

A Novel Visual Tracking Approach Incorporating Global Positioning System in a Ubiquitous Camera Environment

Hsien-Chou Liao and Pao-Tang Chu

Department of Computer Science and Information Engineering,
Chaoyang University of Technology, Taiwan, Republic of China

Abstract: Global Positioning System (GPS) is popularly used in object tracking. Current GPS tracking system is mainly based on a digital map of a GIS (geographic information system) and marks the location of an object on the map. However, it is insufficient to perceive the situation of the object. On the other hand, the deployment of cameras is toward a ubiquitous camera (UbiCam) environment, especially in the urban area. If GPS can be incorporated with UbiCam environment, the visual information acquired from a camera is useful to improve the problem of current GPS tracking systems. Therefore, an approach called GODTA (GPS-based Object Detection and Tracking Approach) is proposed in this paper. GODTA is mainly based on the coordinate transformation method. The transformation is according to the parameters computed from a set of five calibration points. If the transformed image coordinate is within the rectangle range of the camera image, it means the object is within the FOV (field-of-view) of the camera, and then the object can be labeled on the image. Two experiments, fixed position and continuous tracking, were also designed to evaluate the performance of GODTA. The lowest average locating error (*LE*) of the transformed image coordinate is 42 and 38 pixels, respectively. The results show that GODTA is feasible for fulfilling GPS-based visual tracking service and incorporates GPS system and UbiCam environment successfully.

Key Words: Location-based service, object detection, object tracking, coordinate transformation, mobile computing

INTRODUCTION

The evolution of hardware and information technology achieves the digitalization of the surveillance systems. The real-time image or recorded video can be transmitted to any place or device via Internet connection. It enhances the feasibility and efficiency of surveillance systems. This causes the deployment of cameras towards a ubiquitous camera (UbiCam) environment. For example, the area of the Taichung County, Taiwan, is only 163.4 square meters and the population is about 960,000. The number of cameras deployed on the streets is more than 3,000. The UbiCam environment causes the visual surveillance to become an emergent application domain. The moving object detection and tracking are important tasks for visual surveillance. Many researchers have proposed methods for such tasks. For example, Shah et al. (2007) proposed a system, called Knight, for the detection and classification of moving objects. Haritaoglu et al. (2000) proposed a W^4 system. It can detect, track, and identify the activity of a moving object, including talking, walking, etc. Some related studies include Ridder et al. (1995), Stauffer et al. (2000), Chang and Gong (2001), Black et al. (2002), Cucchiara et al. (2003), Micheloni et al. (2005), Hu et al. (2006), and Liao and Cho (2008).

These approaches rely on an image processing technique for object detection and tracking. However, many physical conditions must be considered, such as cluttered background, change in illumination, and occlusion of objects. More issues are encountered for the UbiCam environment, such as spatial relationship establishment, consistent labeling, and camera handoff.

On the other hand, GPS is the most popular positioning system in an outdoor environment since the removal of signal-degrading selective availability (SA) from GPS signals on May 1, 2000. A commercial differential GPS (DPGS) can even provide accuracy to less than a few meters. The application domains of GPS are very broad, such as navigation, vehicle security, positioning for observation of children or elderly relatives, etc. Many products are also developed for children, teens, adults, or elder relatives, such as the GPS phone, GPS jacket, or LiveWire-X5 (real-time GPS tracking device). They are GPS-enabled mobile devices, and also an important trend in recent years. Such devices enable the GPS to make the coordinates of a moving object easily available. For example, Dawoud et al. (2008) proposed a system using a GPS-enabled mobile device with built-in camera. The real-time GPS coordinate and camera image are used to match with virtual images stored in a database in order to compute the accurate position of the moving object.

However, current GPS tracking systems are mainly based on a digital map. It is insufficient to realize the situation of an object simply from its location or trace information. If the GPS-enabled mobile devices could be incorporated with cameras, the real-time image of an object passing through the FOVs (field-of-view) of the cameras would be useful for situation awareness.

According to the trend of GPS-enabled mobile devices and the UbiCam environment, an approach, called GODTA (GPS-based Object Detection and Tracking Approach), is proposed in this paper. The purpose of GODTA is to incorporate GPS system and the cameras, and provides a novel GPS-based visual tracking service. When an object carried a GPS-enabled device moves into the FOV of a camera, the object can be marked and seen on the real-time image. GODTA is based on a 3D-coordinate transformation method. A GPS coordinate of a moving object is transformed to the image coordinate of a calibrated camera. Then, the image coordinate can be used to mark the moving object directly on the image. The performance of the GODTA is dependant on the accuracy of the GPS receiver that is influenced by its chipset and the skyview situation. In order to realize the feasibility of GODTA in the practical environment, two experiments, fixed position and continuous tracking, were designed for two types of GPS receivers and two kinds of skyview situations. A prototype system was also implemented for the experiments. The results show that GODTA is feasible in providing a novel visual tracking service for the UbiCam environment.

APPROACH

A moving object is introduced to carry a GPS-enabled mobile device with the wireless

communication capability. The real-time GPS coordinates of the object can be updated periodically. The purpose of GODTA is to detect and track the moving object in the outdoor environment using multiple cameras. Such a scenario and the model of GODTA are shown in Fig. 1. GODTA consists of several modules. The camera calibration module is used to generate the desired parameters for the coordinate transformation. The GPS coordinate transformation module is responsible for the transformation from the GPS coordinate to the image coordinate of a calibrated camera. A camera detection and object tracking module is responsible for determining which camera image contains the moving object. Then, the moving object is labeled on a 2D map and camera image.

The process of GODTA is separated into calibration and operation phases, as shown in Fig. 2. The DMS (degrees-minutes-seconds) format of the GPS latitude/longitude coordinate is a 3D position on

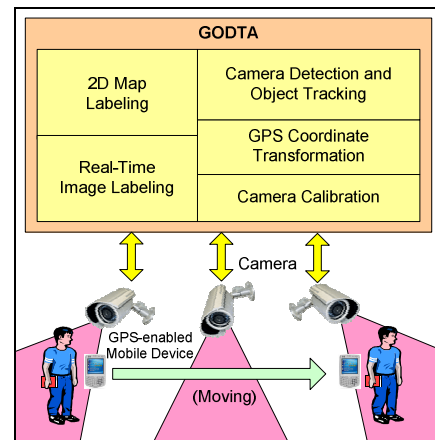


Fig. 1. The scenario and model of GODTA

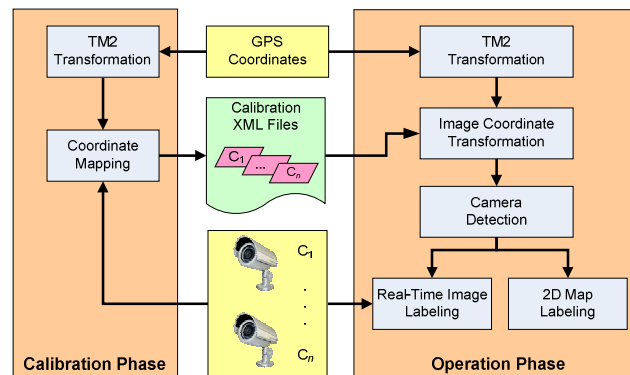


Fig. 2. The GODTA process

the Earth's surface. The coordinate is not a Cartesian coordinate and cannot be used for the image coordinate transformation in the later step. Therefore, it is transformed to a Cartesian coordinate on a 2D plane by using a map projection. A TM2 (2-degree Transverse Mercator) coordinate is the most popular one used for map projection and thus is used in GODTA. Then, a set of calibration points is established and stored in a XML file.

After all the cameras are calibrated, GODTA can start the operation phase. A real-time GPS coordinate of a moving object is transformed to a TM2 coordinate. The TM2 coordinate is then transformed to the image coordinate via the parameters generated in the calibration phase. When the generated coordinate is within the rectangle range of a camera image, it means the moving object is within the FOV of the camera. Then, the image coordinate is used to label the moving object on the real-time image.

TM2 transformation: The transformation is presented as follows. The coordinate acquired from the GPS receiver is in a DMS format. It is transformed to the format of the WGS84 (World Geodetic System 1984) coordinate system. The WGS84 coordinate is then transformed to a TM2 (2-degree Transverse Mercator) coordinate. Assuming that the WGS84 coordinate transformed from the DMS coordinate is represented as (ϕ, λ) and the TM2 coordinate is represented as (x, y) , the transformation occurs according to Eq. 1:

$$\begin{aligned}
 x &= K_0 N \left[A + \frac{(1-T+C)A^3}{6} + \frac{(5-18T+T^2+72C-58e^2)A^5}{120} \right] \\
 y &= K_0 \left\{ M + N \tan \phi \left[\frac{A^2}{2} + \frac{(5-T+9C+4C^2)A^4}{24} + \frac{(61-58T+T^2+600C-330e^2)A^6}{720} \right] \right\}
 \end{aligned}$$

where

$$\begin{aligned}
 K_0 &= 0.999, \lambda_0 = 121^\circ, a = 6378137m, b = 6356752m \\
 e^2 &= \frac{a^2 - b^2}{a^2} \\
 e'^2 &= \frac{e^2}{1 - e^2} \\
 N &= \frac{a}{\sqrt{1 - e'^2 \sin^2 \phi}} \\
 T &= \tan^2 \phi \\
 C &= e'^2 \cos^2 \phi \\
 A &= \cos \phi (\lambda - \lambda_0) \\
 M &= a \left[\left(1 - \frac{e^2}{4} - \frac{3e^4}{64} - \frac{5e^6}{256} - \dots \right) \phi \right. \\
 &\quad \left. - \left(\frac{3e^2}{8} - \frac{3e^4}{32} - \frac{45e^6}{1024} - \dots \right) \sin 2\phi + \dots \right]
 \end{aligned}$$

(1)

Image coordinate transformation: The image coordinate transformation is based on the method proposed by Tsai (Tsai 1987, Lenz and Tsai 1987, and Shapiro and Stockman 2001). The principle of Tsai's method is shown in Fig. 3. First, a set of five coordinate points, such as mappings, are established. Each mapping is composed of one 3D world coordinate and one image coordinate. In GODTA, the 3D world coordinate is a TM2 coordinate transformed from the GPS coordinate. The unit of an image coordinate is in pixels with respect to the original point on the left-top corner.

- x_i, y_i, z_i : It is a 3D world coordinate. Although there are only x and y values in a TM2 coordinate, most of the surveillance area of a camera is limited and can be assumed to be planar. Therefore, the value of z-axis is set as zero.
- u_i, v_i : It is a point on the camera image, e.g. 323, 244. The original point (0, 0) is on the left-top corner.

In order to establish the coordinate mappings, a person carries a GPS-enabled mobile device and moves into the FOV of a camera. The TM2 coordinate and the image coordinate are recorded simultaneously. The person moves to five different positions in the FOV of the camera and five coordinate mappings are recorded in the same way. A mapping example is shown in Fig. 4. The TM2 coordinate is 220919.545 and 2662670.167 and the corresponding coordinate is 135 and 253.

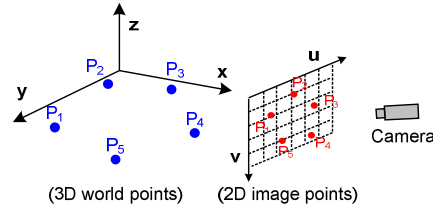


Fig. 3. The 3D world points and corresponding 2D image points for the camera calibration



Fig. 4. A mapping example of TM2 and image coordinates for camera calibration

When the mappings of a camera are established, two parameters for the transformation from the TM2 coordinate to the image coordinate are computed, according to Tsai's method. They are intrinsic and extrinsic parameters as described below:

- **Intrinsic parameters:** These are the true camera parameters. One parameter, focal length f , is used in the coordinate transformation of GODTA. f is the distance from the optical center to the image plane.
- **Extrinsic parameters:** These represent the position and orientation of the camera system in the 3D world. They include the translation (T) and rotation (R) matrices as shown in Eq. 2:

$$T = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}, R = \begin{bmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2)$$

Assuming that there is a new TM2 coordinate (x', y', z') , the corresponding image coordinate (u', v') can be transformed based on (x', y', z') and the above parameters. The computation of (u', v') is shown in Eq. 3:

$$\begin{aligned} u' &= f \frac{r_{11}x' + r_{12}y' + r_{13}z' + t_x}{r_{31}x' + r_{32}y' + r_{33}z' + t_z} \\ v' &= f \frac{r_{21}x' + r_{22}y' + r_{23}z' + t_y}{r_{31}x' + r_{32}y' + r_{33}z' + t_z} \end{aligned} \quad (3)$$

Camera detection: This module is designed to determine the target camera image containing a moving object. From the geographical view, the FOV of a camera is usually an irregular area when there are buildings in front of the camera. For example, the image and the FOV of a camera are shown in Fig. 5. In Fig. 5(a), there is an open area called the “red-brick square” on our campus. The FOV of the camera is marked by the dotted area in Fig. 5(b). The irregular area is not easily represented and checked to determine the target camera. A simple detection method is designed in this study without considering the FOV of a camera. According to the real-time GPS coordinate of the moving object, the nearby cameras can be listed. Then, the TM2 coordinate is transformed to the image coordinate for every listed camera. When the image coordinate (u', v') is within the rectangle range of the image, it means the moving object is within the FOV of the camera. Additionally, when the FOVs of multiple cameras are overlapped and a moving object is within the overlapped area, it is easy to find these cameras.

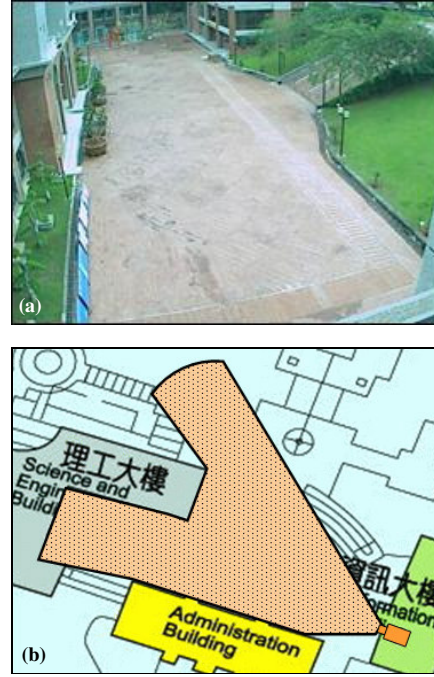


Fig. 5. The image and FOV of a camera (a) the camera image (b) the FOV on 2D map

EXPERIMENTAL STUDIES

Theoretically, GODTA can generate an accurate image coordinate if the calibration points are accurate too. Unfortunately, the current COTS (commercial, off-the-shell) GPS receiver is inaccurate. Its inaccuracy is dependant on its chipset and the skyview situation. Therefore, two experiments were designed to evaluate the performance of the GODTA in a practical environment. One is a fixed point experiment and the other is a continuous tracking experiment. Three kinds of GPS receivers are also used in the experiments to reveal all the important issues in practicing the GPS-based visual tracking service.

Experimental metrics: A Locating Error (LE) is defined to measure the performance of the GODTA. When a GPS coordinate of a moving object is transformed to an image coordinate, denoted as (u', v') , the actual position of the moving object in the image is denoted as (au', av') . LE is defined as the distance between (u', v') and (au', av') , in pixels. A large value for LE means a large error in the GODTA. The computation of LE is illustrated in Fig. 6 and defined in Eq. 4:

$$LE = \sqrt{(u' - au')^2 + (v' - av')^2} \quad (4)$$

Fixed position experiment: This experiment is designed to evaluate the performance of GODTA at fixed positions. The performance is influenced by two factors, the accuracy of the GPS receiver and the skyview of the experimental site. Therefore, two different GPS receivers and two experimental sites are chosen in this experiment. The parameter settings are listed in Table 1.

A tool is also implemented on the Pocket PC for receiving the GPS coordinate and transforming it to a TM2 coordinate. Its screen shot is shown in Fig. 7. The user can select the correct serial port and baud rate and then click the “Start” button. The real-time GPS coordinate and corresponding TM2 coordinate is transformed and displayed on the screen. The “Connection” and “Disconnection” buttons are implemented later for the continuous tracking experiment.

The following experimental steps precede every combination for the GPS receiver and experimental site:

- **Camera calibration:** One person takes a Pocket PC with a GPS receiver and stays at a fixed position. After the GPS receiver is activated, ten real-time TM2 coordinates are recorded for every 30 seconds. Another person saves the real-time image and gets the image coordinate at the fixed position. The average of the ten TM2 coordinates and the corresponding image coordinate is deemed as one calibration point. The above process is repeated until five calibration points are collected.
- **Fixed position test:** One person stays at a randomly-chosen position but not that for the camera calibration. The real-time TM2 coordinate is recorded and the real-time image is saved simultaneously. The above process is repeated at ten different positions.
- **Computation of LE:** Every recorded TM2 coordinate is transformed to an image coordinate (u', v') , according to GODTA. The actual coordinate (au', av') is also acquired manually from the corresponding image. Then, the LE is computed from (u', v') and (au', av') .

The experimental results of the four different combinations are listed in Table 2. Based on the results, conclusions are listed below:

- For two different GPS receivers, the LE of a good skyview is several times smaller than that of a limited skyview. This means that the accuracy of the GPS receiver is influenced by the skyview situation. It causes the LE to also be influenced by the skyview situation.

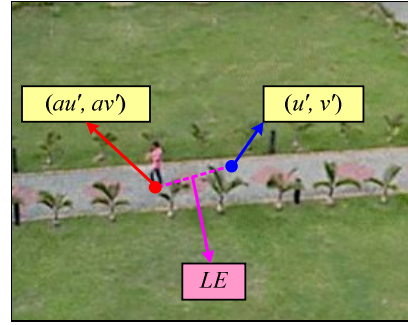


Fig. 6. An example of locating error (LE)



Fig. 7. The screen shot of a tool on Pocket PC

Table 1. The setting of the fixed position experiment

Parameters	Setting	
	A	B
Pocket PC	HP iPAQ hw6515	HP iPAQ hw6515
(GPS Receiver)	(GL20k and GL-LN22 +A-GPS)	(Leadtek 9553X-SiRF Star III+DGPS)
Site	Limited skyview	Good skyview

- For two different sites, the LE of a DGPS receiver (Type B) is several times smaller than that of a general GPS receiver (Type A). This is expected since the DGPS receiver is more accurate than a general GPS receiver.
- For the GPS receiver of Type A, the overall average LE is 553 pixels for both sites. It is impractical to use a 640×480 camera image since the image coordinate is easily outside the image. It also means it is impractical to calibrate a camera by using a general GPS. A camera should be calibrated by a DGPS receiver.

Table 2. The locating errors of the fixed position experiment

GPS receivers	Site A (limited skyview)				Site B (good skyview)		
	No.	(u' , v')	(au' , av')	LE (pixels)	(u' , v')	(au' , av')	LE (pixels)
Type A: HP iPAQ hw6515 (GL20k and GL-LN22+A-GPS)	1	(253, 94)	(288, 436)	344	(1048, 515)	(60, 182)	1043
	2	(454, 259)	(455, 402)	143	(439, 254)	(171, 252)	268
	3	(314, 137)	(367, 320)	191	(484, 227)	(257, 308)	241
	4	(-1890, -13)	(244, 280)	2154	(590, 268)	(346, 424)	290
	5	(-995, 234)	(531, 221)	1526	(653, 214)	(569, 394)	199
	6	(447, 154)	(411, 196)	55	(653, 151)	(467, 238)	205
	7	(363, 107)	(308, 175)	87	(569, 68)	(528, 153)	94
	8	(-74, 50)	(172, 170)	274	(308, 177)	(267, 154)	47
	9	(243, 379)	(252, 132)	247	(397, 110)	(202, 72)	199
	10	(3601, -57)	(248, 115)	3357	(481, 52)	(400, 94)	91
	Average			838	Average		268
Type B: HP iPAQ hw6515 (Leadtek 9553X-SiRF Star III+DGPS)	1	(380, 121)	(288, 436)	328	(59, 198)	(60, 182)	16
	2	(97, 39)	(455, 402)	510	(160, 277)	(171, 252)	27
	3	(64, 31)	(367, 320)	419	(217, 310)	(257, 308)	40
	4	(237, 58)	(244, 280)	222	(362, 362)	(346, 424)	64
	5	(49, 12)	(531, 221)	525	(503, 326)	(569, 394)	95
	6	(93, 19)	(411, 196)	364	(459, 224)	(467, 238)	16
	7	(4, -4)	(308, 175)	353	(502, 157)	(528, 153)	26
	8	(106, 19)	(172, 170)	165	(271, 137)	(267, 154)	17
	9	(-122, -37)	(252, 132)	410	(139, 42)	(202, 72)	70
	10	(-250, -64)	(248, 115)	529	(352, 75)	(400, 94)	52
	Average			383	Average		42

The results also show that GODTA is feasible for a DGPS receiver under good skyview conditions. In the second experiment, a DGPS receiver is used for the calibration phase of GODTA. Then, a general GPS and DGPS are used separately in the operation phase of the GODTA for continuous tracking.

Continuous tracking experiment: In this experiment, the LE of the GODTA is estimated in a continuous moving situation. The experimental parameters are listed in Table 3. Two cameras are deployed at two different sites. Three different GPS receivers, including two general GPSs and one DGPS receiver, are used in the experiment. The cameras are calibrated by using the DGPS receiver. One person takes a Pocket PC with a GPS receiver and then walks along a pre-defined path. The update interval of the GPS coordinate is one second. This means that the coordinate transformation and image recording is done every second.

The experimental site and the deployment of two cameras are shown in Fig. 8. The pre-defined path is marked with a dashed blue line. It starts at the "Car Lot" and ends at the "Design Square". The skyview situation of both sites is good. However, the path goes through the site with limited skyview in the first experiment. The design is mainly to understand the possible consequence when the accuracy of GPS receiver is changed by motion in a certain time period.

A server tool is also implemented for recording the experimental data. The client tool is enhanced to transmit the real-time TM2 coordinate to the server tool via the wireless network. The pre-defined name of a

Table 3. The parameter setting of the continuous tracking experiment

Parameters	Setting
Cameras	2 (AXIS 207W IP Camera)
Sites	Car lot and design square
Mobile devices	Mio A201 (SiRF Star III)
(GPS receivers)	HP iPAQ hw6515 (GL20k and GL-LN22+A-GPS) Mio A201 (Leadtek 9553X-SiRF Star III+DGPS)
Calibration (DGPS receiver)	Leadtek 9553X-SiRF Star III+DGPS
Moving speed	Walk (along a pre-defined path)
Total period	About 220 sec
Update interval	1 sec

moving object is also transmitted by the client tool. The server tool labels the moving object on a 2D map, according to the TM2 coordinate. In addition, the real-time image with a red circle and a name is displayed on the server tool simultaneously when the object is detected in the FOV of a camera. The center of the circle is the image coordinate transformed by GODTA, and its radius is 50 pixels. The moving object can be found in the circle when the locating error (LE) is within 50 pixels. The received TM2 coordinates and captured camera images are stored separately for computation of the LE and playback of the moving trace. A screen shot of the server tool is shown in Fig. 9.

A part of the camera images in the moving trace is shown in Fig. 10. These images illustrate that the moving object can be labeled in the red circle according to the coordinate transformed by the GODTA. It is easy for people to recognize the moving object. It should be noted that a moving object can still be tracked even when the object is totally occluded. An example is shown in Fig. 10.

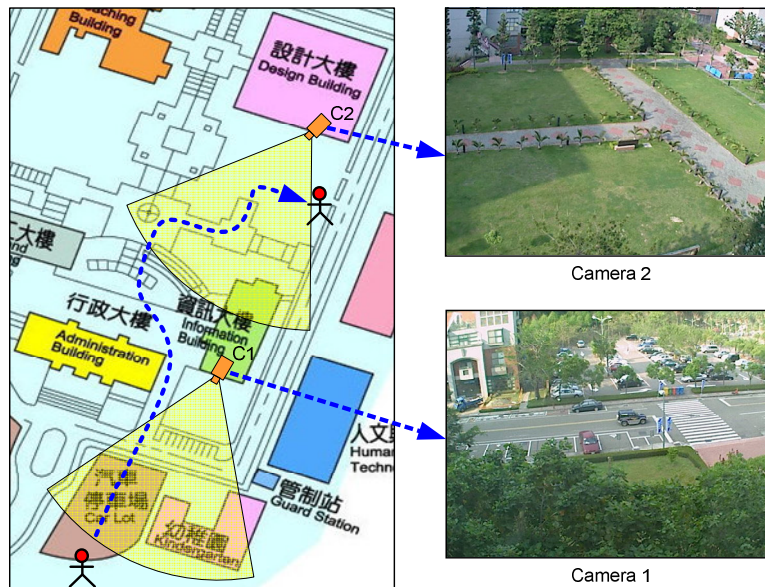


Fig. 8. The experimental site and the views of two cameras

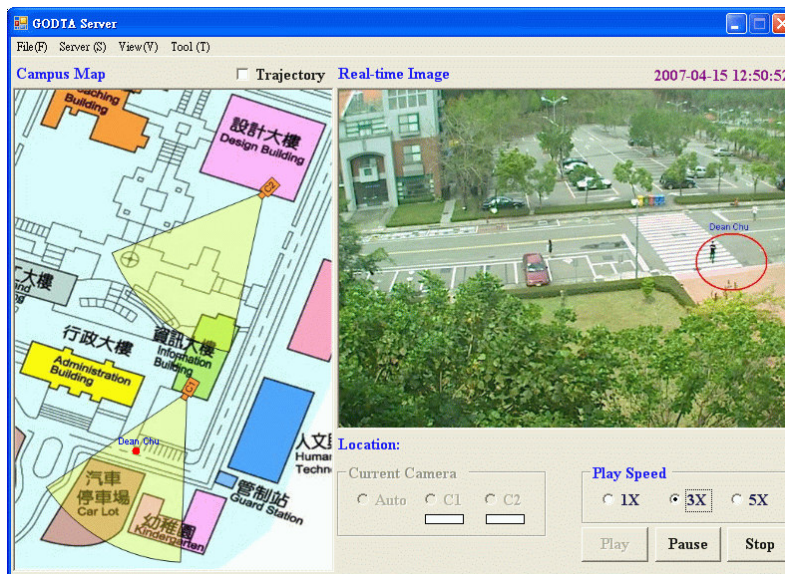


Fig. 9. A screen shot of the server tool

The experiment is repeated three times for every mobile device. When the recorded data is played back and the checkbox "Trace" of the server tool is checked, the moving trace is displayed on the screen as shown in Fig. 11. It is close to the pre-defined path in Fig. 8.

The actual image coordinate is acquired from the recorded image manually. All the LE s of three mobile devices with respect to the moving times are shown in

Figures 12-14, separately. In Fig. 12, the x-axis is the moving time and the y-axis is the LE . The three experiments are represented by three curves. Each point of the curve represents the LE at a specific time. There are two grey areas on the figure. They represent the time when the moving object is occluded by trees and the LE is unable to make calculations from the recorded image.



Fig. 10. A tracking example of a moving object



Fig.11. The moving trace on the 2D map of the server tool

The curves show that, the *LE* is not stable in the moving period since the GPS receiver is influenced by the skyview situation. The change of the *LE* is not consistent for the three experiments. For example, the *LE*s of Ex.1 are larger than that of Ex.3, within zero to 60 seconds. However, the *LE*s of Ex. 1 are smaller than that of Ex.3 after 165 sec. Similar situations can be found in Figures 13 and 14.

Table 4. The experimental results of the continuous labeling experiment

Mobile Devices (GPS Receivers)	Ex.	Samples (valid/total)	<i>LE</i> (pixels)			Overall
			Max.	Min.	Average	
Mio A201 (SiRF Star III)	1	111/235	75	1	40	38
	2	101/219	69	5	31	
	3	95/207	109	1	45	
HPiPAQ hw6515 (GL20k and GL-LN22+A-GPS)	1	98/229	238	3	76	62
	2	94/212	130	15	58	
	3	85/202	134	4	50	
Mio A201 (Leadtek 9553X-SiRF Star III+DGPS)	1	87/216	131	2	40	58
	2	94/212	205	9	77	
	3	91/212	290	8	55	

In the Table 4, a sample is consisted of a received TM2 coordinate and a captured image. A sample is invalid if it is unable to identify its actual position in the image. Therefore, the total number and valid number of samples, the maximum, minimal, average, and overall *LE*s of the three experiments are listed in the table.

The following conclusions can be drawn from the above experimental results:

- The smallest *LE* is found to be 38 by using the first mobile device with a general GPS receiver. It differs from our expectation that the *LE* of the third device with a DGPS receiver would be the smallest. After checking the process and recorded data, it is found that the refresh speed of a GPS coordinate is

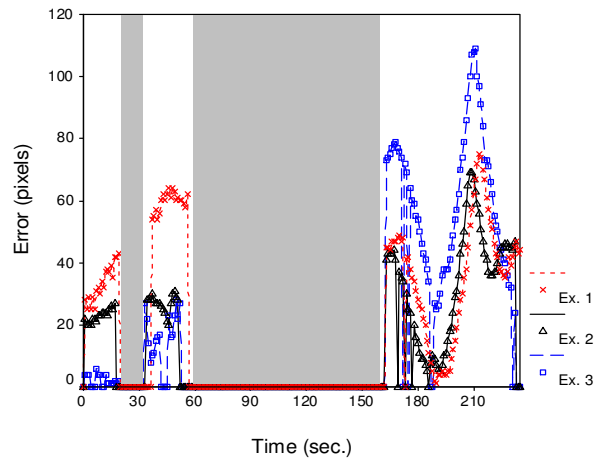


Fig. 12. The estimated *LE* of Mio A201 (SiRF Star III)

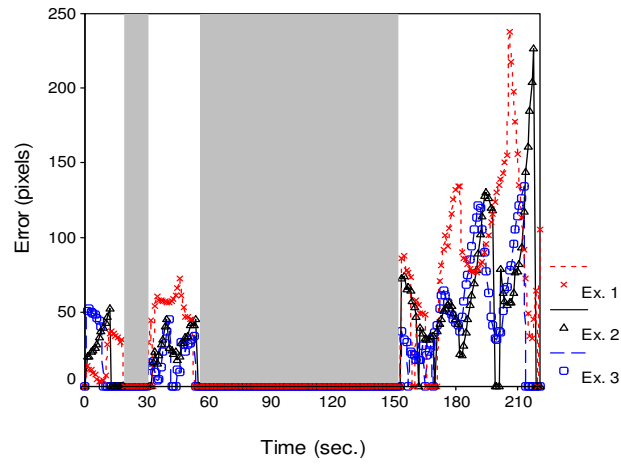


Fig. 13. The estimated *LE* of HP iPAQ hw6515 (GL20k and GL-LN22+A-GPS)

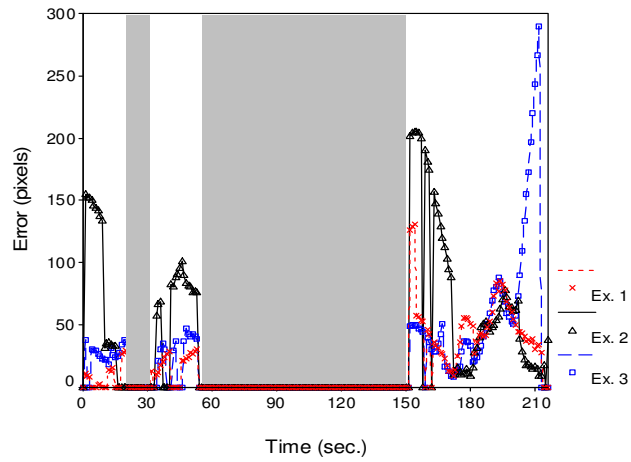


Fig. 14. The estimated *LE* of Mio A201 (Leadtek 9553X-SiRF Star III+DGPS)

too slow. The GPS coordinate is unchanged when the person moves a long distance. It causes the *LE* to increase with time until the coordinate is refreshed. A DGPS receiver may require a longer computation time than a general GPS and thus increases the refresh speed. If the refresh speed of a DGPS is short enough, the *LE* should be smaller than that of a general GPS receiver.

- The overall *LEs* of the first and second devices with a general GPS receiver are 38 and 62 pixels, respectively. It is an acceptable error for the GODTA operated on a general and popular GPS receiver when a camera is calibrated by using the DGPS receiver. It shows that GODTA is feasible in the practical environment.

DISCUSSION

GODTA is the first study to incorporate the GPS system and cameras to provide GPS-based visual tracking service. There is no previous study for comparison. However, the above experimental results and demonstrations show that GODTA can provide a way to detect and track moving objects successfully. When an object carried a GPS-enabled mobile device moves into the FOV of any camera, the object can be marked on the real-time camera image. In advance, some critical issues are found and discussed as follows:

- The inaccuracy and update interval of a GPS receiver limits the applicability of GODTA. Some techniques are used to reduce the noisy measurement, such as a Kalman filter (Simon 2001). The accuracy could be improved with the technical evolution of the GPS receiver. However, GODTA could still be incorporated with image processing techniques to increase its applicability. For example, the candidate area of a moving object generated by GODTA can be used to restrict the processing area and thus increase the efficiency of these techniques.
- Synchronization of GPS coordinates and the camera image: The performance of GODTA is influenced by the synchronization of the GPS coordinates and the camera image. If the acquisition time of a GPS coordinate and the capture time of a camera image are synchronized, it is useful to reduce the *LE*. However, a GPS coordinate is usually delayed caused by the network transmission. A timestamp mechanism can be used for the synchronization of the coordinate and image.
- Real-time tracking of high-speed moving objects: Current experiments focus on low-speed moving objects. If the GODTA is used to track high-speed moving objects, a delay GPS coordinate cannot be used to label the moving object since the object

may have moved outside the camera image. A time-shift technique, which is popularly used in digital TV software, seems suitable for solving this problem. The playback of a real-time image is delayed until the desired coordinate is received. The delay time is adaptable depending on the delay of the GPS coordinate.

Although there are some issues needing further attention in order to fulfill the novel GPS-based visual tracking service on various conditions or moving objects, the proposed GODTA is feasible under normal conditions and can be considered a basis for further development.

CONCLUSION

According to the trend of GPS receivers becoming an integral part of a mobile device, the locations of moving objects can be acquired easily in the near future. The GPS system is incorporated with cameras and an approach, called GODTA, is proposed in this paper. The image coordinate transformed from a GPS coordinate is used to label an object on the camera image directly. The periodical update of the GPS coordinate enables the tracking of a moving object. The experimental results show that GODTA is feasible for a GPS-based visual tracking service. The future works are listed below:

- **An automatic calibration mechanism:** Although the calibration of a camera is a one-time effort, it is still a burden when there are a large number of cameras. An automatic calibration is useful to reduce the burden and increases the practicality of a GODTA. Some image processing techniques can be used for the automatic calibration.
- **Precise labeling of moving objects:** A moving object is unable to be labeled precisely simply according to its inaccurate GPS coordinate. Some feature-based approaches can be used for precise labeling, such as features extracted from the histogram of a moving object (Zhai et al. 2005). The features can be used to distinguish this object from the others when there are many objects within the candidate area.
- **The incorporation of GODTA and a PTZ camera:** A PTZ camera provides the capability for monitoring a wide area. If the GODTA can be incorporated with a PTZ camera, the GPS coordinate can be used to direct the PTZ camera to provide a close view tracking in a wide area.

GODTA is a novel approach to the incorporate a GPS system and cameras. Future studies can aim to fulfill a GPS-based visual tracking service for the upcoming UbiCam environment.

ACKNOWLEDGEMENT

This project was financially supported by the National Science Council under Grant No.: NSC 95-2221-E-324-042.

REFERENCES

- Black, J., T. Ellis, and P. Rosin, 2002. Multi view image surveillance and tracking. Proceedings of IEEE Workshop on Motion and Video Computing, 2002, Dept. Of Comput. Sci., Cardiff Univ., UK., pp: 169-174.
- Chang, T.H. and S. Gong, 2001. Tracking multiple people with a multi-camera system. Proceedings of IEEE Workshop on Multi-Object Tracking, 2001, IEEE, pp: 19-26.
- Cucchiara, R., C. Grana, M. Piccardi, 2003. Detecting moving objects, ghosts, and shadows in video streams. IEEE Trans. on Pattern Anal. and Mach. Intell., 25(10): 1337-1342.
- Dawoud, M., M. Khalil, M.E. El Najjar, El Hassan and B. Ziade *et al.*, 2008. Tracking system using GPS, vision and 3D virtual model. Proceedings of the 3rd International Conference on Information and Communication Technologies: From Theory to Applications, 2008, Fac. and Eng., Lebanese Univ., Tripoli, pp: 1-6.
- Haritaoglu, I., D. Harwood, and L.S. Davis, 2000. W4: Real-time surveillance of people and their activities. IEEE Trans. on Pattern Anal. and Mach. Intell., 22(8): 809-830.
- Hu, W., M. Hu, X. Zhou, T. Tieniu, L., Jianguang and S. Maybank, 2006. Principal axis-based correspondence between multiple cameras for people tracking. IEEE Trans. on Pattern Anal. and Mach. Intell., 28(4): 663-671.
- Lenz, R.K. and R.Y. Tsai, 1987. Techniques for calibration of the scale factor and image center for high accuracy 3D machine vision metrology. Proceedings of the IEEE International Conference on Robotics and Automation, 1987. IBM T. J. Watson Research Center Yorktown Heights, NY., pp: 68-75.
- Liao, H.C. and Y. C. Cho, 2008. A new calibration method and its application for the cooperation of wide-angle and pan-tilt-zoom cameras. Inform. Technol. J., 7(8): 1096-1105.
- Michelsoni, C., G.L. Foresti, and L. Snidaro, 2005. A network of co-operative cameras for visual surveillance. IEE Proc. Vis. Image Signal Process., 152(2): 205-212.
- Ridder, C., O. Munkelt, and H. Kirchner, 1995. Adaptive background estimation and foreground detection using kalman-filtering. Proceedings of International Conference on Recent Advances in Mechatronics (ICRAM '95), Academic Press, pp: 193-199.
- Shah, M., O. Javed, and K. Shafique, 2007. Automated visual surveillance in realistic scenarios. IEEE Multimed., 14(1): 30-39.
- Shapiro, L.G. and G.C. Stockman, 2001. Computer Vision. Prentice-Hall, New York, pp: 444-456.
- Simon, D., 2001. Kalman filtering. Embed. Syst. Program., 14(6): 72-79.
- Stauffer, C. and W.E.L. Grimson, 2000. Learning patterns of activity using real-time tracking. IEEE Trans. on Pattern Anal. and Mach. Intell., 22(8): 747-757.
- Tsai, R.Y., 1987. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. IEEE J. of Robot. and Autom., RA-3: 323-344.
- Zhai, H., P. Chavel, Y. Wang, S. Zhang, and Y. Liang, 2005. Weighted fuzzy correlation for similarity measure of color-histograms. Opt. Commun. 247(1-3): 49-55.