Learning from Outdoor Webcams: Surveillance of Physical Activity Across Environments

J. Aaron Hipp, PhD[1,2*] , Deepti Adlakha, MUD[1], Amy A. Eyler, PhD[1,2], Rebecca Gernes, MSW, MPH[1], Agata Kargol[3], Abigail H. Stylianou[3], Robert Pless, PhD[3]

Washington University in St. Louis, Missouri: [1]Brown School, [2]Prevention Research Center, and [3]Computer Science and Engineering

*Corresponding author information: J. Aaron Hipp, PhD
One Brookings Drive, Washington University in St. Louis, Campus Box 1196
St. Louis, Missouri 63130
[o] 314.935.3868 [f] 314.935.8511 [e] ahipp@wustl.edu

Keywords: webcams, physical activity, built environment, crowdsourcing, outdoor environments

Abstract

There are tens of thousands of publicly available webcams which constantly view the world and share those images. These cameras include traffic cams, campus cams, ski-resort cams, etc. The Archive of Many Outdoor Scenes (AMOS) is a project that aims to geo-locate, calibrate, annotate, archive and visualize these cameras to serve as an imaging resource for a wide variety of scientific applications. Here we report on a multi-disciplinary project to demonstrate and evaluate the potential for webcams to be re-purposed as a tool to evaluate patterns of population-level physical activity behavior in diverse urban built environments.

The AMOS dataset has archived over 560 million images of outdoor environments from 27,000 webcams since 2006. The primary goal is to employ the AMOS dataset and crowdsourcing to develop reliable and valid tools to improve physical activity assessment via online, outdoor webcam capture of global physical activity patterns and urban built environment characteristics. This goal will be accomplished by addressing two subsequent aims:
Aim 1: Develop and test reliability of using publicly available, outdoor webcams to enumerate built environment characteristics and physical activity patterns across thousands of global outdoor environments.
Aim 2: Develop and test reliability and validity of using crowdsourcing to enumerate built environment characteristics and physical activity patterns across thousands of global outdoor environments.

This project's grand scale-up of capturing physical activity patterns and built environments is a methodological step forward in advancing a real-time, non-labor intensive assessment using webcams and crowdsourcing. The combined use of webcams capturing outdoor scenes every 30 minutes and crowdsources providing the labor of annotating the scene allows for accelerated public health surveillance related to physical activity in built environments. The ultimate goal of this public health and computer vision collaboration is to develop machine learning algorithms that will automatically identify and calculate physical activity patterns.

INTRODUCTION

Kevin Lynch's 1960 book, 'The Image of the City', was one of the first to emphasize the importance of social scientists and design professionals in signifying ways that urban design and built environment can be quantitatively measured and improved (Lynch, 1960). It led to enormous efforts to investigate the structure and function of cities, to characterize perception of neighborhoods (J. Jacobs, 1961; Xu, Weinberger, & Chapelle, 2012), and promotion of social interactions (Milgram, Sabini, & Silver, 1992; Oldenburg, 1989). To date, large scale studies seeking to understand and quantify how specific features or changes in the built environment impact individuals, their behavior, and interactions, have required extensive in-the-field observation. However, they only provide a limited view of behaviors, their context, and how each may change as a function of the built environment. These studies are time intensive and expensive, deploying masses of graduate students to conduct interviews about people's daily routines (Milgram et al., 1992) or requiring hand-coding of thousands of hours of video (Whyte, 1980) to characterize a few city plazas and parks. Even current state-of-the art technology to investigate associations between behavior and the urban built environment uses multiple expensive devices at the individual level (GPS and accelerometer) and connects this data to Geographic Information System (GIS) layers known to often be unreliable (James et al., 2014; Kerr, Duncan, & Schipperijn, 2011; Schipperijn et al., 2014).

A key population behavior of interest to our team is physical activity (Adlakha, Budd, Gernes, Sequeira, & Hipp, 2014; Eyler, Brownson, Schmid, & Pratt, 2010; Hipp, Adlakha, Eyler, Chang, & Pless, 2013). Physical activity plays a role in numerous health outcomes including obesity, diabetes, heart disease, and cancer (Office of the Surgeon General, 2011). Over 30% of adults and 17% of children and adolescents in the US are obese (CDC, 2009), with lack of physical activity due to constraints in the built environment being an important influence (O. Ferdinand, Sen, Rahurkar, Engler, & Menachemi, 2012). Lack of safe places to walk and bicycle and lack of access to parks and open space can impact the frequency, duration, and quality of physical activity of residents in urban settings (Brownson, Hoehner, Day, Forsyth, & Sallis, 2009; Jackson, 2003; Jackson, Dannenberg, & Frumkin, 2013). Physical activity may be purposive such as a jog in a park, or incidental such as a ten minute walk from home to a public transit stop. In both purposive and incidental cases the designs of urban built environments influence the decisions and experience of physical activity behaviors. As such, the US Guide to Community Preventive Services (Community Guide) currently recommends the following built environment interventions to increase physical activity behaviors and reduce obesity: (1) community and street-scale urban design and land use policies; (2) creation of, or enhanced access to places for physical activity; and (3) transportation policies and practices (CDC, 2011).

Physical Activity Assessment. Physical activity and built environment research has expanded during the past 20 years (Handy, Boarnet, Ewing, & Killingsworth, 2002; O. Ferdinand et al., 2012). The research has followed traditional patterns of growth beginning with ecological studies of association (Ewing, Meakins, Hamidi, & Nelson, 2003), then local validation of associations via retrospective surveys and researcher-present observation (Bedimo-Rung, Gustat, Tompkins, Rice, & Thomson, 2006; McKenzie & Cohen, 2006). For example, the System for Observing Physical Activity and Recreation in Communities (SOPARC) (McKenzie & Cohen, 2006) was developed to understand physical activity in context with the environment while being unobtrusive. SOPARC continues to be a popular method of assessing physical activity with pairs of researchers positioning themselves in numerous target areas to scan the environment for numbers participating in sedentary to vigorous physical activity (Baran et al., 2013; Cohen, Marsh, Williamson, Golinelli, & McKenzie, 2012; Reed, Price, Grost, & Mantinan, 2012). Presently, natural experiments related to physical activity patterns and built environments are growing in popularity (Cohen et al., 2012). These studies have been of great benefit to the field by informing public health and urban design. While there is now a substantial body of evidence to inform local interventions and policies (Ding & Gebel, 2012; Feng, Glass, Curriero, Stewart, & Schwartz, 2010; Kaczynski & Henderson, 2007; Renalds, Smith, & Hale, 2010; Saelens & Handy, 2008; Sandercock, Angus, & Barton, 2010), currently used methodologies and the use of small, local samples limit the external validity and dissemination of many results, interventions, and policies. There is a need for large-scale, evidence-informed evaluations of physical activity to increase external validity as evident in recent calls for more studies across a greater variety of environments (Cerin, Conway, Saelens, Frank, & Sallis, 2009; Dyck et al., 2012).

Big Data Opportunities. Big data and modern technology has opened up several opportunities to obtain new insights on cities and offer the potential for dramatically more efficient measurement tools (Graham & Hipp, 2014; Hipp, 2013) . The relative ease of capturing large sample data has led to amazing results that highlight how people move through cities based on check-ins (Naaman, 2011; Silva, Melo, Almeida, Salles, & Loureiro, 2012) or uploaded photos (Crandall, Backstrom, Huttenlocher, & Kleinberg, 2009). In addition, GIS, GPS, accelerometers, smart phone applications (apps), and person-point-of-view cameras are each being used in many studies, often in combination (Graham & Hipp, 2014; Hurvitz, Moudon, Kang, Saelens, & Duncan, 2014; Kerr et al., 2011). Apps that track running and walking routes are being investigated for where populations move and how parks and other built environment infrastructure may be associated with such movement (Adlakha et al., 2014; Hirsch et al., 2014).

Though these big data sources offer important contributions to the field of physical activity and built environment research, they are each dependent on individuals to upload data, allow access to data, and/or agree to wear multiple devices. This is the epitome of the quantified-self movement (Barrett, Humblet, Hiatt, & Adler, 2013). A complementary alternative big data source is the pervasive capture of urban environments by traffic cameras and other public, online webcams. This environmental-point-of-view imaging also captures human behavior and physical activity as persons traverse and use urban space.

The Archive of Many Outdoor Scenes (AMOS) has been archiving one image each half hour from most online, publicly available webcams for the last 8 years, creating an open and widely distributed research resource (Pless & Jacobs, 2006). AMOS began to collect images from these 27,000 webcams mapped in Figure 1 to understand the local effects of climate variations on plants. We have used these large collections of up-close, on the ground measurements to suggest corrections to standard satellite data products like NASA’s Moderate Resolution Imaging Spectroradiometer (MODIS) estimates of tree growing seasons (Ilushin, Richardson, Toomey, Pless, & Shapiro, 2013; N. Jacobs et al., 2009; Richardson, Friedl, Frolking, Pless, & Collaborators, 2011). This global network of existing cameras also captures images of public spaces – plazas, parks, street intersections, waterfronts – creating an archive of how public spaces have changed over time and what behaviors are being performed within these spaces.

With its archive of over 550 million captured images, AMOS not only represents 27,000 unique environments, but is capturing concurrent behaviors in and across the environments. Unique and of significance to public health surveillance, the online, publicly available webcams are non-biased in data collection, consistent and thorough (an image each half hour), and timely (images instantly added to the archive and available to the public). The AMOS project provides an opportunity to virtually annotate changes in the built environment and associated physical activity behaviors. This dataset can provide a census of physical activity patterns within captured environments during the past eight years and moving forward. Due to the size of the AMOS dataset, we have used crowdsourcing to help annotate the captured scenes.

Use of crowdsourcing in public health research. Crowdsourcing refers to and utilizes the masses, or crowds, of individuals using the Internet, social media, and social smartphone apps. The crowds participating in these websites and applications are the source of data or the source of needed labor (Kamel Boulos et al., 2011). Crowdsourcing data collection in public health is an emerging field, with examples including the collection of tweets and Google searches that detected an increase in influenza before the increase in subsequent influenza-related hospital visits (Ginsberg et al., 2009; Kamel Boulos et al., 2011). Another potential use of crowdsourcing is as the labor in evaluation or assessment of research hypotheses (Bohannon, 2011; Buhrmester, Kwang, & Gosling, 2011; Office of the Surgeon General, 2011). The present team was the first to publish on the use of crowdsourcing as physical activity annotators (Hipp et al., 2013). A crowdsource marketplace, i.e., Amazon Mechanical Turk, can be used to ask workers to complete Human Intelligence Tasks (HITs) such as drawing a box around each pedestrian in a captured image.

Objectives of Current Work. The primary goal of our ongoing collaboration is to use the AMOS dataset and crowdsourcing to develop reliable and valid tools to improve physical activity behavior assessment. This goal will be accomplished by addressing two subsequent aims:

Aim 1: Develop and test the reliability of using publicly-available, outdoor webcams to enumerate built environment characteristics and physical activity patterns across thousands of global outdoor environments.

Aim 2: Develop and test the reliability and validity of using crowdsourcing to enumerate built environment characteristics and physical activity patterns across thousands of global outdoor environments.

DATA SOURCES

Archive of Many Outdoor Scenes (AMOS). The publicly captured scenes of human behavior, physical activity, and urban built environments are all from the AMOS dataset. AMOS is a Washington University project which aims to capture and archive images from every publicly available, online, outdoor webcam (e.g., traffic cams, campus cams, ski-resort cams, etc. – See Figure 1). This dataset was developed primarily as a basis to research computer vision algorithms for geo-locating and calibrating cameras, and as a demonstration that webcams can be re-purposed as a complement to satellite imaging for large-scale climate measurement (N. Jacobs et al., 2009; N. Jacobs, Roman, & Pless, 2008). Images are digitally captured from each camera every 30 minutes and archived in a searchable dataset.

Our current work builds on the AMOS model system for working with, sharing, and crowdsourcing big data. Figure 2 shows a screenshot of a main data access page, showing (A) one image and (B) the time this specific image was captured. A yearly summary image, indexed by time of year on the x-axis, and time of day on the y-axis is shown in (C). This summarizes a year of images with each pixel as a representation of the image at that time of year and time of day. Pixels can also be represented using principal component analysis to quickly identify images that differ based on precipitation, snowfall, dusk, dawn, etc. This summary serves several purposes. First, it is a data availability visualization, where dark red highlights when the camera was down and did not capture images. Second, it highlights annual patterns such as the summer nights being shorter than winter nights. Third, data capture problems are often visible. Finally, this data visualization is “clickable” so that a user can see, by clicking, the image from a particular time of day and time of year.

Each camera also contains extensive metadata as outlined in the Figure 2: (D) Shows the geo-location of the camera; (E) Shows free form text tags that we and other groups use to keep track of and search for cameras with particular properties; (F) is a new feature added for this present project that allows the tagging of specific images (instead of cameras), and (G) is a pointer to zip-files for data from this camera or a python script to allow selective downloading. When exact camera locations are known, the cameras may be geo-oriented and calibrated relative to global coordinates as shown in Figure 1.

Amazon.com's Mechanical Turk Crowdsource. Virtual audits have emerged as a reliable method to process the growing volume of web-based data on the physical environment (Badland, Opit, Witten, Kearns, & Mavoa, 2010; Clarke, Ailshire, Melendez, Bader, & Morenoff, 2010; C. L. Odgers, A. Caspi, C. J. Bates, R. J. Sampson, & T. E. Moffitt, 2012). Research has also turned to virtual platforms as a way to recruit study participants and complete simple tasks (Hipp et al., 2013; Kamel Boulos et al., 2011). The Amazon.com Mechanical Turk (MTurk) website outsources Human Intelligence Tasks (HITs), or tasks that have not yet been automated by computers. Workers may browse available HITs and are paid for every HIT completed successfully (Buhrmester et al., 2011). MTurk workers are paid a minimum of US$0.01 per HIT, making them a far less expensive option than traditional research assistant annotators (Berinsky, Huber, & Lenz, 2012). MTurk was found as an effective method for survey participant recruitment, with more representative and valid results than the convenience sampling often used for social science research (Bohannon, 2011). MTurk has also been used for research task completion such as transcription and annotation. These have generally been small in scale and MTurk reliability for larger scale data analysis has not been established (Hipp et al., 2013). Within MTurk, our team has designed a unique web-form used with the MTurk HIT that allows amateur workers to annotate our images by demarcating each pedestrian, bicyclist, and vehicle per photograph.

Trained Research Assistants. Trained undergraduate and graduate Research Assistants from the computer science and public health departments at Washington University in St. Louis have annotated images for physical activity behaviors and built environment attributes. For both behaviors and environments, Research Assistants were provided with example captured scenes. Project Principal Investigators supervised the scene annotation process and provided real-time feedback on uncertain scenes. Difficult or exceptional scenes and images were presented to the research group to ensure that all behaviors and environments were annotated in a consistent manner.

METHODS

Annotating Physical Activity Behaviors. We have used 12 traffic webcams located in Washington, DC, to determine initial feasibility of the physical activity behavior research agenda. AMOS has archived a photograph every thirty minutes from Washington, DC, Department of Transportation webcams. Since 2007, Washington, DC, has initiated multiple built environment improvements to increase physical activity behaviors, including a bicycle share program, miles of new bike lanes, and painted crosswalks. For example, a new bicycle lane was added in the middle of Pennsylvania Avenue in spring 2010, and AMOS has an archive of captured images every thirty minutes for the year prior to the installation of the bike lane, and a year following installation.

The MTurk website was used to crowdsource the image annotation. In a pilot project we uploaded each webcam photograph captured by AMOS at the intersection of Pennsylvania Avenue NW and 9th Street NW between 7am and 7pm the first work week of June 2009 and June 2010 to the MTurk webpage (Hipp et al., 2013). There we designed a HIT that allowed MTurk workers to annotate our images by marking each pedestrian, bicyclist, and vehicle in each captured scene. MTurk workers used their computer mouse to hover over the appropriate behavior, e.g., pedestrian activity, and left-click atop each individual pedestrian. Five unique MTurk workers completed this task for all three transportation behaviors per image. The numbers of each type of annotation were then downloaded to a spreadsheet and imported into SPSS.

In related ongoing work, we have used 12 different AMOS webcams that captured other built environment changes at intersections in Washington, DC, between 2007 and 2010. We have made improvements to our MTurk task by asking workers to use their cursors to draw polygons, or boxes, around the specified transportation behavior (walking, cycling, driving). Similar to the first HIT, we used each photograph between 7:00am and 7:00pm during the first week of June proceeding and following a built environment change. Finally, we posted to MTurk every photograph from two of the above intersections between 6:00am and 9:00pm for 19 consecutive months (five months prior to a crosswalk being introduced to the intersections and 14 months post).

MTurk workers were paid US$0.01 or US$0.02 per scene to mark each pedestrian, cyclist, and vehicle in an image and took on average 71 seconds to complete each task. Each image was annotated five unique times. Two trained Research Assistants completed the same task, annotating each image twice. Training took place in two sessions. In the first session, Research Assistants received the same instructions as MTurk participants and completed a practice set of 100 images. In the second session, Research Assistants compared their practice results and discussed differences in analysis. Research Assistants completed the camera annotations in separate forms, and their results were averaged.

Annotating Built Environments. Selecting the appropriate built environment image tags was an iterative process. First, we selected two commonly used built environment audit tools to establish a list of potential built environment tags. These were the Environmental Assessment of Public Recreation Spaces (EAPRS) (Saelens et al., 2006) and the Irvine-Minnesota Inventory (Day, Boarnet, Alfonzo, & Forsyth, 2006). From an initial list of 73 built environment items that we believed could be annotated using captured images we narrowed the final list down to 21 built environment tags. Following the combination of similar terms, we further reduced the potential list of tags based on the inclusion criteria that the tag must be theoretically related to human behaviors.

To establish which of the 27,000 AMOS webcams are at an appropriate urban built environment scale, i.e., those with the potential of capturing physical activity, our team designed an interface that selects a set of camera IDs, and displays 25 cameras per screen. This internal HIT was created to populate a webpage with the 25 unique camera images. Below each image was a green checkmark and a red x-mark. If physical activity behaviors could be captured in the scene, the green checkmark was selected and this tag automatically added to a dataset of physical activity behavior cameras. This process was repeated with trained Research Assistants for reliability and resulted in a set of 1,906 cameras. In addition to the above inclusion criteria, selected cameras must have captured scenes from at least 12 consecutive months. The final 21 built environment tags are presented in Table 1.

To tag each camera, Research Assistants were provided a one-page written and photographic example (from AMOS dataset) of each built environment tag. For example, a written description for a crosswalk was provided along with captured images of different styles of crosswalks from across the globe. A second internal HIT was created similar to the above that populated a webpage with 20 unique camera images, each marked with a green checkmark and a red x-mark. If the provided built environment tag (e.g., crosswalk) was present in the image then the green checkmark was selected and this tag automatically added to the camera annotation. If a Research Assistant was unsure they could click on the image to review other images captured by the same camera or could request the assistance of other Research Assistants of Principal Investigators to verify their selection. This process was completed for all 21 built environment tags across all 1,906 cameras in the AMOS physical activity dataset. To date, the built environment tags have only been annotated by trained Research Assistants. Reliability and validity of tags is a future step of this research agenda. This initial step provided the team a workable set of publicly available webcams to address our two study aims.

DATA ANALYSIS

Physical Activity Behaviors. In the pilot project we used t-tests and logistic regressions to analyze the difference in physical activity behaviors before and after the addition of the bike lane along Pennsylvania Avenue. T-tests were used for pedestrians and vehicles, where the data was along a continuous scale from 0-20 (20 being the most captured in any one scene). Logistic regression was used for the presence or absence of a cyclist in each image.

Reliability and Validity. Inter-rater reliability (IRR) and validity statistics (Pearson’s R, Inter-Class Correlations, and Cohen’s Kappa) were calculated within and between the five MTurk workers and between the two trained Research Assistants. The appropriate statistic was calculated for two, three, four, or five MTurk workers to determine the optimal number of workers necessary to capture a reliable and valid count of pedestrians, cyclists, and vehicles in a scene. Due to each scene being annotated by five unique MTurk workers we were able to test the reliability of ten unique combinations of workers; that is, Worker 1 and Worker 2, Worker 1 and Worker 3, Worker 1 and Worker 4, etc. Similar combinations were used with three workers (ten unique combinations) and four workers (five unique combinations). Each combination was compared to the trained Research Assistants results to measure validity. For all tests we used Landis and Koch's magnitudes of agreement: <0.19 (poor agreement), 0.20-0.39 (fair), 0.40-0.59 (moderate), 0.60-0.79 (substantial) and >0.80 (near perfect agreement) (Landis & Koch, 1977).

RESULTS

Pilot Project. Previously published results reveal that publicly available, online webcams are capable of capturing physical activity behavior and are capable of capturing changes in these behaviors pre and post built environment changes (Hipp et al., 2013). The odds of the traffic webcam at Pennsylvania Avenue NW and 9th Street NW capturing a cyclist present in the scene in 2010 increased 3.5 times, compared to 2009 (OR=3.57, $p<0.001$). The number of cyclists per scene increased four-fold between 2009 (mean=0.03; SD=0.20) and 2010 (0.14; 0.90; F=36.72, 1198; $p=0.002$). Both results are associated with the addition of the new bike lane. There was no associated increase in the number of pedestrians at the street intersection following the addition of the bike lane, as may be theoretically expected with a bicycle-related built environment change, not a pedestrian-related change.

Reliability Assessment. Next, we tested reliability and validity of using publicly available webcams and MTurk HITs to annotate captured scenes for physical activity and transportation behaviors. Reliability statistics varied across MTurk workers based on the number annotating each scene and the annotation task (pedestrians compared to cyclists).

For pedestrians (n = 720 images), pairs of MTurk workers had an agreement average and a Pearson's R-score of 0.562 (range: 0.122 – 0.866). The Inter-Class Correlation (ICC) for three MTurk workers averaged 0.767 (0.330 – 0.944) and four workers averaged 0.814 (0.534 – 0.954). The average for all five workers across the 720 scenes was 0.848 (0.687 – 0.941). The ICCs for four and five workers represented near-perfect agreement. The pair of trained Research Assistants averaged a Pearson's R-score 0.850 (0.781 – 0.925), also representing near perfect agreement.

The averages and ranges of annotator agreement for presence of cyclists in 2007 were as follows (Table 2): two workers (Cohen's Kappa: 0.333; Range: 0.000 – 0.764), three workers (0.553; 0.000 – 0.897), four workers (0.607; 0.000 – 0.882), five workers (0.645; 0.000 – 0.874), and Research Assistants (0.329; 0.000 – 0.602). Annotator agreement with four and five MTurk workers showed substantial agreement. For the pilot project presented above, we used the average of five MTurk workers. When analyzing presence versus absence, majority ruled; if three or more of the five MTurk workers annotated a cyclist was present, then this scene received a 1. If two or fewer annotated a cyclist, the scene received a 0.

The averages and ranges for number of vehicles were as follows: two workers (0.354; 0.000 – 0.769), three workers (0.590; 0.208 – 0.830), four workers (0.653; 0.398 – 0.830), five workers (0.705; 0.592 – 0.837), and Research Assistants (0.885; 0.841 – 0.922). The reliability statistics for four and five MTurk workers again showed substantial rater/annotator agreement, and near perfect agreement between the two Research Assistants.

Validity Assessment. From reliability estimates, we concluded that using four MTurk workers was the most reliable and cost-efficient method. Next, validity statistics were calculated for four MTurk workers and two trained RAs. Validity statistics (Pearson's R) for pedestrians (0.846 – 0.901) and vehicles (0.753 – 0.857) showed substantial to near perfect agreement. Validity (Cohen's kappa) for cyclists (0.361 – 0.494) were in the fair-moderate agreement range.

Built Environment Tags. As provided in Table 1, our final list of built environment tags includes 21 unique items. The number of cameras with the tag present is also presented. 'Buildings' was found the most frequent, present and at a scale to capture human behavior across 1,245 webcams. 'Bike racks' was annotated the fewest times, only occurring in 27 scenes. Figure 4 shows an example map of where each of the cameras with the built environment tag of 'open space' is located.

DISCUSSION

The use of public, captured imagery to annotate built environments for public health research is an emerging field. To date the captured imagery has been static and only available via Google Streetview and Google Satellite imagery (Charreire et al., 2014; Edwards et al., 2013; Kelly, Wilson, Baker, Miller, & Schootman, 2012; Kelly et al., 2014; Candice L. Odgers, Avshalom Caspi, Christopher J. Bates, Robert J. Sampson, & Terrie E. Moffitt, 2012; Rundle, Bader, Richards, Neckerman, & Teitler, 2011; B. T. Taylor et al., 2011; J. R. Taylor & Lovell, 2012; Wilson & Kelly, 2011; Wilson et al., 2012). There have been no attempts to crowdsource this image annotation, nor combine annotation of built environments and images capturing physical activity behaviors. Using an eight-year archive of captured webcam images and crowdsources, we have demonstrated that improvements in urban built environments are associated with subsequent and significant increases in physical activity behaviors. Webcams are able to capture a variety of built environment attributes and our study shows webcams are a reliable and valid source of built environment information. As such, the emerging technology of publicly available webcams facilitates both consistent uptake and potentially timely dissemination of physical activity and built environment behaviors across a variety of outdoor environments. The AMOS webcams have the potential to serve as an important and cost-effective part of urban environment and public health surveillance to evaluate patterns and trends of population-level physical activity behavior in diverse built environments.

In addition to presenting a new way to study physical activity and the built environment, our findings contribute to novel research methodologies. The use of crowdsources (Amazon's Mturk) proved to be a reliable, valid, inexpensive, and quick method for annotating street scenes captured by public, online webcams. While MTurk workers have previously been found to be a valid and reliable source of participant recruitment for experimental research, this is the first research agenda that has found MTurk to be a valid and reliable method for content analysis (Buhrmester et al., 2011, Berinsky et al., 2012, Hipp et al., 2013). Our results indicate taking the average annotation of four unique MTurk workers appears to be the optimal threshold. Our results also show that across each mode of transportation assessed, the average reliability score with four unique workers was 0.691, which is considered substantial agreement (Landis & Koch, 1977).

In addition to substantial agreement between the MTurk workers, the trained RAs yielded substantial agreement with vehicles, near perfect agreement with pedestrians, but only fair agreement with cyclists. The cyclists' statistics were the least reliable, primarily due to the low number of images (only 10% of captured scenes) with a cyclist present. Similar to reliability statistics, validity was near perfect for pedestrians and vehicles, but only fair to moderate for cyclists. These results suggest MTurk workers are a quick, cheap annotation resource for commonly captured image artifacts. However, MTurk is not yet primed to capture rare events in captured scenes without additional instruction or limitations to the type of workers allowed to complete tasks.

Our current big data and urban informatics research agenda shows that publicly available, online webcams offer a reliable and valid source for measuring physical activity behavior in urban settings. Our findings lay the foundation for studying physical activity and built environment characteristics using the magnitude of available globally recorded images as measurements. The research agenda is innovative in: (1) its potential to characterize physical activity patterns over the time-scale of years, with orders of magnitude more measurements than would be feasible by standard methods, (2) the ability to use the increase in data to characterize complex interactions between physical activity patterns, seasons and weather, and (3) its capacity to be an ongoing, systematic public health surveillance system. In addition to increasing the capacity of physical activity research, the methodologies described here are of novel interest to computer vision researchers. Automating algorithms to detect and quantify behavioral transformations due to changes in urban policy and built infrastructure can transform aspects of this research as well.

These findings have several implications related to cost and timeliness for the use of MTurks in content analysis. The total cost for the MTurk analysis was $320.00 for 32,001 images, compared to $1,333.33 for a trained Research Assistant paid at US$10 per hour. This could be due to several cost-saving characteristics of crowdsourcing. MTurk workers are paid for each successful completion of a HIT, compared to hourly wages for a Research Assistant. This allows MTurk to pay multiple workers at a time for each HIT, at a cost substantially lower than the same number of trained Research Assistants. The higher speed and lower cost of crowdsourced analysis is especially suitable for annotating AMOS data, to which thousands of images are added daily. Reliable and rapid image annotation using MTurks could allow for large-scale and more robust analysis of results that would be too costly to complete with traditional analysis. Thus far our team has looked at 12 cameras in one metro area. Future studies could increase the number of cameras annotated for a specific area or time, compare results across metro regions, or analyze environmental effects such as weather, season, and day of the week on mode of transportation.

There are several ethical and human subjects concerns related to publicly available, online webcams and the use of MTurk. With our initial research projects, we have received exempt status for the use of both AMOS and MTurk. The webcams were exempt because we are not collecting individual identifiable private information; this activity is not considered to meet federal definitions under the jurisdiction of an Institutional Review Board and therefore falls outside the purview of the Human Research Protection Office. AMOS is an archival dataset of publicly available photos. The photos are being used for counts and annotation of physical activity patterns and built environment attributes and are not concerned with individual or identifiable information. To date, no camera has been identified that is at an angle and height so as to distinguish an individual's face. The use of publicly available webcams fits with the 'Big Sister' approach to the use of cameras for human-centered design and social values (Stauffer & Grimson, 2000). Related, recent research utilizing Google Street View and Google Earth images have also been HRPO-exempt ("National Center for Safe Routes to School," 2010; Sadanand & Corso, 2012; Saelens et al., 2006; Sequeira, Hipp, Adlakha, & Pless, 2013).

Finally, AMOS is quite literally "Seeing Cities Through Big Data" with applications for research methods and urban informatics. With thoughtful psychometrics and application of this half-billion image dataset, and growing, we believe pervasive webcams can assist researchers and urban practitioners alike in better understanding how we use place and how the shape and context of urban places influence our movement and behavior.

References

Adlakha, Deepti, Budd, Elizabeth L., Gernes, Rebecca, Sequeira, Sonia, & Hipp, James Aaron. (2014). Use of emerging technologies to assess differences in outdoor physical activity in St. Louis, Missouri. *Frontiers in Public Health, 2*(41). doi: 10.3389/fpubh.2014.00041

Badland, H. M., Opit, S., Witten, K., Kearns, R. A., & Mavoa, S. (2010). Can virtual streetscape audits reliably replace physical streetscape audits? *J Urban Health, 87*(6), 1007-1016. doi: 10.1007/s11524-010-9505-x

Baran, Perver K., Smith, William R., Moore, Robin C., Floyd, Myron F., Bocarro, Jason N., Cosco, Nilda G., & Danninger, Thomas M. (2013). Park Use Among Youth and Adults: Examination of Individual, Social, and Urban Form Factors. *Environment and Behavior*. doi: 10.1177/0013916512470134

Barrett, Meredith A., Humblet, Olivier, Hiatt, Robert A., & Adler, Nancy E. (2013). Big Data and Disease Prevention: From Quantified Self to Quantified Communities. *Big Data, 1*(3), 168-175. doi: 10.1089/big.2013.0027

Bedimo-Rung, A. , Gustat, J., Tompkins, B.J., Rice, J. , & Thomson, J. (2006). Development of a direct observation instrument to measure environmental characteristics of parks for physical activity. . *J Phys Act Health, 3*(1S), S176–S189.

Berinsky, Adam J., Huber, Gregory A., & Lenz, Gabriel S. (2012). Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk. *Political Analysis, 20*, 351-368. doi: 10.1093/pan/mpr057

Bohannon, J. (2011). Social Science for Pennies. *Science, 334*, 307.

Brownson, R. C., Hoehner, C. M., Day, K., Forsyth, A., & Sallis, J.F. . (2009). Measuring the Built Environment for Physical Activity: State of the Science. *American Journal of Preventive Medicine, 36*(4 Supplement), S99-123.e112. doi: 10.1016/j.amepre.2009.01.005

Buhrmester, Michael, Kwang, Tracy, & Gosling, Samuel D. (2011). Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science, 6*(1), 3-5. doi: 10.1177/1745691610393980

CDC. (2009). Division of Nutrition, Physical Activity and Obesity. *Available from: http://www.cdc.gov/nccdphp/dnpa/index.htm*.

CDC. (2011). Guide To Community Preventive Services. Atlanta, GA: Epidemiology Program Office, CDC.

Cerin, Ester, Conway, Terry L, Saelens, B.E., Frank, Lawrence D, & Sallis, James F. (2009). Cross-validation of the factorial structure of the Neighborhood Environment Walkability Scale (NEWS) and its abbreviated form (NEWS-A). *International Journal of Behavioral Nutrition and Physical Activity, 6*(1), 32.

Charreire, H., Mackenbach, J. D., Ouasti, M., Lakerveld, J., Compernolle, S., Ben-Rebah, M., . . . Oppert, J. M. (2014). Using remote sensing to define environmental characteristics related to physical activity and dietary behaviours: A systematic review (the SPOTLIGHT project). *Health & Place, 25*(0), 1-9. doi: http://dx.doi.org/10.1016/j.healthplace.2013.09.017

Clarke, P., Ailshire, J., Melendez, R., Bader, M., & Morenoff, J. (2010). Using Google Earth to conduct a neighborhood audit: reliability of a virtual audit instrument. *Health Place, 16*(6), 1224-1229. doi: 10.1016/j.healthplace.2010.08.007

Cohen, D. A., Marsh, T., Williamson, S., Golinelli, D., & McKenzie, T. L. (2012). Impact and cost-effectiveness of family Fitness Zones: a natural experiment in urban public parks. *Health Place, 18*(1), 39-45. doi: 10.1016/j.healthplace.2011.09.008

Crandall, David J, Backstrom, Lars, Huttenlocher, Daniel, & Kleinberg, Jon. (2009). Mapping the world's photos *Proceedings of the 18th international conference on World wide web* (pp. 761-770).

Day, Kristen, Boarnet, Marlon, Alfonzo, Mariela, & Forsyth, Ann. (2006). The Irvine–Minnesota Inventory to Measure Built Environments: Development. *American Journal of Preventive Medicine, 30*(2), 144-152. doi: http://dx.doi.org/10.1016/j.amepre.2005.09.017

Ding, Ding, & Gebel, Klaus. (2012). Built environment, physical activity, and obesity: What have we learned from reviewing the literature? *Health & Place, 18*(1), 100-105. doi: 10.1016/j.healthplace.2011.08.021

Dyck, Delfien Van, Cerin, Ester, Conway, Terry L, Bourdeaudhuij, Ilse De, Owen, Neville, Kerr, Jacqueline, . . . Sallis, James F. (2012). Perceived neighborhood environmental attributes associated with adults' transport-related walking and cycling: Findings from the USA, Australia and Belgium. *International Journal of Behavioral Nutrition and Physical Activity., 9*(1), 70. doi: 10.1186/1479-5868-9-70

Edwards, Nicole, Hooper, Paula, Trapp, Georgina S. A., Bull, Fiona, Boruff, Bryan, & Giles-Corti, Billie. (2013). Development of a Public Open Space Desktop Auditing Tool (POSDAT): A remote sensing approach. *Applied Geography, 38*(0), 22-30. doi: http://dx.doi.org/10.1016/j.apgeog.2012.11.010

Ewing, Reid, Meakins, Gail, Hamidi, Shima, & Nelson, Arthur. (2003). Relationship Between Urban Sprawl and Physical Activity, Obesity, and Morbidity. *American Journal of Health Promotion, 18*(1), 47-57.

Eyler, Amy, Brownson, R. C., Schmid, Tom, & Pratt, M. (2010). Understanding policies and physical activity: frontiers of knowledge to improve population health. *J Phys Act Health, 7*(1543-3080 (Print)), S9-12.

Feng, Jing, Glass, Thomas A., Curriero, Frank C., Stewart, Walter F., & Schwartz, Brian S. (2010). The built environment and obesity: A systematic review of the epidemiologic evidence. *Health & Place, 16*(2), 175-190. doi: 10.1016/j.healthplace.2009.09.008

Ginsberg, Jeremy, Mohebbi, Matthew H., Patel, Rajan S., Brammer, Lynnette, Smolinski, Mark S., & Brilliant, Larry. (2009). Detecting influenza epidemics using search engine query data. *Nature, 457*(7232), 1012-1014. doi: http://www.nature.com/nature/journal/v457/n7232/suppinfo/nature07634_S1.html

Graham, D. J., & Hipp, J. A. (2014). Emerging technologies to promote and evaluate physical activity: cutting-edge research and future directions. *Front Public Health, 2*, 66. doi: 10.3389/fpubh.2014.00066

Handy, SL, Boarnet, MG, Ewing, R, & Killingsworth, RE. (2002). How the built environment affects physical activity: views from urban planning. *Am J Prev Med, 23*, 64 - 73.

Hipp, J. Aaron. (2013). Physical activity surveillance and emerging technologies. *Brazilian Journal of Physical Activity and Health, 18*(1), 2-4. doi: 10.12820/2317-1634.2013v18n1p2

Hipp, J. Aaron, Adlakha, Deepti, Eyler, Amy A., Chang, Bill, & Pless, Robert. (2013). Emerging Technologies: Webcams and Crowd-Sourcing to Identify Active Transportation. *American Journal of Preventive Medicine, 44*(1), 96-97. doi: 10.1016/j.amepre.2012.09.051

Hirsch, J. A., James, P., Robinson, J. R., Eastman, K. M., Conley, K. D., Evenson, K. R., & Laden, F. (2014). Using MapMyFitness to Place Physical Activity into Neighborhood Context. *Front Public Health, 2*, 19. doi: 10.3389/fpubh.2014.00019

Hurvitz, P. M., Moudon, A. V., Kang, B., Saelens, B. E., & Duncan, G. E. (2014). Emerging technologies for assessing physical activity behaviors in space and time. *Front Public Health, 2*, 2. doi: 10.3389/fpubh.2014.00002

Ilushin, D, Richardson, AD, Toomey, MP, Pless, R, & Shapiro, A. (2013). Comparing the effects of Different Remote Sensing Techniques for Extracting Deciduous Broadleaf Phenology *AGU Fall Meeting Abstracts* (Vol. 1, pp. 0542).

Jackson, R. J. (2003). The Impact of the Built Environment on Health: An Emerging Field. *Am J Public Health, 93*(9), 1382-1384. doi: 10.2105/AJPH.93.9.1382

Jackson, R. J., Dannenberg, Andrew L., & Frumkin, Howard. (2013). Health and the Built Environment: 10 Years After. *American Journal of Public Health, 103*(9), 1542-1544. doi: 10.2105/ajph.2013.301482

Jacobs, Jane. (1961). *The death and life of great American cities*: Random House LLC.

Jacobs, Nathan, Burgin, Walker, Fridrich, Nick, Abrams, Austin, Miskell, Kylia, Braswell, Bobby H., . . . Pless, Robert. (2009). The Global Network of Outdoor Webcams: Properties and Applications *ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL GIS)* (pp. 111-120).

Jacobs, Nathan, Roman, Nathaniel, & Pless, Robert. (2008). Toward Fully Automatic Geo-Location and Geo-Orientation of Static Outdoor Cameras *Proc. IEEE Workshop on Video/Image Sensor Networks* (pp. 1-6).

James, P., Berrigan, D., Hart, J. E., Hipp, J. A., Hoehner, C. M., Kerr, J., . . . Laden, F. (2014). Effects of buffer size and shape on associations between the built environment and energy balance. *Health Place, 27*, 162-170. doi: 10.1016/j.healthplace.2014.02.003

Kaczynski, A. T., & Henderson, K.A. (2007). Environmental correlates of physical activity: A review of evidence about parks and recreation. *Lesiure Sciences, 29*, 315-354.

Kamel Boulos, M. N., Resch, B., Crowley, D. N., Breslin, J. G., Sohn, G., Burtner, R., . . . Chuang, K. Y. (2011). Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: trends, OGC standards and application examples. *Int J Health Geogr, 10*, 67. doi: 10.1186/1476-072x-10-67

Kelly, Cheryl, Wilson, Jeffrey, Baker, Elizabeth, Miller, Douglas, & Schootman, Mario. (2012). Using Google Street View to Audit the Built Environment: Inter-rater Reliability Results. *Annals of Behavioral Medicine*, 1-5. doi: 10.1007/s12160-012-9419-9

Kelly, Cheryl, Wilson, Jeffrey S., Schootman, Mario, Clennin, Morgan, Baker, Elizabeth A, & Miller, Douglas K. (2014). The Built Environment Predicts Observed Physical Activity. *Frontiers in Public Health, 2*. doi: 10.3389/fpubh.2014.00052

Kerr, J., Duncan, S., & Schipperijn, J. (2011). Using global positioning systems in health research: a practical approach to data collection and processing. *Am J Prev Med, 41*(5), 532-540. doi: 10.1016/j.amepre.2011.07.017

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Lynch, Kevin. (1960). *The image of the city* (Vol. 11): MIT press.

McKenzie, T.L., & Cohen, D.A. (2006). System for Observing Play and Recreation in Communities (SOPARC). In Center for Population Health and Health Disparities (Ed.): RAND.

Milgram, Stanley, Sabini, John Ed, & Silver, Maury Ed. (1992). *The individual in a social world: Essays and experiments* . Mcgraw-Hill Book Company.

Naaman, Mor. (2011). Geographic Information from Georeferenced Social Media Data. *SIGSPATIAL Special, 3*(2), 54-61.

National Center for Safe Routes to School. (2010).

O. Ferdinand, Alva, Sen, Bisakha, Rahurkar, Saurabh, Engler, Sally, & Menachemi, Nir. (2012). The relationship between built environments and physical activity: a systematic review. *American journal of public health, 102*(10), e7-e13.

Odgers, C. L., Caspi, A., Bates, C. J., Sampson, R. J., & Moffitt, T. E. (2012). Systematic social observation of children's neighborhoods using Google Street View: a reliable and cost-effective method. *J Child Psychol Psychiatry, 53*(10), 1009-1017. doi: 10.1111/j.1469-7610.2012.02565.x

Odgers, Candice L., Caspi, Avshalom, Bates, Christopher J., Sampson, Robert J., & Moffitt, Terrie E. (2012). Systematic social observation of children's neighborhoods using Google Street View: a reliable and cost-effective method. *Journal of Child Psychology and Psychiatry, 53*(10), 1009-1017. doi: 10.1111/j.1469-7610.2012.02565.x

Office of the Surgeon General. (2011). *Overweight and obesity: at a glance* Retrieved from Available from: http://www.surgeongeneral.gov/topics/obesity/calltoaction/fact_glance.html.

Oldenburg, Ray. (1989). *The great good place: Cafés, coffee shops, community centers, beauty parlors, general stores, bars, hangouts, and how they get you through the day*: Paragon House New York.

Pless, R., & Jacobs, N. (2006). *The Archive of Many Outdoor Scenes, Media and Machines Lab, Washington University in St. Louis and University of Kentucky*. Retrieved from: http://amos.cse.wustl.edu/

Reed, J. A., Price, A. E., Grost, L., & Mantinan, K. (2012). Demographic characteristics and physical activity behaviors in sixteen Michigan parks. *J Community Health, 37*(2), 507-512. doi: 10.1007/s10900-011-9471-6

Renalds, A., Smith, T.H., & Hale, P.J. (2010). A systematic review of built environment and health. *Family & Community Health, 33*(1550-5057 (Electronic)), 68-78. doi: 10.1097/FCH.0b013e3181c4e2e5.

Richardson, AD, Friedl, MA, Frolking, S, Pless, R, & Collaborators, PhenoCam. (2011). PhenoCam: A continental-scale observatory for monitoring the phenology of terrestrial vegetation *AGU Fall Meeting Abstracts* (Vol. 1, pp. 0517).

Rundle, Andrew G., Bader, Michael D. M., Richards, Catherine A., Neckerman, Kathryn M., & Teitler, Julien O. (2011). Using Google Street View to Audit Neighborhood Environments. *American Journal of Preventive Medicine, 40*(1), 94-100.

Sadanand, Sreemananath, & Corso, Jason J. (2012). Action bank: A high-level representation of activity in video *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on* (pp. 1234-1241).

Saelens, B.E., Frank, L.D., Auffrey, C. , Whitaker, R.C., Burdette, H.L. , & Colabianchi, N. . (2006). Measuring physical environments of parks and playgrounds: EAPRS instrument development and inter-rater reliability. . *J Phys Act Health., 3*(1S), S190–S207.

Saelens, B.E., & Handy, S. (2008). Built environment correlates of walking: A review. *Medicine and Science in Sports and Exercise, 40*(7), S550 - 566.

Sandercock, Gavin, Angus, Caroline, & Barton, Joanna. (2010). Physical activity levels of children living in different built environments. *Preventive Medicine, 50*(4), 193-198. doi: DOI: 10.1016/j.ypmed.2010.01.005

Schipperijn, J., Kerr, J., Duncan, S., Madsen, T., Klinker, C. D., & Troelsen, J. (2014). Dynamic Accuracy of GPS Receivers for Use in Health Research: A Novel Method to Assess GPS Accuracy in Real-World Settings. *Front Public Health, 2,* 21. doi: 10.3389/fpubh.2014.00021

Sequeira, Sonia, Hipp, Aaron, Adlakha, Deepti, & Pless, Robert. (2013). Effectiveness of built environment interventions by season using web cameras *141st APHA Annual Meeting (November 2-November 6, 2013)*.

Silva, Thiago H, Melo, Pedro OS, Almeida, Jussara M, Salles, Juliana, & Loureiro, Antonio AF. (2012). Visualizing the invisible image of cities *Green Computing and Communications (GreenCom), 2012 IEEE International Conference on* (pp. 382-389).

Stauffer, Chris, & Grimson, W. Eric L. (2000). Learning patterns of activity using real-time tracking. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 22*(8), 747-757.

Taylor, Bronwen T. , Peter, Fernando, Adrian, E. Bauman, Anna, Williamson, Jonathan, C. Craig, & Sally, Redman. (2011). Measuring the Quality of Public Open Space Using Google Earth. *American Journal of Preventive Medicine, 40*(2), 105-112.

Taylor, John R., & Lovell, Sarah Taylor. (2012). Mapping public and private spaces of urban agriculture in Chicago through the analysis of high-resolution aerial images in Google Earth. *Landscape and Urban Planning, 108*(1), 57-70. doi: 10.1016/j.landurbplan.2012.08.001

Whyte, William Hollingsworth. (1980). *The Social Life of Small Urban Spaces.*

Wilson, Jeffrey S., & Kelly, Cheryl M. (2011). Measuring the Quality of Public Open Space Using Google Earth: A Commentary. *American Journal of Preventive Medicine, 40*(2), 276-277. doi: http://dx.doi.org/10.1016/j.amepre.2010.11.002

Wilson, Jeffrey S., Kelly, Cheryl M., Schootman, Mario, Baker, Elizabeth A., Banerjee, Aniruddha, Clennin, Morgan, & Miller, Douglas K. (2012). Assessing the Built Environment Using Omnidirectional Imagery. *American Journal of Preventive Medicine, 42*(2), 193-199. doi: 10.1016/j.amepre.2011.09.029

Xu, Zhixiang, Weinberger, Kilian Q, & Chapelle, Olivier. (2012). Distance Metric Learning for Kernel Machines. *arXiv preprint arXiv:1208.3422*.

Figure 1. Map of cameras captured by the Archive of Many Outdoor Scenes (AMOS).

Figure 2. Screenshot of an AMOS data access page.

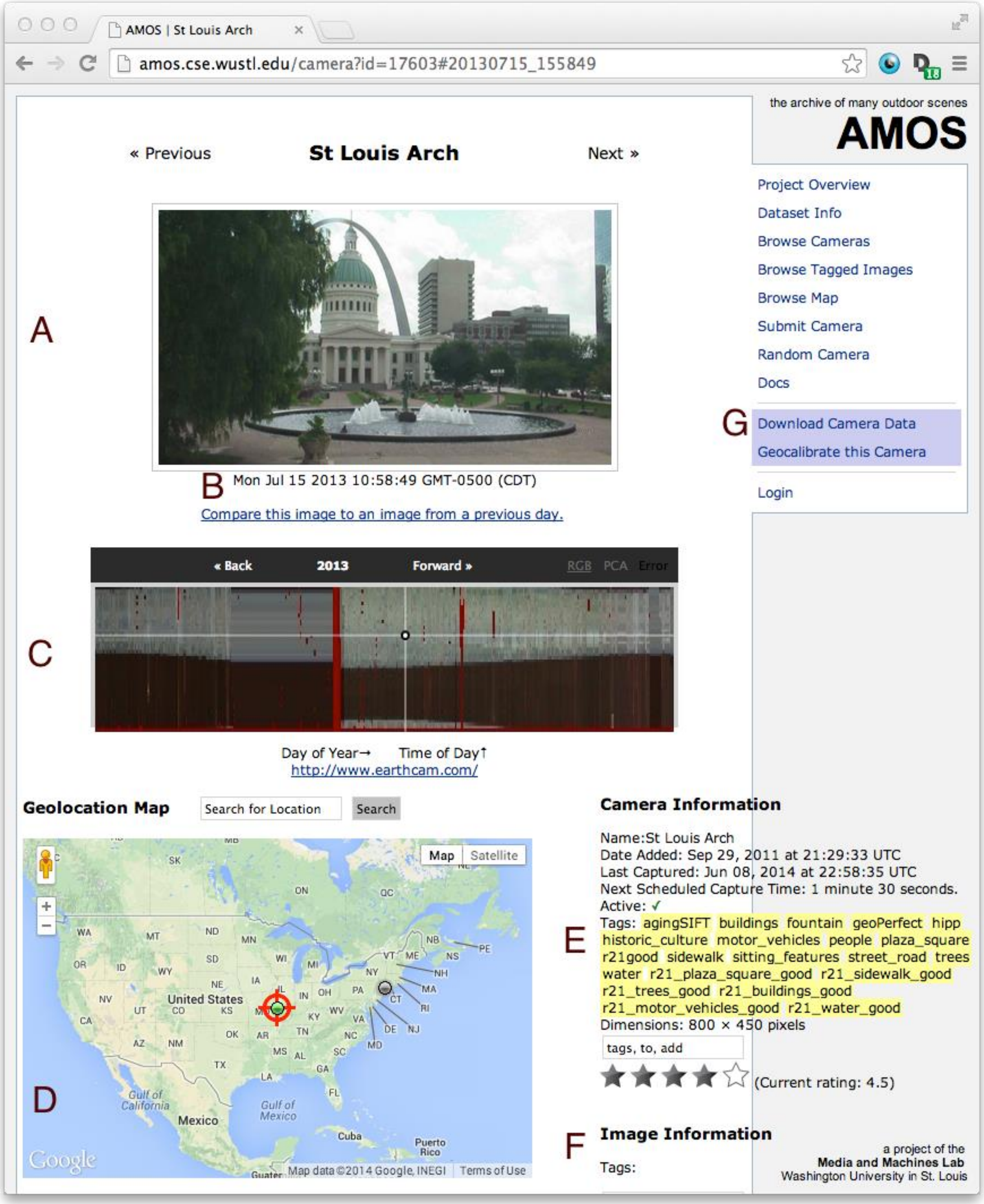

Figure 3. Reliability results for annotation of pedestrians in720 webcam scenes.

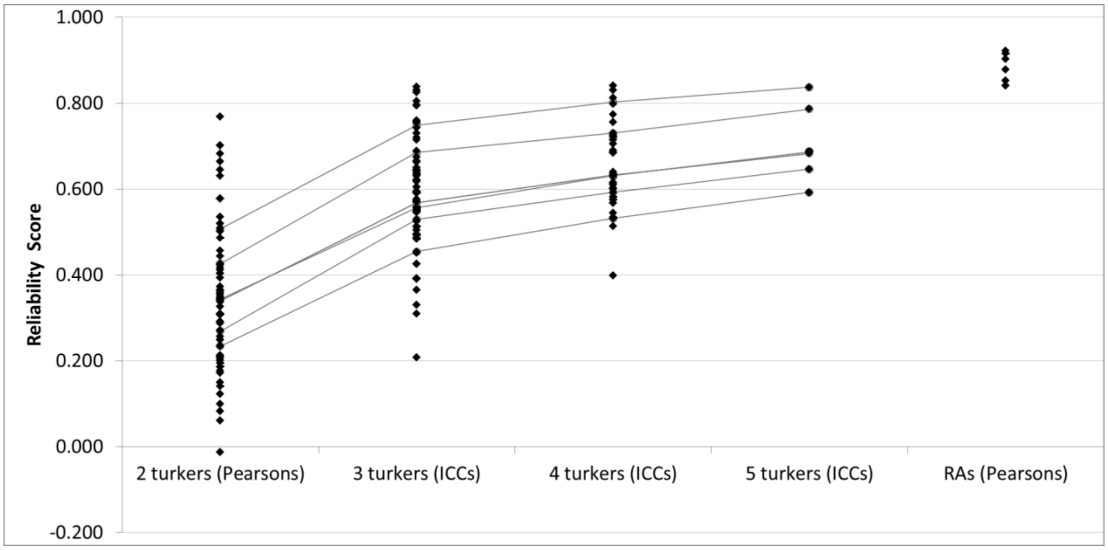

Figure 3. Location of AMOS webcams tagged with 'open space'.

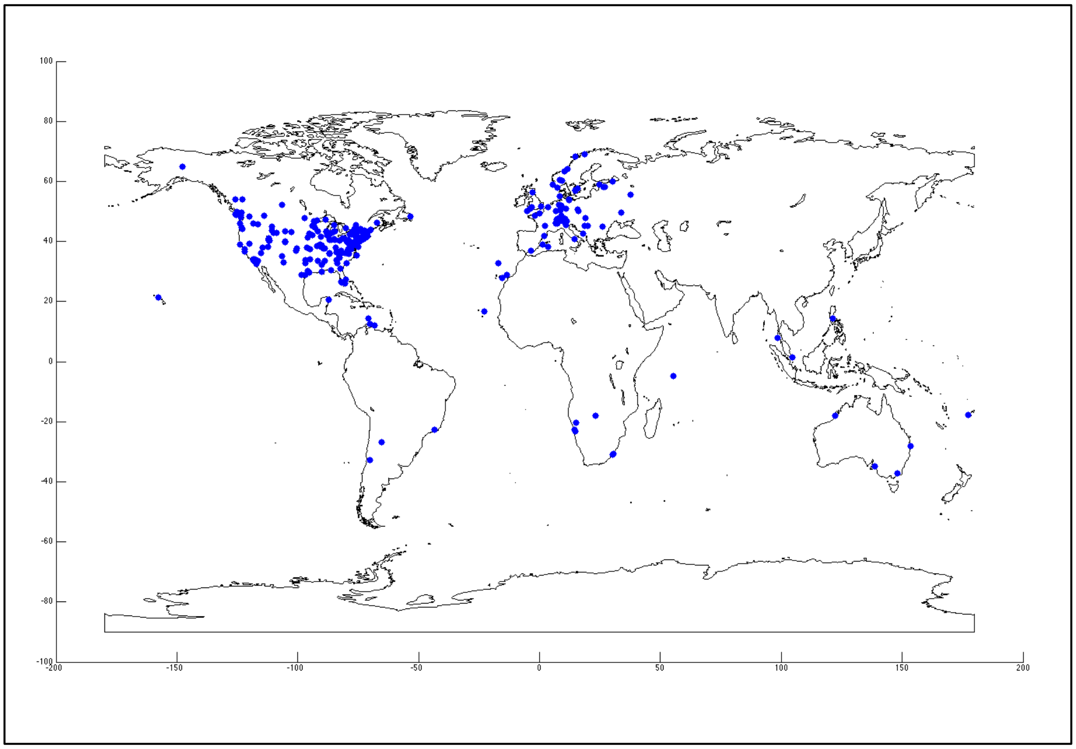

Table 1. List of Built Environment tags used to annotate AMOS webcam images.

No.	**Built Environment Tag**	**Number of Cameras with Built Environment Tag Present**
1.	Open space	769
2.	Sidewalk	825
3.	Plaza/ Square	174
4.	Residential/ Homes	97
5.	Trees	1,058
6.	Buildings	1,245
7.	Street, intersection	621
8.	Bike lane	71
9.	Athletic fields	60
10.	Speed control	185
11.	Trail path	154
12.	Street, road	1,029
13.	Signage	59
14.	Commerce Retail	382
15.	Play features	42
16.	Sitting features	166
17.	Motor vehicles	1,048
18.	Crosswalk	576
19.	Bike racks	27
20.	Water	326
21.	Snow	169

Table 2: Inter-rater reliability coefficients for trained Research Assistants and Mechanical Turk workers.

Camera ID	Counts	N	RAs Correlation	2 MTurkers Range	2 MTurkers Correlation Average	3 MTurkers ICC Range	4 MTurkers ICC Range	5 MTurkers ICC
919	Pedestrian[a]	74	.779**	.122 - .769**	.344	.330*** - .825***	.534*** - .728***	.687***
	Bicyclist[b]	74	n/a[c]	n/a[c]	n/a[c]	n/a[c]	.000- (-.019)	.000***
	Motor Vehicle[a]	74	.916**	.122 - .769**	.344	.330*** - .663***	.534*** - .721***	.687***
920	Pedestrian[a]	121	.925**	.382** - .800**	.604	.724*** - .881***	.819*** - .874***	.896***
	Bicyclist[b]	121	n/a[c]	n/a[c]	n/a[c]	.267* - .773***	.507 *** - .737***	.665***
	Motor Vehicle[a]	121	.852**	-.013 - .502**	.235	.208*** - .630***	.398*** - .614***	.592***
929	Pedestrian[a]	120	.822**	.342** - .812**	.577	.706*** - .897***	.812*** - .882***	.874***
	Bicyclist[b]	120	.524***	.425*** - .635***	.566	.706*** - .897***	.882*** - .857***	.874***
	Motor Vehicle[a]	120	.902**	.257** - .702**	.508	.605*** - .838***	.730*** - .841***	.837***
930	Pedestrian[a]	125	.900**	.659** - .860**	.784	.875*** - .944***	.917*** - .954***	.941***
	Bicyclist[b]	125	.547***	.317*** - .663***	.426	.477*** - .685***	.578*** - .756***	.704***
	Motor Vehicle[a]	125	.878**	.177* - .577**	.340	.390*** - .730***	.575*** - .685***	.683***
942	Pedestrian[a]	126	.781**	.353** - .534**	.450	.679*** - .749***	.765*** - .780***	.803***
	Bicyclist[b]	126	.602***	.281** - .764***	.552	.680*** - .868***	.783*** - .871***	.852***
	Motor Vehicle[a]	126	.841**	.099 - .411**	.269	.392*** - .617***	.544*** - .634***	.646***
996	Pedestrian[a]	128	.893**	.526** - .709**	.615	.790*** - .856***	.851*** - .877***	.889***
	Bicyclist[b]	128	.301***	.169 - .659***	.317	.437*** - .788***	.645*** - .792***	.774***
	Motor Vehicle[a]	128	.922**	.291** - .682**	.381	.619*** - .742***	.690*** - .774***	.786***

Reliability Coefficients: columns RAs through 5 MTurkers.

Significance levels: **p* < .05, ** *p* < .01, *** *p* < .001
Abbreviations: RA: Research Assistant; ICC: Intraclass Correlation Coefficients
[a] Pearson correlations were used in this calculation
[b] Kappas were used in this calculation
[c] Items could not be calculated due to insufficient variance between raters