Democratizing the visualization of 500 million webcam images

Joseph D. O'Sullivan, Abby Stylianou, Austin Abrams and Robert Pless
Department of Computer Science
Washington University
Saint Louis, Missouri, 63130
Email: osullj@wustl.edu, astylianou@wustl.edu, abramsa@cse.wustl.edu, pless@wustl.edu

***Abstract*—Five years ago we reported at AIPR on a nascent project to archive images from every webcam in the world and to develop algorithms to geo-locate, calibrate, and annotate this data. This archive of many outdoor scenes (AMOS) has now grown to include 28000 live outdoor cameras and over 630 million images. This is actively being used in projects ranging from large scale environmental monitoring to characterizing how built environment changes (such as adding bike lanes in DC) affects physical activity patterns over time.**

But the biggest value in a very long term, widely distributed image data-set is the rich set of before data that can be analyzed to evaluate changes from unexpected or sudden events. To facilitate the analysis of these natural experiments, we build and share a collection of web-tools that support large scale, data driven exploration. In this work we discuss and motivate a visualization tool that uses PCA to find the subspace that characterizes the variations in this scene, This anomaly detection captures both imaging failures such as lens flare and also unusual situations such as street fairs, and we give initial algorithm to clusters anomalies so that they can be quickly evaluated for whether they are of interest.

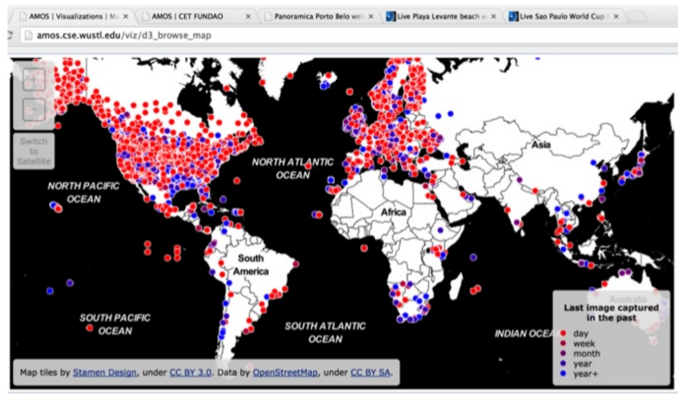

Fig. 1. As of 2014, the AMOS database comprises over 630 million images, captured from public webcams around the world. This paper discusses aspects of visualizing this dataset in terms of data availability, patterns of change, and anomalies within scene.

I. Introduction

The amount of imagery available online to the public is astounding, comprised of satellite feeds, Google Street View, Flickr and Facebook uploads, and webcam imagery. Equally amazing is the efforts that companies expend to index this imagery so that anyone can find the imagery they need or want – pictures to suggest what a particular highway exit will look like or to remember what, exactly happened at a party the night before. This indexing is usually done based on geographic coordinates of the image data, or semantic/text tags that are included with the image data.

Image data captured over time by a webcam is not well served by these indexing methods, because all images from a webcam come from the same geographic coordinates, and because webcams constantly share images and rarely have any sort of annotation included with those images. This paper aims to develop tools to improve the visualization and indexing of these sets of images that are taken over the course of many years from a single viewpoint. The challenges to visualizations of outdoor data over long time periods include the fact that large variations are caused by variations that may not be of interest to a particular application. These nuisance variables include

- lighting and weather variations,
- image artifacts such lens flare and specular reflections, and
- occluding objects such as cars that, in some cases, take up a substantial fraction of the image.

While there are attempts to define a formal information theory of outdoor images that explicitly captures the effects of these nuisance variables [9], [11], [12]; the approach in this paper is to create specific visualizations can help users quickly look through large amounts of data and find the types of variations or images that are most interesting to them. These visualizations are built into the interface for a very large archive of webcam imagery called "The Archive of Many Outdoor Scenes" (AMOS) [7], [6], [8]. That dataset includes imagery from about 28000 webcams, captured one picture per half hour, for the last 8 years. The specific contributions of this paper include:

- a description of AMOS and documentation of the types of image variations within it,
- an exposition of visualization tools appropriate for very long image sets from static images and screen-shots of its implementation as a live web service in AMOS, and
- examples of how these visualization tools illustrate different types of long term changes.

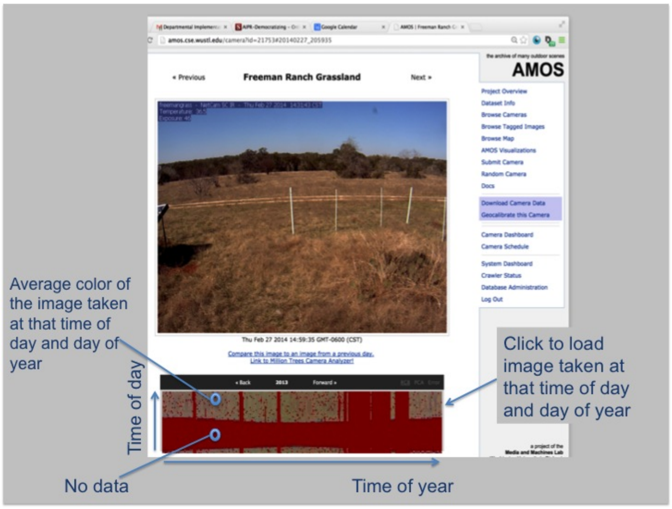

Fig. 2. The basic format of our camera visualization page shows an image from the camera and an annual summary image. This summary image is organized by time of day and time of year, and each pixel in the summary images is the average color of the entire image captured on that day at that time. Dark red indicates that the camera did not capture an image at that time, in this case the camera only operates in the daytime and had some week-long and one several month long gap in service.

All the visualization tools are available to the public live through the website: http://amos.cse.wustl.edu.

II. Background and Related Work

In the last 5 years there has been an incredible variety of very large image data sets that have been collected and shared. Some of the most interesting datasets attempt to be comprehensive, for example, the ImageNet project giving example images for all noun categories in the English Language [5], a data set called 80 million tiny images aims to be the definitive resource for scene classification examples [13]. Imaging resources such as Flickr allows users to geotag images creating fascinating maps of where and when people take pictures around the world [4], and long running crowd source projects that explicitly seek to acquire pictures that are well sampled across the earth [1].

In this paper we focus on the problem domain of images that are captured from a camera that remains in the same location and views the same scene. In this simplified context, the image variations are very limited and often well represented by a low-dimensional subspace [7]. Work that focuses on this problem domain often seeks representations of common scene variations in order to detect anomalies [3], [10] or to find patterns of behaviors that help to label image regions by their function attributes [14], or to label times of day or times of week that correspond to behavioral changes [15], [2]

To our knowledge, our paper is the first to explicitly seek to create web accessible visualization tools capturing the variation in long term image data sets. Making these tools publicly available and not reliant on expertise in Matlab or R or another statistical programming packing is a requirement to make these broadly accessible.

III. Current Visualizations

Our AMOS dataset currently contains the web addresses of over 32000 webcams, about two-thirds of which are currently publicly sharing images. For each camera, we capture one image each half hour, and our current database comprises over 630 million images. One primary question for data sets collected from disparate locations and systems is simply data availability. This leads to our first visualization, which offers the ability to navigate an image data set based on an annual summary image. This is shown in Figure 2 which is a screen capture of our interactive system.

This interface has three parts. On the right of the screen is a navigation bar to visit other parts of our site. The top shows an image from the current camera. The bottom of the screen (in red and gray) is an annual summary image. Each location in this summary image corresponds to a time of day and a time of year. The pixel at that location is colored by the average (r,g,b) color of the entire image that was captured at that time of day and day of year, unless no image was captured in which case that pixel is colored dark red.

This is a ”clickable” interface, a user can select and see the image from a time of day and day of the year. The camera shown in this figure views a farm, and in this (r,g,b) summary image it is possible to see that the summertime is slightly greener than winter. Such effects are often much smaller than expected because there is substantial natural variation due to weather, snow, lighting, etc., and because many cameras implement some sort of white balancing that encourages the average color of the image to be gray.

Visualizing Variations: Because the raw color of images is often uninformative, we consider another simple summary of each image. The PCA decomposition provides an approach to take a large collection of images: $I_1, I_2, I_3, \ldots I_n$, and to represent each image as a linear combination of a mean image and basis images, so that:

$$I_i \approx I_\mu + c_{i1}B_1 + c_{i2}B_2 + c_{i3}B_3,$$

where I_μ is the mean image, B_1, B_2, B_3 are basis images, and c_{i1}, c_{i2}, c_{i3} are the coefficients that describe what linear combination of B_1, B_2, B_3 are needed to reconstruct I_i. PCA solves for the basis images B_i that capture the maximum variation within the dataset, and the coefficients for each image represent how they vary within that basis. Therefore, we use the coefficients c_{i1}, c_{i2}, c_{i3} as the $(r, g, b) - color$ to create a false color image that summaries how a scene changes over time.

Figure 3 gives on example of this summary image, showing a years worth of images captured from a webcam on the Notre Dame campus. In the annual summary image, the colors no longer have a direct relationship to the colors that appear in the image, but pixels on the summary image that are the same color are similar images (or, more specifically, their approximate PCA reconstruction is similar) because they have similar coefficients. In the particular case shown, the images at night have a blue-ish green false color and are very consistent (because the night-time appearance in this scene is dominated by a few lights that are constantly on). The daytime images often vary over the course of the day because the sun changes position, and this is reflected in the false-color summary image

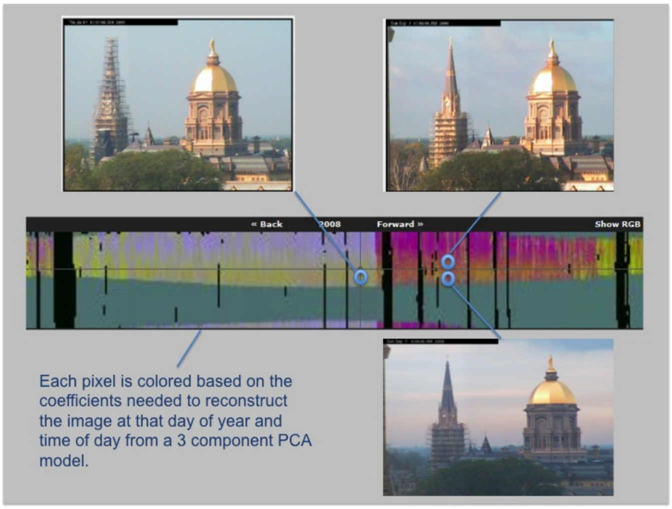

Fig. 3. One alternative summarization approach computes a 3-component PCA decomposition of all images in the year and then represents each image by the 3 coefficients used to best reconstruct it. Using those 3 coefficients as the (r,g,b) color in the summary image create a false-color image that highlights important change in the scene, such as the slight camera viewpoint shift that happens in late July, and variations in scene appearance as a function of cloudiness and different sun positions.

Fig. 4. For large image sets where the camera does not move, the space of images that is view is often low dimensional and a low dimensional PCA reconstruction often reconstructs images quite well. The residual image captures regions of the scene that are not well reconstructed; here those include the flowers on the near tree (left) and most of the flowering tree on the right.

as a color change from greenish to purple-ish in the first half of the year, and from orange to fuchsia in the second half of the year. Near the middle of the year, the camera slightly shifted viewpoints; because PCA is a linear basis that does not code for motions, reconstructing the images before and after the shift require significantly different coefficients and the shift becomes easy to see in this summary image.

Reconstruction Error: The PCA coefficients capture how one image varies with respect to the most common variations in the data set, but image may also vary in other ways. The image variation not captured by the PCA reconstruction is captured in the residual; the difference between the image and its reconstruction. This is computed as follows:

$$R_i = I_i - I_\mu + c_{i1}B_1 + c_{i2}B_2 + c_{i3}B_3,$$

An example of an original image and its approximate PCA reconstruction is shown in Figure 4. The residual image shows the magnitude of the reconstruction error as a colormap from blue (low error) to red (greatest error). This highlights that in this scene, the flowering trees on the right and the left are unusual with respect to the PCA basis, which makes sense because they flower only for a short amount of time.

Fig. 5. Another alternative summarization highlights the reconstruction error of each image by displaying those errors for each day and time of day. Images with the highest reconstruction error are often unusual images like those with strong lens glare (left), or times when a camera is briefly moved or zoomed in compared to its usual viewpoint (right).

To create a tool that allows anyone to understand the unusual images from a camera throughout a year we use this residual computation to create another annual summary image. In this case, we create a pixel for each image whose magnitude corresponds to the sum of the squared residual errors. This creates an annual summary image whose brightest points correspond to the images that were least well reconstructed by the PCA bases. Figure 5 shows an example of this image again on the Notre Dame webcam, with two example images that had large reconstruction error. One of these images was a case where the sun was in the field of view and a large lens flare dominates the image appearance. The other image comes from a day when the camera was zoomed in dramatically and the image was not similar to the rest of the year.

IV. Characterizing Anomalies

The images that are highlighted in the visualization shown in Figure 5 are typical of natural outdoor images that have the largest residuals, because lens-flare and large camera motion create changes all over an image. While those effects are sometimes of interest, it is often other types of changes that are more relevant or interesting to an application. Most commonly, systems first flag anomalies and then show all anomalies to a user to find those that are interesting. For large datasets, however, there may be many anomalies so it is useful to consider how to provide additional support for finding those that are of interest.

The approach that we explore here is to support the exploration of anomalous images by clustering features of the residual images. The goal is to sort those images that are not well reconstructed by the PCA basis into groups that have different explanations of why they are not well reconstructed. Then a user can focus on the clusters that are of interest to them. Our algorithm for this process is the following:

1) Compute the PCA basis and reconstruction for all images for a camera from a year
2) Compute the magnitude of the residual (which is the sum of squared error of the PCA reconstruction) for each image
3) Select all images whose reconstruction error is in the highest 3%.
4) Use k-means clustering with 5 clusters to categorize these anomalies into clusters.

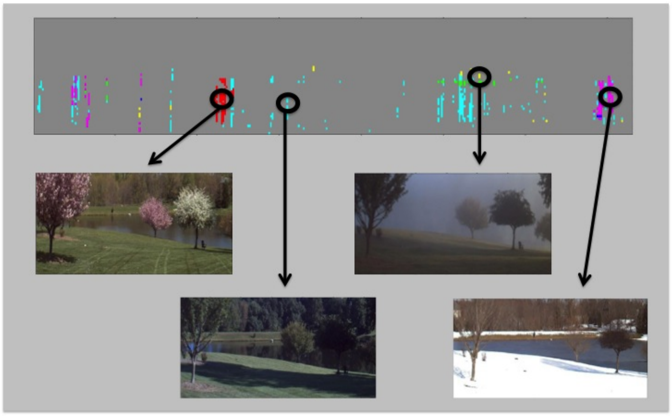

Fig. 6. Results of the anomaly clustering algorithm for the golf course camera shown in Figure 4. Pixels labeled with the same color represent images that were clustered together because they had residual errors in the same places, and representative pictures from some of the clusters are shown.

5) Create a visualization showing these clusters in the annual summary image.

We explored several features of the residual image, but found the most interesting and relevant clusters when we dramatically sub-sampled the squared residual image to $24 \times 32 pixels$, unwrapped that image into a 24×32 element vector and directly run k-means clustering on those vectors. These creates a clustering that captures *where* in the image there was an anomaly, but because the residual images are so sub-sampled, it is a coarse estimate of where the anomaly occurs so very small changes (like a specular reflection moving slightly over a curved metal surface) are clustered together.

After running that algorithm the 3% of the images with the worst reconstruction error are each assigned to a cluster. We create a false color summary image by placing a pixel at the time of day and time of year each image was captured. We color this pixel in the summary image based on its cluster.

We offer two examples of this clustering. The first example, shown in Figure 6 is from a webcam that looks at a golf course. We can impute meanings to these clusters by viewing example images and by exploring where (in time of year and time of day) those clusters appear. For this first example, we see:

- the pink cluster occurs only for a few days in April and last during the entire day — this corresponds to the few days when the trees in the scene are flowering,
- the light blue cluster is scattered throughout the day and on many different days. Example images show harsh shadows across the scene, which are difficult for a low-dimensional PCA model to capture because they move,
- the yellow cluster occurs only in the mornings because they correspond to foggy days and fog only appears in this scene in the morning, and
- the purple/pink cluster only appears in the winter and corresponds to snow on the ground.

This example is informative because three of these clusters (foggy days, snow on the ground, and harsh shadows) are not likely to be the type of anomalies one is most interested in. But looking through exemplar images of each cluster allows one to ignore whole clusters at a time.

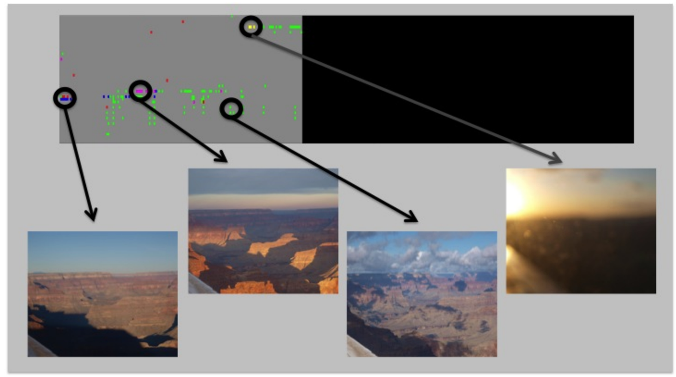

Fig. 7. Results of the anomaly clustering algorithm for a camera from the Grand Canyon. Pixels labeled with the same color represent images that were clustered together because they had residual errors in the same places, and representative pictures from some of the clusters are shown.

Another example scene is from the National Park Service webcam looking at the Grand Canyon. We ran this on images from the first half of 2014. The color coded visualizations of the clusters in this case shows:

- a pink cluster corresponding to shadow patterns in the very early morning,
- a dark blue cluster corresponding to shadow patterns in the late morning,
- a yellow cluster corresponding to the early morning condition when the sun is in the field of view, and
- a green cluster corresponding to different patterns of shadows from partly cloudy days.

V. Discussion and Conclusions

All visualizations of large scale data present challenges. We believe that creating systems and tools that allow web-based interaction for the public offers exciting potential for two reasons. First, from the perspective of allowing people to explore data-sets, creating web-visualizations makes this accessible to users that don't have the expertise to be able to use systems based on Matlab or R. Second, from the perspective of creating new data sets, building a system based on logging public webcam feeds, allows anyone to add their own webcam to our system and therefore benefit from having these visualizations for themselves or sharing them. Collectively, this allows a more diverse user based with a broader set of interests to explore large scale data visualization. We hope that sharing these tools broadly will encourage a very diverse set of users to explore data visualization in novels ways.

Acknowledgment

The authors would like to thank Nathan Jacobs for long term support and insightful comments, and support from: NSF DEB-1053554, NSF EF-1065734, NSF IIS-1111398, and NSF IIA-1355406. Any opinions, findings, and conclusions or recommendations expressed in this material are those of

the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] www.confluence.org.

[2] A. Abrams, J. Tucek, J. Little, N. Jacobs, and R. Pless. Lost: Longterm observation of scenes (with tracks). In *Applications of Computer Vision (WACV), 2012 IEEE Workshop on*, pages 297–304. IEEE, 2012.

[3] M. D. Breitenstein, H. Grabner, and L. Van Gool. Hunting nessie-real-time abnormality detection from webcams. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pages 1243–1250. IEEE, 2009.

[4] D. J. Crandall, L. Backstrom, D. Huttenlocher, and J. Kleinberg. Mapping the world's photos. In *Proceedings of the 18th international conference on World wide web*, pages 761–770. ACM, 2009.

[5] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 248–255. IEEE, 2009.

[6] N. Jacobs, W. Burgin, N. Fridrich, A. Abrams, K. Miskell, B. H. Braswell, A. D. Richardson, and R. Pless. The global network of outdoor webcams: Properties and applications. In *ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL GIS)*, pages 111–120, Nov. 2009.

[7] N. Jacobs, N. Roman, and R. Pless. Consistent temporal variations in many outdoor scenes. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–6, June 2007.

[8] N. Jacobs, R. Souvenir, and R. Pless. The global webcam imaging network. In *Applied Imagery Pattern Recognition Workshop (AIPR)*, pages 1–8, 2009.

[9] A. Ravichandran and S. Soatto. Long-range spatio-temporal modeling of video with application to fire detection. In *Computer Vision–ECCV 2012*, pages 329–342. Springer, 2012.

[10] E. Ricci, G. Zen, N. Sebe, and S. Messelodi. A prototype learning framework using emd: Application to complex scenes analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(3):513–526, 2013.

[11] S. Soatto. Actionable information in vision. In *Machine learning for computer vision*, pages 17–48. Springer, 2013.

[12] S. Soatto. Visual scene representations: Sufficiency, minimality, invariance and approximations. *CoRR*, abs/1411.7676, 2014.

[13] A. Torralba, R. Fergus, and W. T. Freeman. 80 million tiny images: A large data set for nonparametric object and scene recognition. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(11):1958–1970, 2008.

[14] M. W. Turek, A. Hoogs, and R. Collins. Unsupervised learning of functional categories in video scenes. In *Computer Vision–ECCV 2010*, pages 664–677. Springer, 2010.

[15] G. Zen, J. Krumm, N. Sebe, E. Horvitz, and A. Kapoor. Nobody likes mondays: foreground detection and behavioral patterns analysis in complex urban scenes. In *Proceedings of the 4th ACM/IEEE international workshop on Analysis and retrieval of tracked events and motion in imagery stream*, pages 17–24. ACM, 2013.