# VISUAL TRACKING OF PEDESTRIANS JOINTLY USING WI-FI LOCATION SYSTEM ON DISTRIBUTED CAMERA NETWORK

*T. Miyaki†, T. Yamasaki‡and K. Aizawa‡*

Dept. of Frontier Informatics†and Dept. of Information and Communication Engineering‡
The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

**ABSTRACT**

Object tracking with multiple cameras is a fundamental problem in wide-area surveillance application, but it has difficulties to achieve accurate and stable performance because of disjoint shot areas or initial object identification problems. We propose a novel object tracking method which jointly uses estimated location information of the target derived from a set of Wi-Fi signal strength values with video images from cameras. Apart from sensor fusion techniques proposed in the past which use another kind of sensors (eg., Global Positioning System (GPS), pressure sensors on the floors to detect foot steps, laser-range scanners, etc), our approach can cover wide-areas both indoor and outdoor in low cost because of propagation characteristics of Wi-Fi signals. Particle filtering is applied to achieve tracking the target from video images, with this Wi-Fi location estimation. This paper describes system architectures and the experimental results of the pedestrian tracking technique with Wi-Fi location estimation.

## 1. INTRODUCTION

Surveillance system is getting more important because of our growing concern about security. So fundamental image processing techniques such as object detecting and tracking is demanded to get higher and higher performance. Today, image sensors and processing units are getting cheaper and high imaging quality. Building a multimedia sensor network which consists of multiple cameras and multiple image processing units costs not so high for important application. On multimedia sensor networks, object tracking with multiple cameras is one of a key feature [1][2][3].

But it is difficult to achieve more accurate and stable tracking performance by only using a series of image information. To make realize this objective, sensor fusion techniques are proposed in many application areas which use another kind of sensors with video information from cameras (eg., Global Positioning System (GPS), pressure sensors on the floors to detect foot steps, laser-range scanners[4], etc.) Because of the characteristics of sensors to be used, these solution often tends to restrict target area of surveillance application. GPS cannot be used indoors because radio waves from GPS satellite do not penetrate into buildings and pressure sensors and laser-range scanners cannot be set in wide areas because of the cost problem. A location estimation system which can be used easily in wide area is strongly needed for surveillance application even if it's accuracy of estimated location is lower than that of GPS like systems.

We propose a novel visual tracking method which additionally use location information of the target by utilizing signal strength values of multiple Wi-Fi access point (AP) to detect current location of the Wi-Fi client. There proposed two types of positioning systems over wireless networks. One is called Time Difference of Arrival (TDOA), the other is called Signal Strength models. The former approach needs accurate time synchronization, because of this reason the client which receive wireless signals have to be special one for this purpose. Signal Strength model is much easier to use. The client scan signal strength from multiple APs and match the data with location database which has a set of estimated location of APs. As client-side does not need processing power, light-weight equipments like Wi-Fi enabled mobile phones or PDAs are enough to estimate location of the user[5][6]. With this Wi-Fi location estimation, particle filtering[7][8] is applied to achieve tracking the target from video images.

Recently, city-wide public wireless LAN services have become increasingly popular. Around our campus, we developed outdoor wireless LAN infrastructure in collaboration with NTT Communications Corporation. We also developed our distributed camera network system in the same area which is a kind of multimedia sensor networks composed of 36 cameras and 12 camera processing nodes connected via the wireless LAN service. On this environment, we implemented the proposed joint tracking method and get an experimental result which support validity of the purpose. With this proposed method, such a novel application like networked "keeping an eye on" system for children passing by the street will be accomplished using urban infrastructure with a feasible cost. As higher security is requested in urban life, especially for children, it is important technology.

This paper shows system architectures and the experimen-

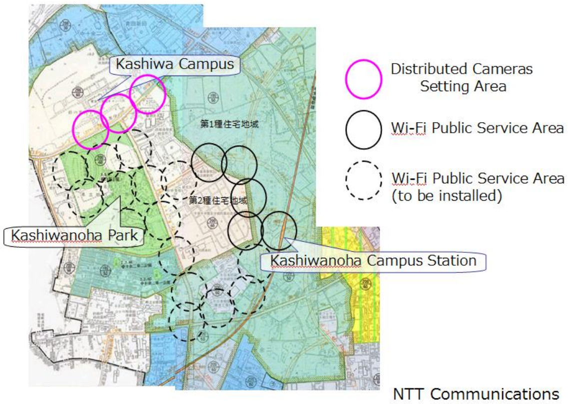


**Fig. 1**. Target area of pedestrian tracking using distributed camera network system

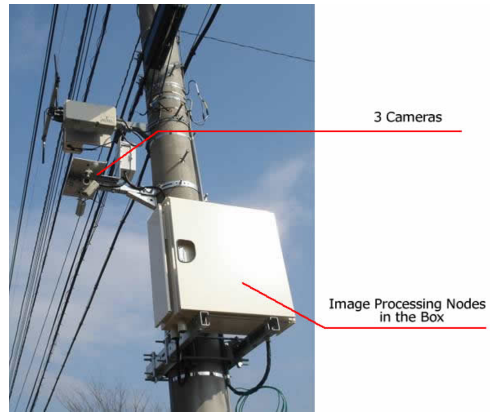


**Fig. 2**. Distributed camera node setup


**Fig. 3**. Screen shot of PlaceEngine client software.

tal results of the pedestrian tracking technique with Wi-Fi location estimation. Following sections describe our system to develop convenient joint tracking system on top of wireless LAN infrastructure that supports wide-area pedestrian tracking.

## 2. DISTRIBUTED CAMERA SYSTEM

To make realize our joint tracking algorithm, distributed camera network system is installed around the campus. We installed 12 camera nodes, which are composed of three cameras, a processing unit, and a wireless LAN interface on 12 power poles along the street in front of our campus (Fig. 1). The captured image size is 640x480. The processing unit is contained in an iron box. The cameras are set so that the view range overlaps with each other. Data transfer from the sensor nodes to the central control server is achieved using the HOTSPOT service provided by NTT Communications, which is a wireless LAN connection service available in public space. As a result, the system is easy-to-implement and low cost. Fig. 2 shows an installed sensor node. It is attached to the electric pole at about 3.5m from the ground.

Wireless LAN AP has been laid on some streets of the area. Circles plotted in Fig. 1 show covered area by this network. This system is also used as a public infrastructure for urban network system.

## 3. JOINT TRACKING ALGORITHM

### 3.1. Wi-Fi Location Estimation

We adopted the object tracking method which uses location information of the target by utilizing signal strength values of multiple Wi-Fi access point (AP) to detect current location of the Wi-Fi client. The precedence work shows several approaches about this location estimation technique[5][6].

Wi-Fi location estimation algorithms depend on an initial training phase. This involves war drivings around the neighborhood with a Wi-Fi equipment and an attached GPS device. The Wi-Fi card periodically scans its environment to discover wireless networks while the GPS device records the latitude-longitude coordinates of the war driver when the scan was performed.

Our system use PlaceEngine[9] as Wi-Fi location estimation engine. In this system, client software scans 802.11 beacon packets and submits these information (MAC address of the Wi-Fi AP (48bit), Service Set Indicator (SSID) and Received Signal Strength Indicator (RSSI) for each scanned APs) to location database server which stores number of pairs of MAC address and its location (latitude-longitude coordinates) via wireless network. The server estimates the user's location from scanned Wi-Fi signal strength by Centroid algorithm[5] and returns estimated location information in the format as shown in Table 1. Thus the wireless networked environment is used for both location estimation and data transfer.

### 3.2. Pedestrian Tracking by Multiple Cameras

Particle filtering algorithm is suitable for pedestrian tracking on video sequences.

**Table 1**. Result data format from location estimation system.

| latitude | longitude | number of APs | scanned time |
|---|---|---|---|
| 139.761988 | 35.71413 | 4 | 1169179446 |
| 139.761984 | 35.714128 | 7 | 1169201291 |
| ... | ... | ... | ... |

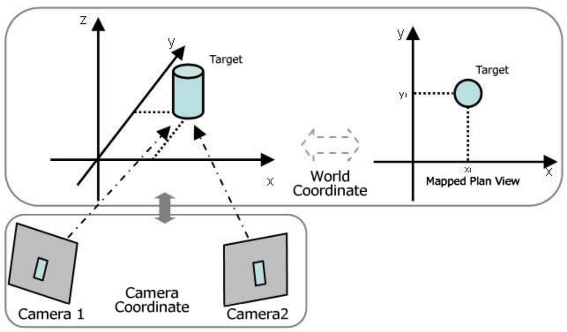


**Fig. 4**. Mapped plan view from multiple cameras.

It should be done in a distributed way at the processing nodes because of the limited network bandwidth. The result of the tracking have to be handed over from one camera to the adjacent camera. The handover tracking is described below.

#### *3.2.1. Object Tracking using Particle Filter*

Particle filter (sequential Monte-Carlo method) is well known for its good object tracking performance in computer vision application[7][8]. We modeled the target by a simple three dimensional cylindrical object, which is shown in Fig. 4. Images captured from cameras are degenerated to two dimensional planar images. By calibrating cameras before capture, the captured images can be mapped into three dimensional X-Y-Z coordinate ("World coordinate" in Fig. 4). We evaluate the likelihood of each particles on this X-Y coordinates which is approximately the same as the ground plane.

#### *3.2.2. Handover*

Two types of handover are required. One of them is a handover between cameras of the same node. The other is the handover between the nodes. When the between node handover is needed, only the features of the object model is transmitted to the other node. Then, minimum amount of communication is enough between the nodes. Fig. 5 shows examples of the handover between the nodes.

**Fig. 5**. Handover tracking by particle filter. (Camera images superposed by predicted rectangles.) Person tracked in the left image is handed over. Right image is the view of one adjacent camera.

### 3.3. Joint Tracking

Based on these two algorithms, particle filtering and Wi-Fi location estimation, we propose the pedestrian tracking algorithm which can be used in wide-area and can be one of the solutions to initial object detection problems in traditional particle filtering approach.

These two algorithms has complemental characteristics. Particle filtering on video can accurately detect and track the position of targets by no means inferior to GPSs, but cannot cover the areas outside the shot scene and hard to identify targets one by one. On the other hand, Wi-Fi location estimation system can be used in everywhere as long as the client equipment can receive at least one Wi-Fi AP independent of indoor or outdoor, but the location estimation is not always stable and the accuracy is affected by the environment (eg., number of APs, database of location engine, etc.).

Our focus is to combine these characteristics by switching the two algorithms. Switching strategy is decided by a priority of characteristics. In the area where cameras can shoot, the system gives priority to tracking accuracy. In the area where cameras can not cover, Wi-Fi tracking is given higher priority. For this strategy, switching is occurred at the boundary of covered area by cameras. The boundary is known in advance to every camera nodes, Cameras which has such boundary calculate the background subtraction to detect object penetration. When both the location information and color features of detected object matched with the one detected by Wi-Fi location estimation, the tracking strategy is switched.

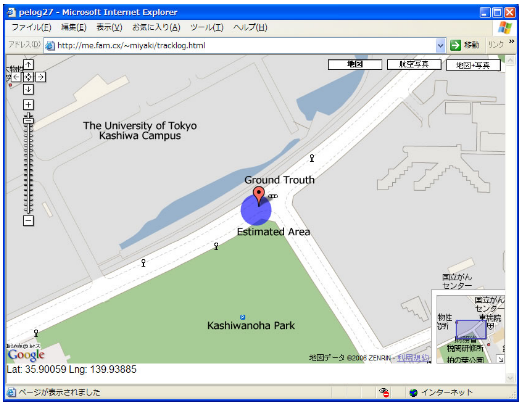


**Fig. 6**. Predicted area of the target by Wi-Fi location estimation and ground truth by GPS.

## 4. EXPERIMENTAL RESULTS

### 4.1. Initial Training of Wireless Positioning System

Before the experiment, the Wi-Fi location estimation system needs to be calibrated for target areas by making registration of training dataset. We registered Wi-Fi signal strength data at 20 different point in the camera node area (approximately $100m^2$). On the estimation phase, maximum error is 18m compared with GPS detected result.

### 4.2. Target Detection

In this experiment, we evaluate the scene which the target walked into camera-covered-area (CCA) from the outside of CCA. The target has PDA which runs Wi-Fi location engine every 10 seconds. The camera nodes receives the User ID and color features information of the target from PDA when the target location by Wi-Fi estimation is close to camera boundary. Fig. 6 shows the estimated location (areas filled with circle) and the real location derived from GPS. As shown in Fig. 7, algorithm switching from Wi-Fi location estimation to particle filtering is successfully accomplished.

## 5. CONCLUSION

A distributed camera system which achieve pedestrian tracking is described. On this system, object tracking algorithm consists from two parts: one is traditional video based particle filtering and the other is Wi-Fi based location estimation using wireless signal strength which PDAs receive from Wi-Fi AP. With this joint tracking algorithm, targets enter from areas which can not be covered by any cameras are properly detected.

From experimental results confirmed on our distributed camera network system, two algorithms, Wi-Fi location estimation system and a particle filter algorithm, are smoothly switched by interacting using wireless network infrastructure.

**Fig. 7**. Pedestrian tracking using joint tracking algorithm.

## 6. REFERENCES

[1] R. Cucchiara, "Multimedia surveillance systems," in *Proc. of ACM VSSN*, 2005, pp. 3–10.

[2] D. Makris, T. Ellis, and J. Black, "Bridging the gaps between cameras," in *Proc. of IEEE CVPR*, 2004, vol. 2, pp. 205–210.

[3] C. Micheloni, G.L. Foresti, and L. Snidaro, "A network of co-operative cameras for visual surveillance," in *Proc. of IEE VISP*, 2005, pp. 205–212.

[4] H. Zhao and R. Shibasaki, "A novel system for tracking pedestrians using multiple single-row laser-range scanners," in *IEEE Trans. on Systems, Man and Cybernetics*, 2005, pp. 283–291.

[5] Y.C. Cheng, Y. Chawathe, A. LaMarca, and J. Krumm, "Accuracy characterization for metropolitan-scale Wi-Fi localization," in *Proc. of ACM MobiSys*, 2005, pp. 233–245.

[6] A. LaMarca et al., "Place Lab: Device Positioning Using Radio Beacons in the Wild," *Proc. of IEEE Pervasive*, 2005.

[7] M. Isard and A. Blake, "Condensation – conditional density propagation for visual tracking," *Int. J. of Computer Vision*, vol. 29, pp. 5–28, 2004.

[8] Z. Khan, T. Balch, and F. Dellaert, "MCMC-based particle filtering for tracking a variable number of interacting targets," in *IEEE Trans. on PAMI*, 2005, pp. 1805–1819.

[9] PlaceEngine, "http://www.placeengine.com/," SonyCSL.