

GeoFaceExplorer: Exploring the Geo-Dependence of Facial Attributes

Connor Greenwell
University of Kentucky
csgr222@uky.edu

Richard Souvenir
UNC Charlotte
souvenir@uncc.edu

Scott Spurlock
UNC Charlotte
sspurloc@uncc.edu

Nathan Jacobs
University of Kentucky
jacobs@cs.uky.edu

ABSTRACT

The images uploaded to social networking websites are a rich source of information about the appearance of people around the world. We present a system, GeoFaceExplorer, for collecting, processing, browsing, and analyzing this data. GeoFaceExplorer allows for the crowdsourced collection of human facial images, as well as automated and interactive visual analysis of the geo-dependence of facial appearance and visual attributes, such as ethnicity, gender, and whether or not a person is wearing glasses. As a case study, automated approaches are applied to detect common facial attributes in a large set of geo-tagged human faces, leading to several analysis results that illuminate the relationship between raw facial appearance, facial attributes, and geographic location. We show how the distribution of these attributes differs in ten major urban areas. Our analysis also shows a similar expected distribution of ethnicity within large urban areas in comparison to manually collected U.S. census data. In addition, by applying automated hierarchical clustering to facial attribute similarity, we find a large degree of overlap between discovered regional clusters and geographical and national boundaries.

Categories and Subject Descriptors

H.2.8 [Database Applications]: Image databases; H.2.8 [Database Applications]: Spatial databases and GIS

Keywords

faces, geolocation, images, facial attributes

1. INTRODUCTION

What do people look like in southern Sudan? Is it common to wear a sari to a cricket match in Mumbai? What types of hats do teenagers wear in Nigeria? Addressing these questions using traditional surveying techniques or census data

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions at Permissions@acm.org
SIGSPATIAL '14, November 04-07 2014, Dallas/Fort Worth, TX, USA
Copyright 2014 ACM 978-1-4503-3133-3/14/11 ...\$15.00
<http://dx.doi.org/10.1145/2676440.2676443>.



Figure 1: Human facial appearance differs for many reasons, including ethnicity, gender, and hair style. In this work, we present a system for exploring the relationship between such visual attributes and geographic location.

would be costly and not likely to provide the flexibility for fine-grained analysis or to answer new questions. Another avenue of analysis is enabled by the convergence of two phenomena. First, every day, a growing number of (geo-tagged) images are uploaded to social media sites. On one popular social media site¹, geo-tagged photos are uploaded at a rate of around 500 per minute, or 260 million per year. Second, the state-of-the-art algorithms in computer vision have reached a level of accuracy and robustness that allows detailed scene information (e.g., people, objects, background) to be automatically extracted from images. We present a system for visualizing and interacting with a large collection of crowdsourced images designed to support queries on the distribution of facial appearance and features from around the world.

We present a case study for using geo-tagged images as a source for information about people and their appearance across regions of the world. Our focus is on faces extracted from geo-tagged images (Figure 1), and builds upon the Ge-

¹<http://twitter.com>

oFaces dataset [4], an ever-growing repository of such facial image patches, by incorporating automatically extracted visual attributes such as gender, ethnicity, and facial hair. Our system provides methods of visualizing the distribution of attributes or relative distribution of value pairs within a given attribute and supports interactive analysis from the global down to city scales. In addition, our system uses these extracted visual attributes to cluster faces with relation to geo-location.

1.1 Related Work

The advent of inexpensive GPS-enabled cameras and persistent connectivity has led to a profusion of geo-tagged imagery that may be freely downloaded from social media sites. Such images have been used to extract a wide variety of geospatial information, including 3D scene models [1], local weather conditions [11], land-cover type [7], and architectural styles [2]. To our knowledge, only one previous work [4] has used automated approaches for exploring the relationship between facial appearance and geographic location. Our work significantly extends this work by performing deeper, attribute-based analysis which reduces the challenges of pose and lighting normalization.

Various approaches to navigating and exploring face imagery have been proposed. In terms of exploring facial features, Kumar et al. developed a search engine for faces where users can narrow down lists of faces by numerous facial attributes that were labeled off-line by an automatic facial feature detector [5]. However, unlike our system, this method does not consider the location where the photo was taken.

This work is most related to recent efforts in automatically extracting contextual cues from natural images. Doersch et al. examined the geo-dependence of architectural features in cities around the world [2]. In another application, scene-level features were used to aid in image geo-localization [3]. These methods overlook the very thing that humans photograph most; other humans. Our system aims to close this gap by providing a method for exploring how the changing appearance of people over space and time can lead to insights into trends in fashion, migration, and coiffure.

1.2 Contributions

The main contributions of this work are:

- examining the geo-spatial variations in human appearance using visual attributes extracted from a large database of geo-tagged faces;
- presenting interactive, map-based visualizations representing distribution of visual attributes; and
- automatically discovering clusters of facial similarity and comparing these regional clusters to geographical and national boundaries.

2. SYSTEM OVERVIEW

We introduce GeoFaceExplorer, a system for exploring large datasets of geo-tagged face images from around the world. The system is composed of three main modules:

- **Dataset construction:** New images can be added to the repository automatically, by scraping social media sites, or manually, by allowing users to provide geo-tagged images.
- **Facial Image processing:** New images are processed to detect faces and estimate various facial attributes.



Figure 2: For these images, the face detector automatically identified each face and provided estimates for the pose direction and detection confidence.

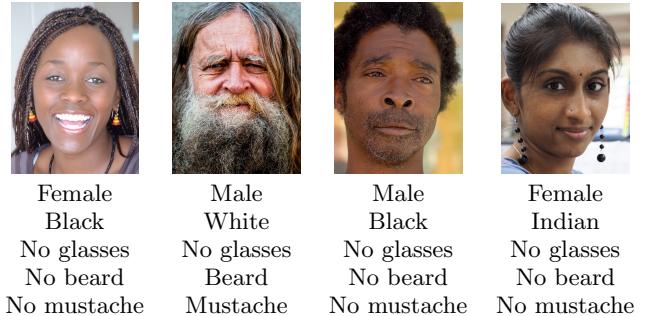


Figure 3: For each face in our dataset, the attribute detector automatically extracts values for multiple facial attributes.

- **Data visualization:** Users explore visualizations of various analysis methods applied to the extracted features.

The remainder of this section provides additional details about each aspect of GeoFaceExplorer.

2.1 Dataset construction

We seed the system with an initial set of images by following an approach similar to the method used to build the GeoFaces [4] dataset. Geo-tagged images with face-related tags (e.g., face, portrait, men, family, friends) are downloaded from Flickr². For each image, a commercial face detector³ is used to identify (nearly frontal) faces and fiducial points. From 3.14 million images, 2.65 million face patches are extracted. For each face patch, the detector reports the estimated pose direction and detection confidence, as shown in Figure 2. For our initial dataset, we filter candidate patches based on these values. Specifically, we retain images with an estimated pose of zero degrees (directly facing the camera), and we empirically determine that a detection confidence greater than 600 filtered most of the non-face false positive detections. This resulted in an initial dataset of 1.2 million faces.

Over time, the dataset can continue to grow by tapping additional social media sites and a wider variety of query terms. Further, using the web interface of GeoFaceExplorer, users can upload their own geo-tagged images directly. In either case, once high-confidence, geo-tagged frontal face images are obtained, the next step is to extract appearance information.

²<http://flickr.com>

³http://www.omron.com/r_d/coretech/vision/okao.html

2.2 Facial Image Processing

Our system automatically extracts features from every face image we collect. Many approaches have been proposed for analyzing human face images. For example, Kumar et al. developed a method for pairwise face verification by comparing sets of human-describable features and visually descriptive similes [6]. Another approach built generative models for opposing facial attributes (smiling-to-frowning, etc.) [8]. Xiong et al. recently introduced IntraFace, a tool for identifying human facial features [10]. Our system incorporates IntraFace for facial image processing.

IntraFace computes many facial attributes from an image path; the publicly accessible version of the software provides values for five attributes: beard, gender, glasses, race, mustache. Except for glasses (“Eye-glasses”, “Sun-glasses”, “No glasses”) and race (“Asian”, “Black”, “Indian”, “White”), the attributes are binary. Figure 3 shows example faces with the attribute values estimated by Intraface. Currently, we rely on the attributes automatically extracted from the images, but plan to allow these to be modified by users. For example, for the third image in Figure 3, the automated algorithm classified the image as “No beard” and “No mustache” even though stubble is visible.

The image processing pipeline (face detection, alignment, attribute detection) takes roughly 5 seconds per image, with most of the computation spent on attribute detection. Without any optimization, processing the initial collection of Flickr images would take roughly 2.5 CPU-months. GeoFaceExplorer has been optimized to process images in parallel to efficiently handle large numbers of user-supplied images.

2.3 Data Visualization

The GeoFaceExplorer interface allows users to visualize trends across the globe or within smaller regions. For example, a user can highlight a region of interest and inspect a number of statistics and visualizations calculated from the faces found in the region. Figure 5 shows three visualizations provided by GeoFaceExplorer.

Scatter Plot Figure 5a shows all of the locations where an image containing a face with a particular attribute value was collected.

Density Map Figure 5b uses kernel density estimation to calculate and map the density of a particular facial attribute value in a given region of interest. Given data (x_1, x_2, \dots, x_n) and a kernel function K_h , an estimate of the density of the data at x is given by:

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i). \quad (1)$$

In this work, we use Gaussian kernel functions (isotropic in latitude/longitude) with bandwidths that vary depending on the desired scale of analysis.

Relative Density Histograms Figure 5c shows the relative density of pairs of attribute values. For fixed-sized bins, the intensity represents the relative frequency, λ , of a pair of labels:

$$\lambda(n_1, n_2) = \frac{(n_1 + p) - (n_2 + p)}{(n_1 + p) + (n_2 + p)} \quad (2)$$

where n_1 and n_2 are the number of facial images in a region of interest with a particular attribute value

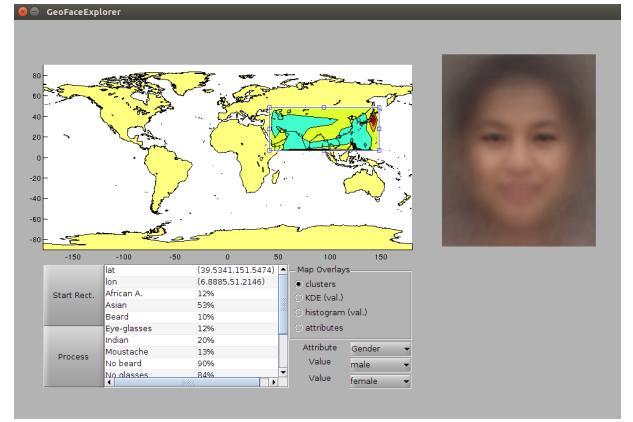


Figure 4: A GeoFacesExplorer window showing the relative density contour map of “male” (red) versus “female” (blue) in a selected region of Asia.

and p is a pseudo-count used to reduce the influence of outliers (for all results in this paper $p=10$).

2.4 GeoFaceExplorer Summary

GeoFaceExplorer is an interactive system for exploring the geo-tagged faces of the world. Figure 4 show an example of a user highlighting a region of interest and inspecting statistics, map-based visualizations, and representative images calculated from the faces found in the region. In the next section, we describe case studies of the types of analysis possible with GeoFaceExplorer.

3. DATA EXPLORATION

The GeoFaceExplorer system supports both interactive exploration and automated analysis. In the remainder of this section, we show a variety of interesting insights obtained by using this system in both ways. We use GeoFaceExplorer to analyze images at various spatial scales, from global to local. The local analysis focuses on regions surrounding major urban areas, as these are the locations with the highest image density.

3.1 Interactive Analysis

The most interesting insights come from the relative distribution of paired attribute values at both global and regional scales. We computed the relative distribution of each pair of attribute values (e.g., “Male” versus “Female”, “Beard” versus “No beard”) and plotted the relative density histogram. At the global scale, the bin size is 5 degrees latitude by 5 degrees longitude (an area roughly 340,000 sq. kilometers, or about the size of Germany). For the set of images in each bin, the color of the tile represents the relative frequency of the attribute pair at that location. Similarly, we zoomed in to regions of high facial image density to observe finer-grained patterns. At this scale, the bin size was about 6.5 kilometers square. Figure 6 shows the relative density histograms for several attribute pairs for London, Los Angeles, New York, and the whole world.

Many of the global attribute distributions are unsurprising and follow expected geographic distributions. For example, comparing the relative frequency of the “Asian” and “White” attribute values for race reveals distinct, and oppo-

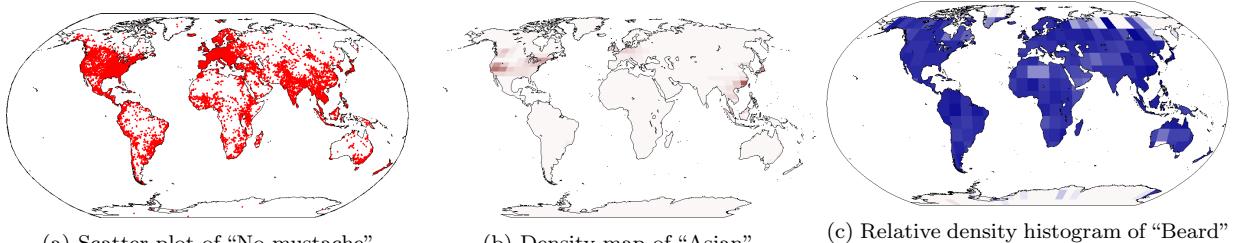


Figure 5: Sample visualizations provided by GeoFaceExplorer.

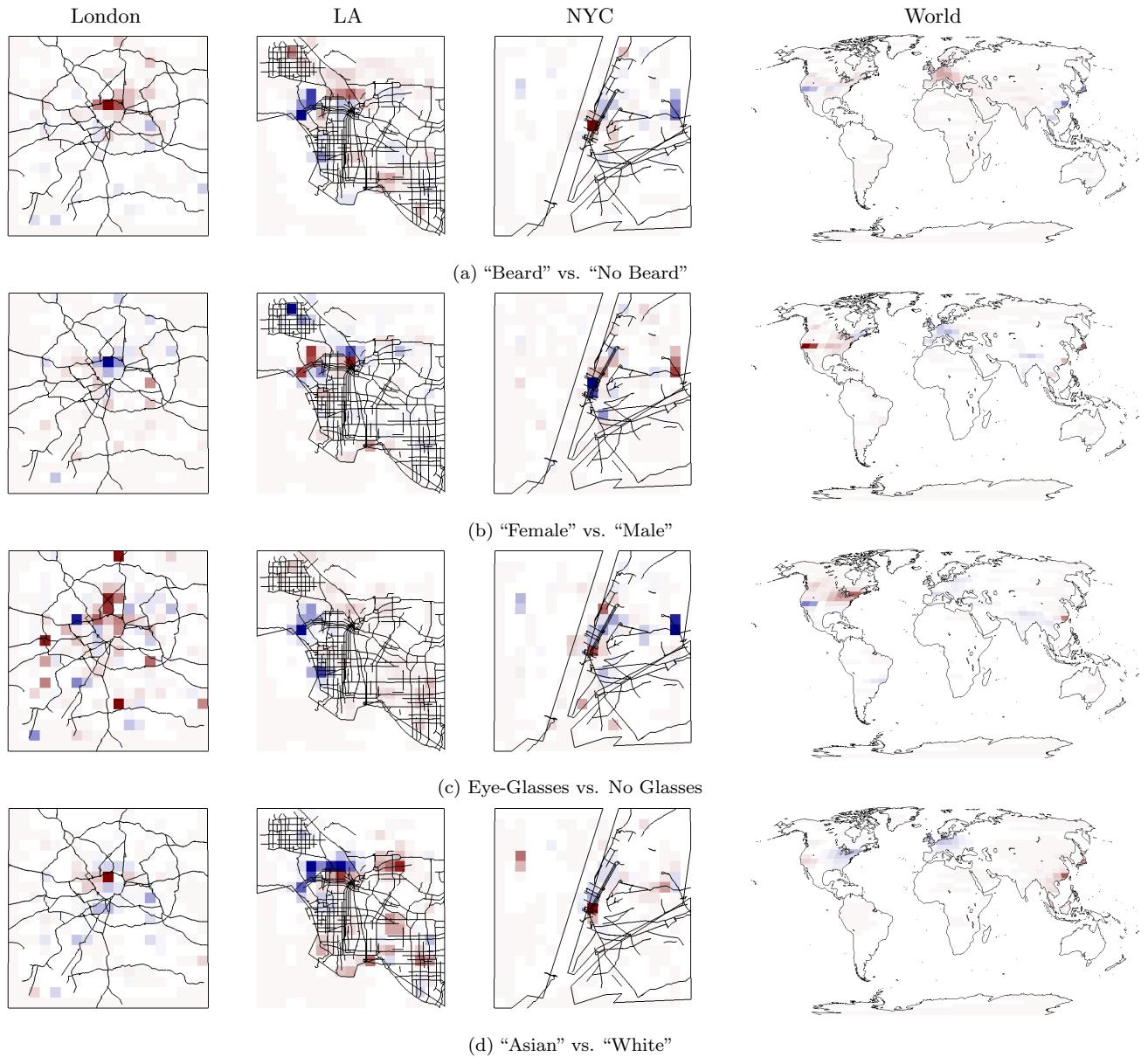


Figure 6: Maps showing the relative densities of paired attributes in selected cities and globally. In each graph, red represent higher concentrations of the first attribute value, and blue represent higher concentrations of the second.

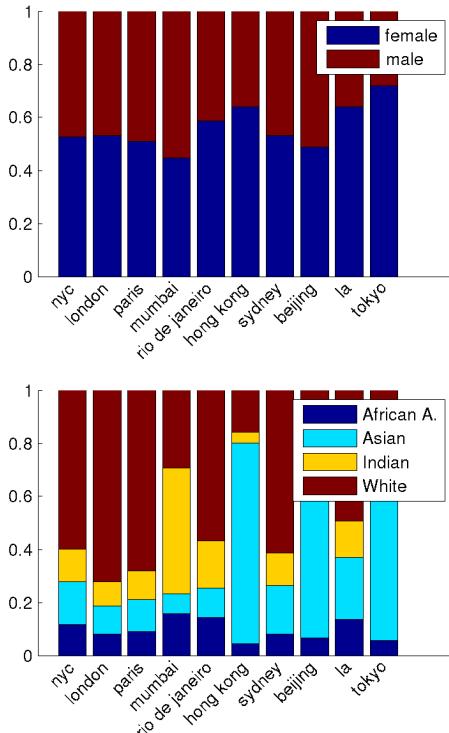


Figure 7: Stacked bar charts representing the distribution of gender (top) and race (bottom) of images collected from 10 major world cities.

site, modes in Southeast Asia and Europe. Similarly, the red “hot spots” in London, LA, and NYC correspond to regions of these cities with large Asian populations (e.g., Chinatown in NYC). Other patterns are perhaps more unexpected. Our dataset contains a higher proportion of “Female” images in eastern Asia and western United States and a higher proportion of “Male” images in the Middle East. Also, there is a high proportion of “Beard” faces (compared to “No Beard”) in the heart of downtown London. On first glance, it is not always clear if these types of observations are due to particular biases in the dataset or cultural norms. However, GeoFaceExplorer allows the user to look at individual and composite summary images for further investigation. For example, upon exploring the over-representation of beards in London, one finds other correlated attributes. London “Beard” images are correlated with “Male” faces, and “Asian” faces. A similar pattern emerges from the maps of New York City, around the Chinatown area in Manhattan.

Beyond the map-based visualizations, GeoFaceExplorer provides summary statistic for any selected local regions. For the regions surrounding 10 major cities (New York, London, Paris, Mumbai, Rio de Janeiro, Hong Kong, Sydney, Beijing, Los Angeles, Tokyo), the stacked bar charts in Figure 7 show the ratios of certain attribute values in each city. Despite inherent biases in the data due to source (e.g., geotagged images uploaded to social media sites or GeoFaceExplorer), we observe that the “Race” and “Gender” attributes distribution roughly correspond to the 2010 U.S. Census demographic data for both Los Angeles and New York.

3.2 Automated Analysis

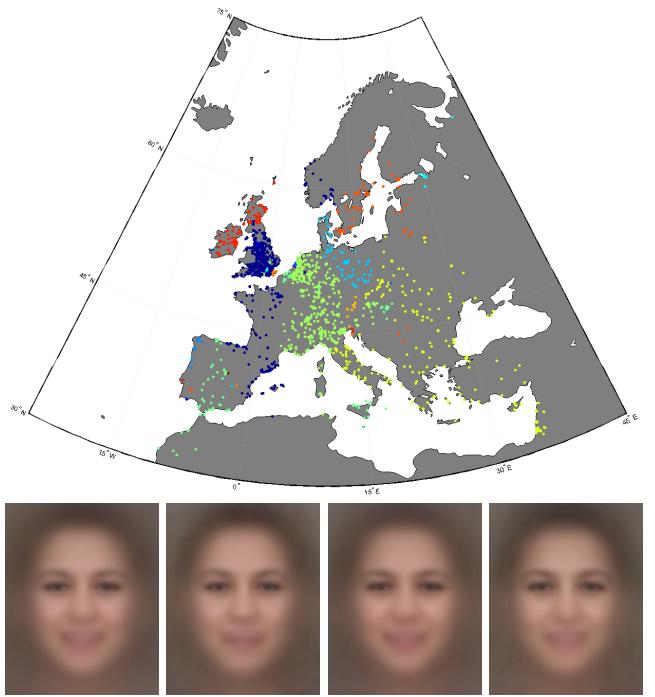


Figure 8: Continent level clustering and average faces. Left to right: England/western France, Scandinavia, Germany/Northern Italy, Italy/Greece/Turkey.

To facilitate discovery of underlying patterns in the collected faces, GeoFaceExplorer also supports automated analysis based on data clustering. Of particular interest is the relationship between global location and facial attributes, which can be investigated by grouping faces that are nearby spatially as well as similar in terms of attributes. Well-known unsupervised clustering techniques can be applied to this problem. We employ the normalized cuts algorithm [9] on a graph-based representation of the dataset, where nodes are facial images and edges encode attribute similarity. Only spatially proximate face nodes are linked by edges. The clustering process identifies groups of faces from similar regions that share similar attributes.

Interestingly, although the clustering process is unsupervised, the discovered clusters tend to align with geographically meaningful regions. Figure 9, shows the results of an experiment with 15 clusters. Regions such as Eastern United States, Central America, and Europe are evident in the clustering output. Average face images from several groups illustrate the correspondence between clusters and geographic regions.

The same alignment of discovered clusters and established regions can be seen at the continent level. Figure 8 shows the results for a similar experiment limited to an area roughly encompassing Europe. Faces are grouped together in a manner that combines both the geographic location and similarity of facial images. We observed similar partitioning in other sub-regions.

4. CONCLUSION

We presented GeoFaceExplorer, a system that facilitates

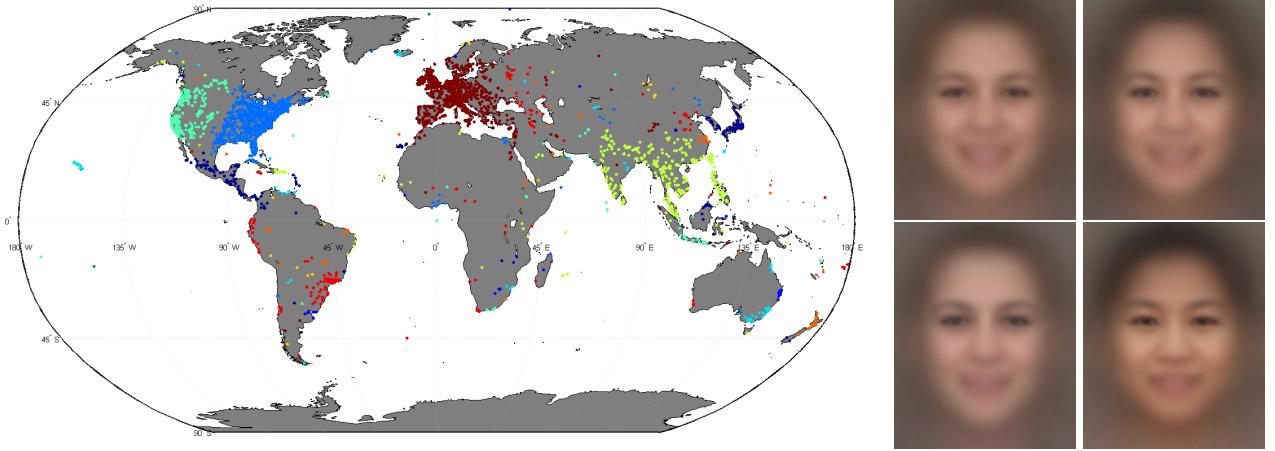


Figure 9: World level clusters and average faces for each of the four largest clusters. Clockwise from top-left: Eastern US, Western US, Southeast Asia, and Europe.

the interactive exploration of facial appearance across the world from a collection of geo-tagged crowdsourced images. The system seamlessly blends both automated and interactive visual analysis to support a variety of data exploration tasks.

GeoFaceExplorer can be used with two types of crowdsourcing: direct, where users of the tool provide images, and indirect, where the images of Internet users are collected from social media sites. This second formulation allows us to bootstrap the system with millions of images to provide the coverage necessary to observe large-scale patterns. Nonetheless, the current database is sparse in large portions of the world, including Africa, South America, and Central Asia, while areas such as America, Japan, and Western Europe contain a large number of images. Since the popularity of image sharing websites varies significantly from country to country, we expect that the spatial coverage of our system will improve dramatically as we add additional social media sites to our collection module.

Several improvements to GeoFaceExplorer are planned for the future. We will extend the interface, with the goal of supporting new types of data exploration and more expressive queries. A significant focus will be on enabling queries that highlight changes in the distribution of facial appearance over time. Also planned are improvements to the facial image processing module, initially focused on the current facial attribute extraction methods, as well as on adding additional attributes, such as headwear, eye color, and facial expression.

Acknowledgements

We gratefully acknowledge Mohammad T. Islam for providing a foundational dataset for our initial experimentation. This work was partially supported by the NSF (CNS-1156822) and DARPA (D11AP00255). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the sponsor.

5. REFERENCES

- [1] S. Agarwal, Y. Furukawa, N. Snavely, I. Simon, B. Curless, S. M. Seitz, and R. Szeliski. Building

- Rome in a Day. *Communications of the ACM*, 54(10):105–112, 2011.
- [2] C. Doersch, S. Singh, A. Gupta, J. Sivic, and A. A. Efros. What Makes Paris Look like Paris? *ACM Transactions on Graphics (SIGGRAPH)*, 31(4), 2012.
- [3] J. Hays and A. A. Efros. im2gps: Estimating Geographic Information from a Single Image. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008.
- [4] M. T. Islam, S. Workman, H. Wu, R. Souvenir, and N. Jacobs. Exploring the Geo-Dependence of Human Face Appearance. In *IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2014.
- [5] N. Kumar, P. N. Belhumeur, and S. K. Nayar. FaceTracer: A Search Engine for Large Collections of Images with Faces. In *European Conference on Computer Vision (ECCV)*, 2008.
- [6] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar. Attribute and Simile Classifiers for Face Verification. In *IEEE International Conference on Computer Vision (ICCV)*, 2009.
- [7] D. Leung and S. Newsam. Proximate Sensing: Inferring What-is-Where from Georeferenced Photo Collections. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.
- [8] D. Parikh and K. Grauman. Relative Attributes. In *IEEE International Conference on Computer Vision (ICCV)*, 2011.
- [9] J. Shi and J. Malik. Normalized Cuts and Image Segmentation. *IEEE Transactions Pattern Analysis and Machine Intelligence (PAMI)*, 22(8):888–905, 2000.
- [10] X. Xiong and F. De la Torre. Supervised Descent Method and its Applications to Face Alignment. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013.
- [11] H. Zhang, M. Korayem, D. J. Crandall, and G. LeBuhn. Mining Photo-Sharing Websites to Study Ecological Phenomena. In *International Conference on World Wide Web (WWW)*, pages 749–758. ACM, 2012.