Learning the Behavior of Users in a Public Space through Video Tracking

Wei Yan
Department of Architecture
& Computer Science Division
University of California, Berkeley
weiyan@uclink.berkeley.edu

D. A. Forsyth
Computer Science Division
University of California, Berkeley
daf@cs.berkeley.edu

Abstract

The paper describes a video tracking system that tracks and analyzes the behavioral pattern of users in a public space. We have obtained important statistical measurements about users' behavior, which can be used to evaluate architectural design in terms of human spatial behavior and model the behavior of users in public spaces. Previously, such measurements could only be obtained through costly manual processes, e.g. behavioral mapping and time-lapse filming with human examiners. Our system has automated the process of analyzing the behavior of users. The system consists of a head detector for detecting people in each single frame of the video and data association for tracking people through frames. We compared the results obtained using our system with those obtained by manual counting, for a small data set, and found the results to be fairly accurate. We then applied the system to a large-scale data set and obtained substantial statistical measurements of parameters such as the total number of users who entered the space, the total number of users who sat by a fountain, the time that each spent by the fountain, etc. These statistics allow fundamental rethinking of the way people use a public space.

1. Introduction

Modeling human spatial behavior in built environments is an area of great interest to architects and their clients. Adequate observations will reveal significant information about people's spatial preferences, which are essential considerations in architectural design [1]. 'Good' spatial behavior is an indicator of successful architectural design [2], whereas 'bad' spatial behavior can be an indicator of wasted resources and the cause for users' dissatisfaction [3]. For example, the steps depicted in Figure 1 are poorly designed, resulting in pedestrians using a narrow ramp, in the city of Zilina, Slovakia [4]. If the designers of Zilina city center could have predicted this pattern of behavior, they

Figure 1. Pedestrians prefer the narrow ramp to the shallow stairs Zilina, Slovakia [4]

would likely have chosen a different design based on a more appropriate design norm. However, without a better understanding of human spatial behavior, appropriate design norms are difficult to set up.

Behavioral mapping [5] —a manual procedure of recording detailed observations of who does what where for a specified time interval — has been used to derive data from the field. However, it is costly to collect and analyze these manual observations. Time-lapse filming was use by William Whyte [6] and others. However, manual analysis of the pictures is enormously time consuming and tedious. William Whyte used over 100 man-hours in front of the film viewer to evaluate six and a half hours of video.

We demonstrate the application of computer vision to collect and analyze mass data of human behavior in a public space through video tracking techniques. We describe a system used to analyze behavioral data collected in a roughly square subspace in the center of Sproul Plaza at the University of California, Berkeley. The space contains distinctive paving, a fountain with low seating edge, large area of steps, and a few benches (see Figure 2). The results of the analysis, which are statistical measurements, can be used for learning and modeling users' behavior in public spaces. For example, it will be used for validating and calibrating a user model in a public space, which is

currently under development [3]. The major measurements we are interested in are listed below:

How many people entered the plaza within the specified time interval? This is an indicator of the level of service of the plaza – how many people were using the space in the time interval.

How many people came to sit by the plaza fountain, on the steps, or on the benches, respectively? Because a good public space attracts more passengers to stop, sit, and look around, we can evaluate whether the design of the space is successful or not. These measurements also help architects know what type of seating is more preferred by users.

What is the probability that a user who entered the plaza chose to sit by the fountain, on the steps, or on the benches, respectively? This will be a property of a user model that will affect a user's decision of whether to sit or pass by, and if to sit, where to sit.

What is the probability that a person chose to sit by the fountain depending on the number of people who were already sitting there? The probability will tell us how others affect a person's behavior. People like to watch other people. It is generally true that in public space, people are attracted to other people and seek to place themselves near others [4]. Our measurement will help us get the quantitative proof about this qualitative statement and show how significant the effect is.

How long did each person sit in a place and what is the distribution of their time of stay? Time of stay is an important indicator of how people like a public space.

Where did people walk? This will reveal the spatial preferences of users' walking behavior and will reflect the settings and functionalities of the surrounding architectural environment.

How many people wandered between the fountain and places around? People who like the space more than others will spend more time around the fountain, chatting, eating, and taking pictures there.

The measurements will help us get a quantitative assessment of many generally believed qualitative design rules. The main focus of research in tracking humans in outdoor environments has been (a) identifying humans (b) clustering behaviors from tracking geometries and (c) finding unusual behaviors (e.g. [7], [8], and [9]). Our focus is on accurate measurement of particular

going to apply the results to model the behavior of users in public spaces and evaluate architectural design. Our tracking system consists of two components being a head detector and data association. Our system can track people fairly accurately and efficiently (section 2). In section 3, we show how to obtain statistical measurements of the relevant behaviors. Section 4 demonstrates that these measurements are reasonably reliable, and shows parameters extracted from a large scale data set. In particular, we have validated our conclusions by 1) comparing the results obtained by our system, for a small data set, with those obtained by manual counting; and 2) choosing some results from a one hour data set and comparing them with manual counting.

2. Description of the tracking system

In order to obtain the statistical measurements we need, we tracked (a) people sitting by the fountain, on the steps, or on the benches, respectively, and (b) people walking and wandering in the plaza.

Heads are good for detecting and tracking people because they can be easily and reliably detected. Their distinctive appearance – dark color in the video – can help us detect and link them across time. Our approach is to detect heads in each frame and link heads from frame to frame through data association. We manually set up target regions for tracking each behavior. For example, we set the fountain area to be a target region to track people sitting there (Figure 2a), and the area excluding the fountain to be a target region to track people walking in the plaza (Figure 2b). We use background subtraction and intensity thresholding to detect people's heads and link them from frame to frame. We compared the results of our program count with a manual count for a small data set and found that they are reasonably close (section 4.1). Finally we applied the tracking method to 8 hours of video and obtained the measurements we need (section 4.2).

Background models. Because our tracking system does not need to be real time, we can build background models by manually selecting video clips in which there were few people in the plaza and then averaging the images in those clips. The reasons for using such a straightforward method to model the background are twofold: 1) the background did not change much in terms of lighting during the two hours of each day when we

(a)

(b)

Figure 2. Sproul Plaza and target regions, which are highlighted for tracking people by the fountain (a), and in the plaza (b), respectively.

behavioral parameters, specified in advance, because we are

took the video, so our static background model can work; 2) the

effect caused by people sitting in the background for long periods makes it difficult to apply adaptive background models (such as those of [7]), but our manually created background model can eliminate that effect. Because our viewpoints in the video were slightly different from day to day, we have built different background models for different days.

2.1 Detector

In the video, most people's heads and many people's upper bodies are dark. This means we can detect people accurately and fairly simply by detecting heads and/or upper bodies. The dark pixels can be separated from the background by intensity thresholding and then grouped into blobs for tracking. We apply intensity thresholding to an image after background subtraction. Thresholded regions are grouped into blobs. This approach is robust because one usually has either a dark head or a dark upper body, and because we have avoided difficulties in identifying whole bodies for people walking or sitting in groups. In particular, because heads are small, one has a good chance that counting heads gives an accurate count of people in a group, because their heads tend not to overlap. Detecting dark pixels that represent heads and/or upper bodies has another advantage in detecting people by the fountain. Because the running fountain water is always bright, our approach can eliminate most of the disturbance coming from the fountain water, which would otherwise need extra work to be modeled as background.

2.2 Data association

We need to track people over frames in order to obtain some measurements, such as how long each person sits by the fountain. Without knowing who in frame i+1 corresponds to who in frame i, we cannot obtain the measurements we need. For example, considering estimates of how long people stay by the fountain, without estimates of the true correspondence we must average over all possible correspondences. This can be done (even weighting with possible costs [10]), but it leads to unacceptably heavily smoothed estimates for our problem. We applied the Hungarian Algorithm to link blobs from frame to frame, where the entries of the cost matrix are distances between people's location on the two frames. For a blob in frame i+1 that is not found corresponding to any blob in frame i, if it is close to the border of the target region, it is regarded as a new person arriving; otherwise it will be ignored as a noise point. If a blob in frame i has no correspondence in frame i+1, we create a virtual blob whose position is extrapolated from the real blob's location and velocity. Virtual blobs can be linked to real blobs, meaning that short drop-outs do not perturb our track. Virtual blobs can be propagated forward by more than one step, but are eventually reaped if they do not match a real blob within an acceptable number of steps – we interpret this as a person leaving. Identifying people who are crossing is solved by assuming that people normally keep their speed and direction when they walk and therefore applying velocity as a measurement to the data association.

3. Tracking people in the video

3.1 People by the fountain

We track people by the fountain, on the steps, or on the benches with similar approaches, therefore we only describe in detail our approach to track people by the fountain. We want to obtain the following measurements: How many people came to sit or stand by the fountain during the specified time interval? How long did each person stay by the fountain and what is the distribution of time of stay? What is the probability that a person chose to sit by the fountain depending on the number of people who were already sitting there?

We first set up an oval area as the target region on top of the fountain (Figure 2a), which is intended to include only heads and upper bodies of people, in order to implement the head detecting approach. We exclude the region containing people's lower bodies and legs to eliminate repeated counts caused by their colors, which may contain dark pixels. Then we detect people in each single frame and link people over frames. The steps of detecting people by the fountain in a single frame are described below.

On top of an original image we create a target region (Figure 3a). We do normal background subtraction, binary thresholding, size thresholding, and get the foreground blob image (Figure 3b). We now multiply this mask with the original image and get a foreground image (Figure 3c). We then do intensity thresholding

Figure 3. The steps of detecting people's heads and upper bodies in a single frame. Although it is a fairly complicated scene, where many people overlap, the detector finds 8 persons, while the result of a manual count is 10 approximately – it's hard even for a manual count to have an accurate result in this case. (a) People by the fountain and the target region is highlighted in the oval. (b) Blobs after background subtraction. (c) Foreground image. (d) Blobs after intensity thresholding and merging closely located blobs.
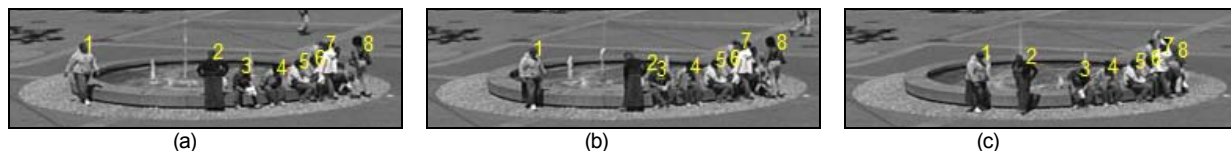

Figure 4. Image sequence shows linking people by the fountain over frames (a), (b), and (c). IDs of people are superimposed on the scene images. Note that ID 2 was labeled poorly on frame (b), but was recovered by frame (c). By linking people over frames we can tell how long people stay by the fountain (Figure 8).

on the foreground image and merge closely located blobs, which are the head and upper body of the same person, and we get the blobs representing people's heads and upper bodies (Figure 3d). Note that the two blobs of water shown on the left side in Figure 3c are now eliminated.

After obtaining the coordinates of blobs representing people in each frame, we can start to link people from frame to frame, using the Hungarian Algorithm, where the entries of the cost matrix are distances between people's location in the two frames. Figure 4 shows an image sequence of linking people by the fountain. People's IDs are superimposed on the scene images. From the IDs we know the result is good: each person was labeled with the same ID in different frames. Note that ID #2 was labeled poorly on frame (b), but was recovered by frame (c). After linking people over frames, we count people that have appeared by the fountain longer than a preset time threshold and regard them as people who stayed by the fountain.

3.2 People in the plaza

We track people in the plaza in order to obtain the following measurements: How many people entered the plaza within the specified time interval? Where did people walk? What are the paths? How many people wandered between the fountain and the places around? With the measurements from tracking people both by the fountain and in the plaza, we can get the probability that people entered the plaza chose to stay by the fountain.

The process of detecting people in the plaza is similar to that by the fountain. We first set up a target region in the plaza as highlighted in Figure 2b. Then we detect people in each single frame. The difference is that people move faster and in a larger space in the plaza than by the fountain, so the probability of people overlapping or connecting is reduced. We can simply use blob detection without intensity thresholding to detect most people walking in the plaza. However, it is common for blobs to be linked by cast shadows at foot level; since we have enough geometric context to know an up-vector, we know where the feet are and can avoid this problem by cutting off the lower 2/3 of each blob. Figure 5 show how this works. This does not work for large blobs, which might be formed by several people walking together. For blobs of sufficiently large area we instead threshold intensity for the blob. Figure 6 shows how this works.

To link people walking in the plaza from frame to frame, we used similar methods as that in linking people by the fountain. We assume that people normally keep their speed and direction when they walk so we use people's velocity as the measurement to help link people. This is a particularly useful trick (Figure 7).

4. Results

We collected 8 hours of video data in Sproul Plaza in 4 days, from 3PM to 5PM on each day, during the Berkeley summer sessions 2003. While we were taking the video in the plaza, we also studied users' behavior through direct observations, which proved helpful later when validating our system. The original video has the resolution of 720x480 pixels with RGB mode, and 30 fps frame rate. We actually use gray scale images at 10 fps converted from the video.

4.1 Comparing program results with manual count and observations

We tested our methods of detecting and data association with a 6 minute (10 fps) clip of video and compared the results with a manual count. We obtained reasonable results compared with the manual count in both detecting people in a single frame and
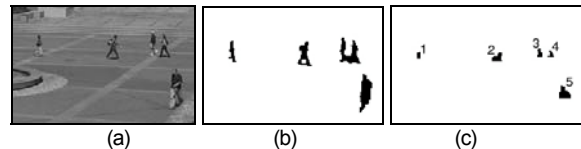

Figure 5. Frame (a) shows five people walking across Sproul Plaza. Simple background subtraction as in (b), results in an undercount because blobs are connected by cast shadows. But the geometric context tells us that this is likely to occur at legs, which are low on a blob. This suggests cutting the lower 2/3 of a blob, as in (c), resulting in the correct count of five.
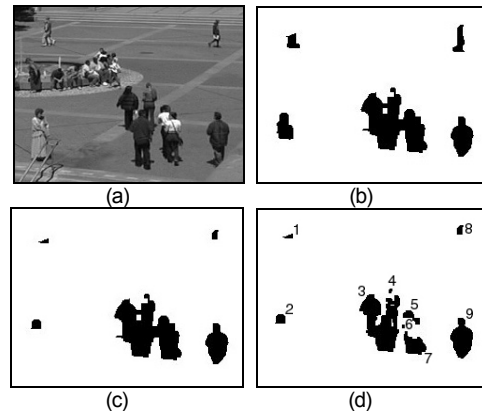

Figure 6. Frame (a) shows 9 people walking across Sproul Plaza. Simple background subtraction as in (b), results in an undercount because blobs are linked when many people overlap. We cut the lower 2/3 of small blobs for the same purpose shown in Figure 5, but leave a large blob uncut, as in (c). Intensity thresholding for the large blob splits it into several blobs, as in (d). Although the blob-people match is not perfect, we have an improved match and a good count of 9.

Figure 7. Image sequence shows linking people walking in the plaza over frames (a) and (b). People's IDs are superimposed on the scene images. Note that the two persons with ID 4 and 5 were crossing in (a). Using velocity extrapolated, ID 4 and 5 were labeled correctly after they crossed in (b). The images show the data association succeeded.

identifying people over frames.

Sitting people. We count the number of people who entered and stayed in the fountain area for more than 30 seconds. A manual count gives 26 and the program counts 32. Because the

scene is complicated and people walk in and out of the fountain area quite often, it's hard to have an accurate manual count and this degree of agreement is reasonable.

We also calculated the distribution of people's time of stay by the fountain and compared it with manual count (Figure 8). This is particularly difficult to do manually, as one must run video back and forth. The comparison between manual and
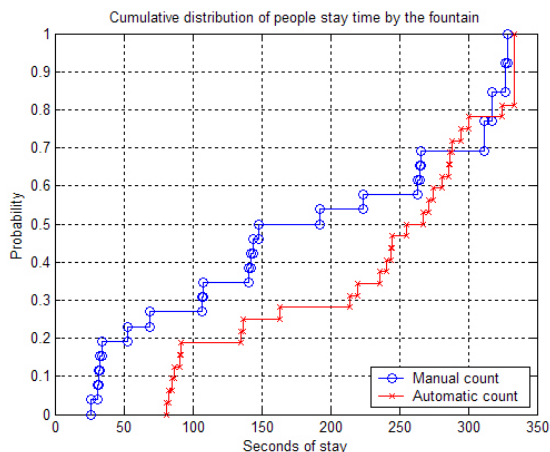

Figure 8. The cumulative distribution of times of stay at the fountain, measured by hand (o's) and automatically (x's). Note that the overall structure of the CDF's is fairly similar, with most people staying for either a minute or so or of the order of five minutes. Given the difficulty in obtaining an accurate manual count and the relatively short period (approximately six minutes) the two compare rather well.

automatic counts appears to be reasonable.

A further confirmation is that our program identifies one person who sits on the steps for a period of 20 minutes in the main dataset. We were able to confirm by hand that this, in fact does occur.

Walking people. We tracked people walking in the plaza. The paths in Figure 9 show the behavior pattern of walkers in the plaza, similar to what we have found through observation in the plaza, e.g. most people passed the plaza between the steps of the fountain and the Sproul Hall, which is an office building in Berkeley campus. The people's walking paths also show some interesting behavior pattern of people, e.g. people walking from the north of the plaza to the stair on the west tended to pass a small area at the end of the stair rail more often than we expected. But both our observation and the concrete wear on the pavement in the plaza have confirmed this behavior pattern (Figure 9).

Wandering people. We also tracked people wandering around the fountain. We define wandering people as those people whose starting points are close to the fountain and spent more than 1 minute in the plaza. Only if a newly tracked blob representing people is close enough to the fountain did we treat it as a wandering person. Figure 10 shows the paths of tracked wandering people in the test. The result shows that many people wandered between the fountain and Sproul Hall, which agrees with our observations in the plaza – many people took pictures in front of Sproul Hall.

Figure 9. The result of a 6-minute tracking of walking people in the plaza. The walkers' paths are superimposed on an image of the scene. The paths show the behavior pattern of walkers in the plaza, similar to what we have found through observation in the plaza, e.g. most of the people passed the plaza between the steps of the fountain and the Sproul Hall, which is an office building on Berkeley campus. The people's walking paths also showed some interesting behavior pattern of people, e.g. people walking from the north of the plaza to the stair on the west tended to pass a small area at the end of the stair rail more often than we expected. But both our direct observation in the plaza and the concrete wear on the pavement in the plaza have confirmed this behavior pattern (the arrow on the image shows the wear pattern in the concrete).

Figure 10. Paths of tracked wandering people are superimposed on the scene. People's starting points are marked as circles.

4.2 Large-scale tracking results

All the above material suggests that predictions made by automatic tracking are in line with those made by a manual analysis. We have run the tracking system with 8 hours of video – 2 hours each for four days, a volume too large for manual evaluation -- and obtained the following results pertaining to the behavior of users of this space:

The total number of people who entered the target region in the plaza (Figure 2 b) during the 8 hours is 10257. Figure 11 shows the numbers of people on each day. Because of the configuration of our target regions, when we count people entering the plaza, we may have missed a few people, including: 1) those who came from the north side of the fountain and left without entering both the fountain and the region we count for walkers; and 2) those who were by the fountain at the beginning.

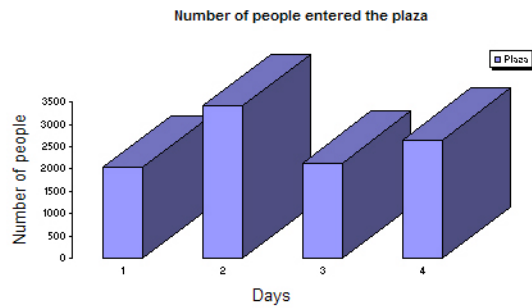We know both cases to be uncommon, so the total people entered the plaza is 10257+n, where n is a small number.

Figure 11. Numbers of people entered the plaza in different days.

Total numbers of people who sat in different places during the 8 hours are shown in Table 1, and numbers of people in each day are in Figure 12. We can see that most people sat by the fountain. This is consistent with our observations in the plaza.

Table 1. Total number of people sat in different places

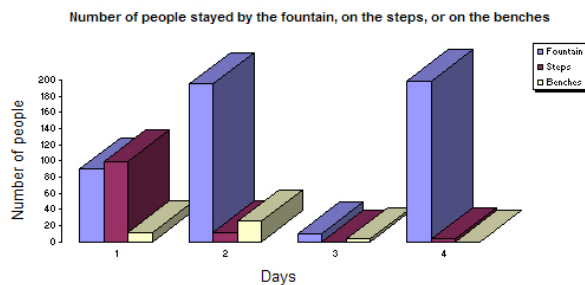|	Fountain	Steps	Benches
Total people	494	113	39

Figure 12. Numbers of people sitting in different places in different days. From the chart, we see that most people sat by the fountain. This is consistent with our observations in the plaza.

Probabilities of people who entered the plaza chose to sit in different places were calculated by dividing the numbers of sitters on each type of seating by the total number of walkers entering the plaza, which is 10257+n, where n is a small number. The probabilities are shown in Table 2.

Table 2. Probabilities of people entering the plaza chose to sit in different places. Note that about 1 in 20 people sat by the fountain, which is much more popular than the steps or the benches.

	Fountain	Steps	Benches	Total
Prob	494/(10257+n) approx 5%	113/(10257+n) approx 1%	39/(10257+n) approx 0.4%	6.4%

Probability that a person chose to sit by the fountain depending on the number of people already sitting there. From Figures 13 and 14, we can see that people tend to sit when there are other people already sit there. This confirms the belief that in public space, people are attracted to other people and seek to place themselves near others [4].
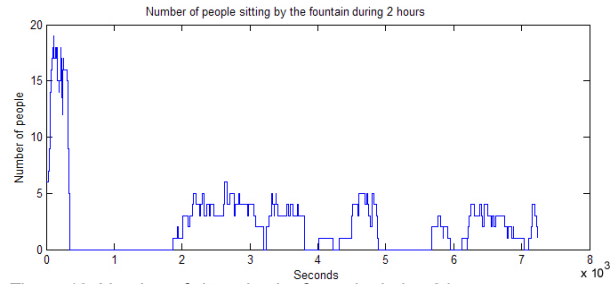

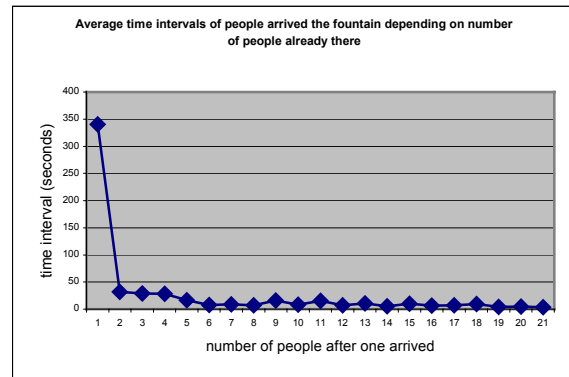Figure 13. Number of sitters by the fountain during 2 hours.

Figure 14. A plot of the average time interval between the arrival of the i-1'th person and the arrival of the i'th person against i, measured for two 2 hour days; we omit two of our days of data here because a manual evaluation confirmed people were behaving unusually because the days were unusually hot. In particular, the fountain has no people at it for intervals of approximately 6 minutes on average. Once the first person has arrived, on average it takes under a minute for the second one to turn up, and numbers three and four follow shortly. The tail of the graph is untrustworthy due to there being relatively few instances.

Distribution of time of stay is shown in Figure 15. Although more people sat by the fountain than on the benches and steps, people stayed for much longer time on the steps than by the fountain or on the benches.
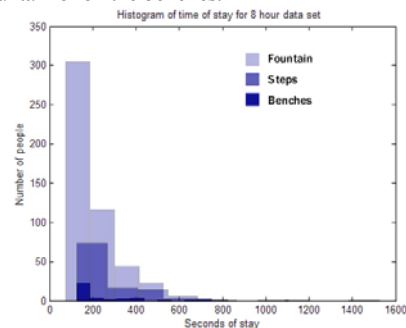

Figure 15. Distribution of time of stay of people by the fountain, on the steps, or on the benches. Although more people sit by the fountain than on the benches and steps, from this histogram of distribution of time of stay, we see that people stayed for much longer time on the steps than by the fountain or on the benches. For example, a person sit on steps for about 20 minutes and this result is confirmed by our manual count.

People's walking paths are shown in Figure 16. The paths suggest that the whole Sproul plaza is more of a street than a square, where most people walked between the south side and

the heart of campus. Paths also show the interesting behavior patterns e.g. people walking from the north of the plaza to the stair on the west tended to pass a small area at the end of the stair rail more often than we expected. But both our direct observation in the plaza and the concrete wear on the pavement in the plaza have confirmed this behavior pattern (see the arrow on Figure 9).
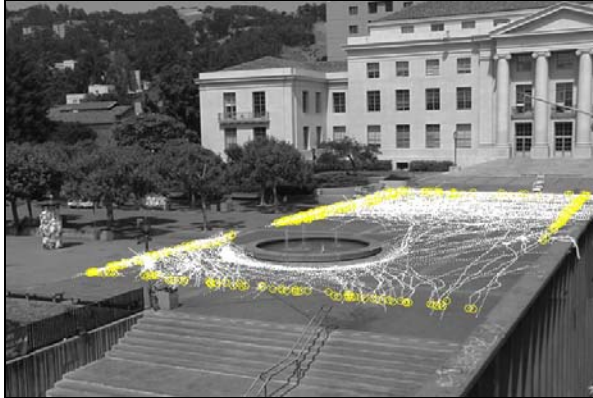

Figure 16. People's walking paths of 20 minutes on a scene image. The paths show us clearly the walking behavior pattern: most people walked between the south side and the heart of campus, which is shown by a large number of horizontal paths. We also have checked the walking paths for the 4 day data set and they all look similar, which means the walking behavioral pattern of users is constant in the data set.

Total number of people wandered between the fountain and places around is 132 during the 8 hours. Figure 17 shows the numbers of wandering people on each day. The paths of wandering people are shown in Figure 18. We tracked walking people who started from the fountain and spent more than 1 minute in the plaza before leaving and we defined them as wandering people. Our observations tell us some people normally like to look around and enjoy the space and many of them like to take pictures in front of Sproul Hall's impressive entrance. Our tracking results confirmed this by showing some people wandered between Sproul Hall and the fountain.
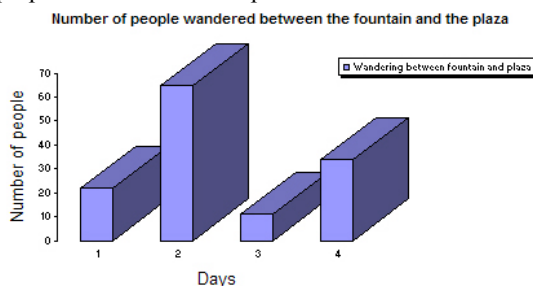

Figure 17. Numbers of people wandered between the fountain and the plaza.

Results showing abnormal behavior patterns are depicted in Figure 12. Some of our video clips were taken on two of the hottest days in the summer. The results of the video tracking showed us some abnormal users' behavior patterns in those two days. In one of them, there were almost no people sitting by the fountain (Figure 12, day 3). It was simply so hot that no one wished to stay in the sun. On the other day, there were a large number of people who stayed at the fountain (Figure 12, day 4). By manual observation, these were teenagers having an event in the plaza. They were playing and wading in the fountain water for quite a long time (Figure 19).
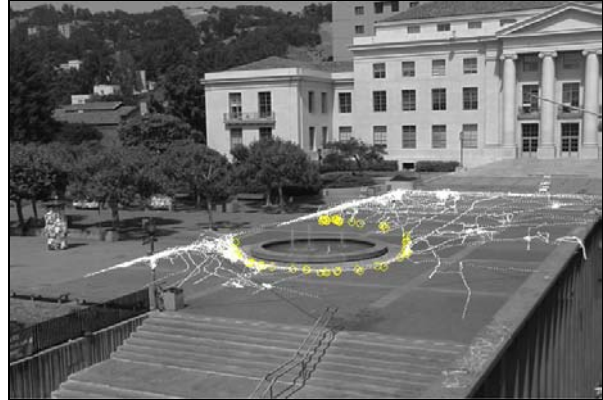

Figure 18. People's wandering paths on a scene image. We tracked walking people starting from the fountain and spent more than 1 minute in the plaza before leaving as wandering people. Our direct observations in the plaza tell us some people normally like to look around and enjoy the space and many of them like to take pictures in front of Sproul Hall's impressive entrance. Our tracking results confirmed this by showing some people wandered between Sproul Hall and the fountain.

Figure 19. On a very hot day, many people were wading and playing in the water. We do not use data from this day to estimate arrival probabilities at the fountain, because this behavior is unusual.

5 Conclusions: what do people do in public?

From the results of analyzing the large-scale data, we have some important findings about users' spatial behavior in the public space.

A total of 6.4% of the people who entered the plaza chose to stay by the fountain, on the benches, or on the steps for a while. The number is high, if we consider students' tight schedules in the summer sessions at Berkeley. It shows us the use of the plaza is quite different from what many people think about the plaza. The target areas we are interested in are at the center of the whole Sproul Plaza. The whole plaza is a linear design creating promenade for students who walk between the south side of campus, where most dormitories and other student housing is located, and the libraries and classroom buildings in the heart of campus. The plaza makes many people think it is simply a pedestrian thoroughfare, except when some events happen occasionally and turn the plaza into a crowd-gathering place. From our tracking results we know that it is used on a daily basis

by many people – approximately one a minute – who want to stay for a while, sitting by the fountain, on the steps or on the benches. This suggests the plaza is, indeed, a good public space, and attracts many people to come, stop, sit and look around.

Although architects normally regard benches as primary seating, our tracking results show us in the plaza the secondary seating, which includes the fountain and steps, are more popular, especially the fountain, which is dominantly more attractive to people who want to sit than the benches and steps are. The fountain became a magnet that attracts people to linger in the plaza space and encourages social interaction. This result helps architects know the importance of putting a fountain into a plaza. An architectural explanation to this is that the visual and aural attraction of moving water in the fountain is universal. A "noisy" fountain located close to seating may successfully screen out surrounding traffic noises, help immeasurably in creating pleasant ambience and producing stress-reducing effects [11].

From Figure 13 and 14, we can see that people tend to sit when there are other people already sit there. This confirms the belief that in public space, people are attracted to other people and seek to place themselves near others [4].

People's walking paths also showed some interesting behavioral patterns of people, e.g. people walking from the north of the plaza to the stair on the west tended to pass a small area at the end of the stair rail more often than we expected. But both our direct observation in the plaza and the concrete wear on the pavement have confirmed this result. This phenomenon may suggest a human spatial behavior rule: walkers with a destination in mind tend to keep their direction until they have to change.

The results from the two hot days tell us that the use of a public space is affected greatly by the climatic variation, but design can still help provide outdoor spaces for people to gather during tough weather. For example, a fountain in the plaza can still attract many people in even the hottest day.

Finally, we can conclude that computer vision techniques are likely to be remarkably effective at producing a statistical picture of how people behave in public which, though rough, can be an important source of detailed guidelines for architects. In particular, we have been able to obtain parameters from a large body of data with much greater efficiency than was previously possible.

6. References

[1] D.C. Miles, C.R.S., et al., *Plazas for people*, Project for Public Spaces, Inc., 1978.

[2] E. Steinfeld, "Toward artificial users", *Evaluating and predicting design performance*. Y. E. Kalay, ed., John Wiley & Sons, New York, 1992.

[3] W. Yan and Y. E. Kalay, "Simulating the Behavior of Users in Built Environments", *Journal of Architectural and Planning Research*, 2003.

[4] J. Gehl, *Life between buildings: using public space*, Arkitektens Forlag, Copenhagen, 1996.

[5] W.H. Ittelson, New York (City), City University of New York, Center for Research in Cognition and Affect, et al., *Environment and cognition*, Seminar Press, New York, 1973.

[6] W. Whyte, *The Social Life of Small Urban Spaces*, Project for Public Spaces, Inc., 1980.

[7] C. Stauffer, and W.E.L Grimson, "Learning Patterns of Activity Using Real-Time Tracking", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8), pages 747-757, 2000.

[8] I. Haritaoglu, D.Harwood, and L.Davis, "Who, when, where, what: A real time system for detecting and tracking people", *Proceedings of the Third Face and Gesture Recognition Conference*, pages 222-227, 1998.

[9] H. Zhong and J. Shi, "Finding (Un)Usual Events in Video", *tech. report CMU-RI-TR-03-05*, Robotics Institute, Carnegie Mellon University, May, 2003.

[10] T. Huang and S. J. Russell, "Object Identification: A Bayesian Analysis with Application to Traffic Surveillance", *Artificial Intelligence*, 103(1-2), pages 77-93, 1998.

[11] C. C. Marcus and T. Wischemann, "Campus outdoor spaces", *People places: design guidelines for urban open space*, C. C. Marcus and C. Francis, ed., Van Nostrand Reinhold, New York, 1990.