiple experiments have averaged facial features together and found through experimentation averaged features were more attractive than individual faces. This phenomenon is known as koinophilia [3]. This phenomenon of people being attracted to the average is referred to in evolutionary psychology as koinophilia. This goes back to when humans were in tribal societies and people developed their sense of attraction based on preserving genetic bloodlines. This preference for average phenotypes was a strategy for genetic stability. By being attracted to averages humans are subconsciously programmed to preserve their genes [1]. This phenomenon is relative to geographic locations where the same experiment will find different average attractive features in different geographic locales. This is underscored by studies which have found different standards of attractiveness in different geographic areas [2].

Previous research into experiments in koinophilia phenomena showed people are subconsciously more attracted to the average of the faces in their subconscious relative to geographic history. “If individuals of a koinophilic species occupied different geographic areas with different distributions of phenotypes, the average phenotype in each area would be different, and individuals from other areas would be avoided as mates, leading to reproductive isolation” [3]. By exploiting this phenomenon we will demonstrate that one can understand geographic origins of people from a crowd through a crowdsourced revision of the koinophilia experiment where one side of the experiment has known geography. By crowdsourcing facial averages from users of known geography on one side of the experiment, you could potentially determine geographic origins of other sides of faces by performing random forest facial averages judged on relative and competitive attractiveness. By judging selective groups of averages features you could potentially tag the geographic origins of the faces in the crowd or if the crowd’s geographic origins are known you could potentially tag the geographic origins of the judging user.

A key advantage of the classification of geographic origin is it is a feature that is chronically locked. The majority of approaches for behaviorally tracking and classification of video surveillance are time-dependent in nature and rely on time relevant feature data [6][7][8]. The perception of average features is relevant only to perceptions developed at a specific point in time and is, therefore, time-invariant and reliant on the use of known sources of chronological data. Therefore, the approach outlined in this paper can process similarity transforms at a higher computational efficiency when creating examples since temporal warping is not required for understanding relevant features of the data since temporal-specific features are used in the process.

In the following sections, I first break down potential applications for this research. Following that, I take an in-depth look at the process designed for the simulation to geotag an unknown selection of participants and I additionally break down the methodology that can algorithmically average faces to analyze koinophilic phenomena.

II. APPLICATIONS

There are several potential applications for this method relevant to industries including law enforcement and advertising.

A. Applications to Law Enforcement

This project has applications to law enforcement in the following scenarios. The simulation could be used to geotag members of a crowd of people. In addition, this process could
be used to interpret intelligence like the location of a photo including multiple people. If one can look at an event with a large number of people like a concert, this process could additionally be used to locate members of a known geographic origin within a crowd.

By starting with large groupings of people averaged together and then reducing those groupings downward it's possible to find and search for individuals in large crowds of people and process a Big Data resource of surveillance. If someone took the next door neighbor for example of someone you were looking for and had them run this simulation, the way they chose clusters of attractive faces would be more likely to include members of a similar geographic origin. By processing large groups and reducing the overlap to smaller groups made into composite faces, it is possible to locate someone within a crowd of people by exploiting koinophilia phenomena.

This scenario is most useful by abstracting large scale facial recognition, a process that is highly computationally intensive when scaled to N people to a much simpler process that is less computationally intensive since averages can be formed at a constant rate of computation time with varying N people.

The majority of video surveillance technologies available today are passive in nature. They can be used to review footage after an incident has happened. The approach of this technology can be used to actively search for a person amongst a big data source of video surveillance and provide an active approach to surveillance that can be used to respond to queries in real time. By providing an inexpensive solution towards surveillance processing it becomes possible to better utilize the resources of inexpensive surveillance cameras versus the typical barriers of expensive manpower and computation required to provide active video surveillance [6].

III. METHODOLOGY

The following is a breakdown of the methodology of exploiting koinophilia phenomena. The process starts with a phase of data gathering faces are gathered to build the simulation as discussed in section A. This is followed by an experimental phase where participants judge the relative attractiveness of a mixture of averaged faces as discussed in section B. When the geographic origins are known of the participants in either A or B, the hypothesis of this paper is that the geographic origins of the corresponding part can be interpreted.

A. Crowdsourcing Faces

When designing the initial experiment I came across a feature built into Facebook that would allow me to download face data to create the simulation. Facebook, allows friend lists to be sorted by hometown. In addition, profile pictures are publicly available and can be easily downloaded. Examples of profile pictures pulled from Facebook can be seen in Figure 1. This means that to add a location to the simulation you can pull the public friend's list of a profile in a given city and crowsource the faces of people living in a geographic area. An important thing to note about extracting pictures from social media is there is a source of subjective error introduced to the experiment due to different standards of facial expressions associated with profile pictures in different parts of the world. For our experiment, the standard closed mouth smile was much

B. Applications to Advertising

There are many industries like sales and marketing which benefit from a locally tailored experience. People feel more comfortable buying from someone they trust and people tend to trust those who they can relate to geographically. This creates an opportunity to use koinophilia to exploit marketing.

If a company is looking to enter a specific market and has prepared to advertise including commercials and other sources which include human faces, they could use the methodology outlined in this paper to adapt the faces used in their campaign to best match the geographic origins of the target market. Furthermore, this creates a new layer to geographically targeted advertising, a practice that is commonplace today and could increase the effectiveness of said advertising.

This process could also be adapted to hiring for sales positions. A firm could use this procedure to filter their candidates for in-person sales experiences based on their relative impressionability in terms of their relation to the average in the intended market. Furthermore, with nationwide operations, this process could be used with an entire company to relocate members of the sales force to the markets they are able to relate to the most to improve overall sales efficiency. When someone looks familiar to you, this familiarity can be exploited to develop trust.

Another marketing application is for kiosks and unsupervised sales operations operating in a location with varied geographic traffic like airports or other travel hubs. This process could be used to automate and localize products to meet localized sales interest. For example, people from the western USA might like one version of a logo versus people from the Eastern USA on average and with an automated machine that can print said logos onto merchandise, it's possible to deliver targeted sales in these environments.

When paired with other sources of behavior processing from video surveillance, it becomes possible to deliver targeted advertisements that meet local interests or to understand local interests based on the average responses of groups of people classified by geographic origin. By placing advertising in a geographically diverse area, it becomes possible to understand markets of best fit by the geographic classifications that have the best overall responses to the targeted advertising.
more common in the group of pictures from Rome, Italy than in the Boca Raton, Florida group where an open-mouthed smile was more common. With increased quantities of profile pictures introduced into the procedure, this margin of error can be standardized.

Another scenario is to grab a series of photos from public locations in a known area. This could be possible by pulling frames from webcams of known geography that are publicly available. There are a variety of databases available for this type of content which can be used.

Another factor that needs to be accounted for in the faces chose for the procedure is gender. The input data needs to be classified for male and female faces and the experimental procedure should be performed with faces from the targets preferred sex.

In addition, a concern to address is the race of the participants. If someone grows up in a specific geographic area, their perception of average attraction expands beyond their own race and includes the proportional mixture of the race of those in their geographic origin. Therefore, faces should be gathered from multiple ethnicities in a proportion that accurately represents the geographic origin for the best outcome of the procedure.

Figure 1: A selection of Facebook profile pictures pulled from publicly available Facebook data

B. Performing the Simulation

The following is an experimental procedure designed to geographically classify an unknown party. After a data-set of faces have been collected from a selection of known geographies, they should be averaged together. I employed the OpenCV implementation for facial feature averaging developed by Malik [4]. Below are several faces created with this procedure and their experimental outcomes.

![Figure 1: A selection of Facebook profile pictures pulled from publicly available Facebook data](image1.png)

To better understand how these averaged faces are created, the following is a summary of the algorithm employed that was designed by Malik and based on research into koinophilia [4]. The first step of the algorithm is facial key point detection. "The script tries to find human faces in an image and extract 68 landmarks. These are points on the face such as the corners of the mouth, along the eyebrows, on the eyes, and so forth. We'll need these landmarks to map the different faces onto each other" In the example used for the implementation of this simulation I employed a trained model that specialized in this task. This model was a dlib model and detected 68 key points.

Once key points are detected, facial averaging can proceed. First, the images are standardized to a coordinate system with key features like eyes being set to specific coordinates of a homogenized and scaled image. This is done with a similarity transform. To create the similarity transform, the OpenCV methods for warpAffine are used. An interesting feature of this algorithm is that it considers a third point which would form an equilateral triangle between points to meet the requirements of the warpAffine method.

The next step involves using Delaunay triangulation to standardize the features of the face. Overall the faces have been aligned in the previous step, but this allows the alignment of individual features.

An expansion to the initial methodology designed occurs in this specific step. Through experimentation, I discovered that the results of the simulation could be influenced by the facial expressions of the averaged faces. I developed a classifier that is a convolutional neural network that can understand the facial expressions of a given face. This method returns metrics for 9 key emotions. These metrics can be used to train a convolutional neural network to make additional adjustments during the phase of Delaunay triangulation to warp facial features to a standardized expression across outcomes. This could be achieved by focusing on the average classified emotional metrics and making transformations that are positive and negative to adjust to meet the average expressions.

Finally, these aligned features can be averaged together. “To calculate the average image, we can simply add the pixel intensities of all warped images and divide by the number of images” [4].

Additional optional expansion to this simulation is to standardize the external facial features. Experiments conducted on the rate of matching facial features to composite faces found there was no change in matching adjustments to external features [5]. This means that the visual noise created from the process on the external features of the face like the hair and neck could be reduced by replacing these features with less distorted features from a single participant from the sample as seen in Figure 4.
Figure 2: An average of 3 Female Faces from Boca Raton, Florida which scored an average of 9.4 on a 10 point scale where 10 is the most attractive.

Figure 3: An average of 5 Female Faces from Rome, Italy which scored an average of 5.3 on a 10 point scale where 10 is the most attractive.

In the experiment, a group of people orders a series of photos from most to least attractive. These faces include varying mixtures of either pure locations or mixes between varying locales. In our experiment, I surveyed 20 individuals who have spent the majority of their lives in Florida and found that 95% of the participants placed as their top 3 most attractive faces those which were composed of a majority of faces geographically originating in Boca Raton, Florida including Figure 2 which appeared in 100% of the top 3 selections. Figure 3, on the other hand, appeared in 15% of the top 3 selections. This experiment showed a clear correlation between the source of the faces used in the composition and the relative attractiveness of the composites. This procedure could further be expanded with more locations and by weighting the percentage a face is composed of a geographic origin, it is possible to estimate the geographic origins of a participant.

It’s possible to perform the procedure using a bot-net of Tinder or another dating app profiles where bots set to target locations could be used to experimentally crowd-source the geography of an unknown party. An advantage of this approach is the ability to specify bots gender and the classifying participants also have entered their preferred gender into the application which allows the proper experimentation to occur without adjustments for sexual preference. The advantage to this approach is it’s possible to perform the procedure where the botnet operates in a known set of geographies and you can, therefore, interpret the geographic origins of an unknown set of composite faces based on their area of highest phenotypical attraction match.

Figure 4: This is an example of the composite face generated in Figure 2 processed through the additional procedures that were an expansion to the Malik method [4] wherein external features are replaced with those from a single face to create a more natural and less noisy composite face.

IV. CONCLUSION

The procedure outlined in this paper can exploit an unknown subject’s perception of averages to determine their geographic origin. Geographic origin is an actionable feature that has applications for law enforcement, advertising, and many other industries. The methodology used to test this phenomenon relied on a small sample group of two target geographies. Future research will determine the applications for scaling to a big data solution of hundreds of geographies with thousands of faces. The permutations of faces generated by this process will additionally increase factorially. This means that the time required to sort faces will increase factorially as well.
The procedure outlined above can clearly make a determination between two geographic origins of an unknown participant. This application is still useful situationally when sorting between several groups or when a hypothesis is being tested of potential geographic origins with prior insight used to reduce the number of potential locations.
REFERENCES


