Networked Cellular Motion Detection System by Using Pyroelectric Infrared Sensor for Surveillance

Minoru Kanamaru[1], Hiroyuki Kawai[1], Hisato Kobayashi[2], Junya Tatsuno[3], Hiromi Mochiyama[4] and Nobuaki Kobayashi[1]
[1]Department of Robotics, Kanazawa Institute of Technology, Japan
[2]Department of Engineering and Design, Hosei University, Japan
[3]Department of Mechanical Engineering, Kinki University, Japan
[4]Department of Intelligent Interaction Technologies, University of Tsukuba, Japan
hiroyuki@neptune.kanazawa-it.ac.jp

Abstract— **In this paper, we propose a networked cellular motion detection system by using pyroelectric infrared sensors for surveillance. The networked sensor technology has a potential capability to solve some of our most important scientific and societal problems. But, difficulties of processing are always big problems in case of such huge amount of information acquired by the distributed vision systems. The proposed method gets a hint from information processing of human hearing organs and compound eyes of insects. By the constructed testbed as a motion detection cell, we confirmed that the proposed method can be utilized to detect human behavior and realized as a real system.**

I. Introduction

The networked vision systems have a potential capability to solve some of our most important scientific and societal problems such as security monitoring for huge public space. Such networked system can acquire huge amount of information, but we face to the serious problems, i.e., how we can handle such huge amount of information and how we can retrieve our necessary intelligence. Even if distributed data processors assist the computational task and reduce the network traffic; it might be very difficult for the central processor to rebuild and to analyze the information of the whole space from the information gotten by decentralized processing [1].

On the other hand, living things processes information very efficiently. Human being recognizes voice or sound by an adroit way. In terrestrial vertebrates, sound waves in the air enter the outer ear, strike the tympanic membrane [2]. The sound waves are converted to fluid waves in the cochlea by a series of mechanical couplings in the middle ear. The fluid waves cause vibration of the basilar membrane, on which sit sensory hair cells in the Corti's organ [3]. Our brain can recognize the sound or the voice in real time, by which hair cells are oscillating and how big their magnitude. Namely, our brain retrieves the necessary information from the sound waves by monitoring the dynamical motions of hair cells. In other words, each hair cell compresses the information in the allotted frequency band ideally. The information format is changed from sound to dynamical motion.

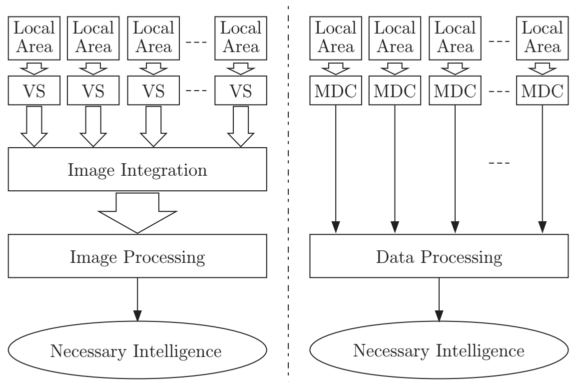

Fig. 1. Information flow. Left: With conventional image processing. Right: With information transformation. VC and MDC mean a vision system and a motion detection cell, respectively.

There also exists the similar supervisory monitoring in the natural world. Some insects can recognize flying baits at unbelievable instantaneous moment. They do not have enough computational power in their brain to execute the image processing. Thus, the hint must lay in their compound eyes and the consecutive neural network [4]. Each cell of the compound eyes must send a very simple signal to the neural network; it must not visual image.

We can get two aspects from the natural world.

1) Supervisory Monitoring: The central processor does not treat local data directly; it retrieves necessary intelligence from Meta data acquired by local agents.
2) Changing Information Category: The central processor does not treat image data directly; it retrieves necessary intelligence from different kind of value.

Based on the above concept, the networked cellular vision system has been proposed in our previous work [5]. Although estimated trajectories of human motions can be obtained by using the networked cellular vision system, it is not easy to construct a real system.

In this paper, we propose networked cellular motion detection system by using pyroelectric infrared sensors instead of the vision system. Because our final objective is the realization of the real system as a surveillance system,

Fig. 2. Conventional pyroelectric infrared sensor and fresnel lens.

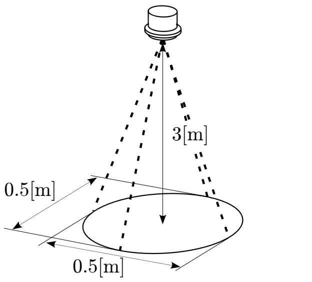

Fig. 3. Field of view of the pyroelectric infrared sensor in the 3D workspace.

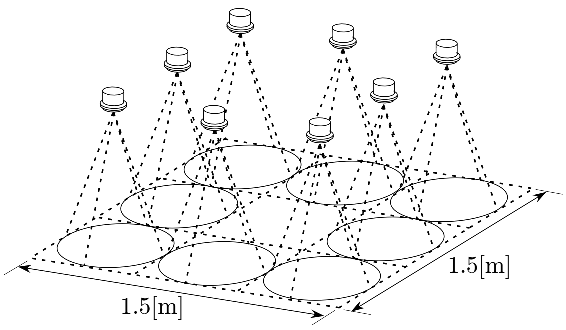

Fig. 4. Field of view of the motion detection cell in the 3D workspace.

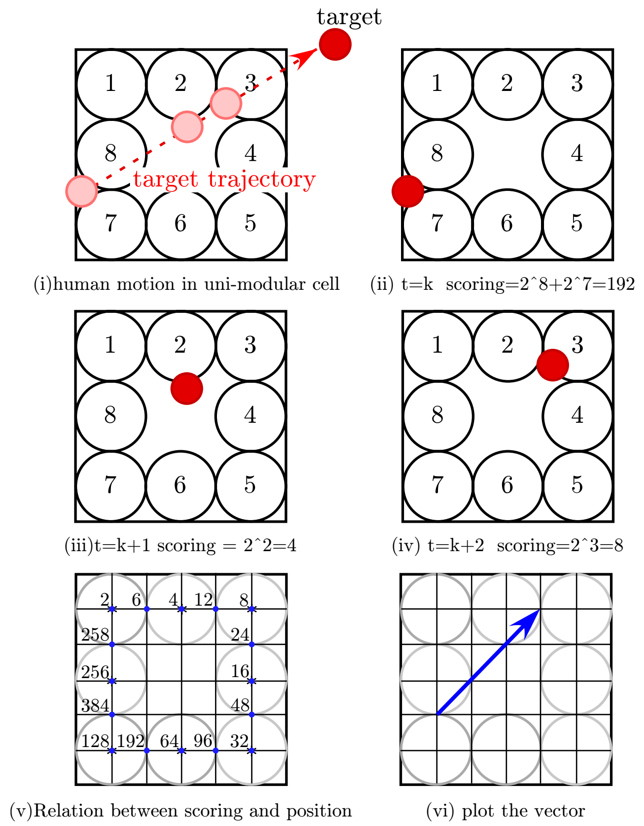

Fig. 5. Vector generation for human motion in the motion detection cell.

we show the constructed testbed as a motion detection cell. While the cellular vision [5] changes image information into driving voltages in order to obtain the vector of human motion, the motion detection cell obtain the vector of human motion directly. In Fig. 1, the left block-diagram shows the information flow of the conventional image processing and right one shows the new scheme with the information transformation; where each motion detection cell send only the data of its autonomous movement to the central unit.

II. Cellular Motion Detection System

A. Basic Structure of Motion Detection Cell

In this section, we propose a uni-modular sensor device, which we call the motion detection cell, by using the pyroelectric infrared sensor and fresnel lens as shown in Fig. 2. The motion detection cell is constructed by connecting eight conventional pyroelectric infrared sensors. The pyroelectric infrared sensor has an ability to detect the change of incoming infrared light, and which used to detect a human. Since an output voltage from the sensor is a faint signal, the signal is converted to the digital signal via the operational amplifier and the window comparator to handle easily. The fresnel lens has an ability to narrow the monitoring area of the pyroelectric infrared sensor.

We configure the networked sensing system by connecting large number of these motion detection cells. Let us assume that each sensor is hanging from the ceiling of 3[m] height and watching down a circle 0.5[m] in diameter vertically as represented in Fig. 3. We let this area be the responsible monitoring area of the sensor. Each sensor send output data to the central processor as one(digital signal) while sensors catch target point. Thus, the monitoring area of the motion detection cell is 1.5[m] $\times$ 1.5[m] as shown in Fig. 4.

Next, we describe how to process the target moving on the cell. Fig. 5 (i) shows an example of target moving in a cell and monitoring area of each sensor intuitively. Eight circles show the monitoring area of each sensor, and each number means the sensor number. Four markers and dashed line show the target for each instant of time and sample target motion, respectively. Fig. 5 (ii) to (iv) show the situation of the target for each instant of time. By processing the output data from sensors, the central processor determines components of a vector: entry point, exit point and time of each. The central processor calculates $scoring$ to determines a vector coordinate. Let us assume that reacting sensor number is $sensornum(i)$, equation of the $scoring$ is shown as follow.

$$scoring = 2^{sensornum(i)} \tag{1}$$

Let us assume that a vector coordinate gets the center point and the boundary point of each monitoring area, the resolution of a cell is five-by-five. Therefore, the entry and

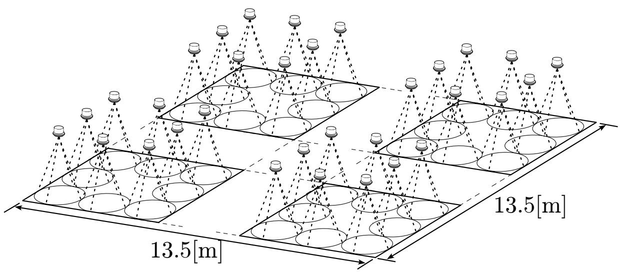

Fig. 6. Cellular motion detection system for square space.

exit points of the vector should be selected from 16 points as shown in Fig. 5 (v).

The following is the central processor work of explaining how to generate a vector.

1) The central processor calculates the *scoring* when the adequate time T_1 goes on after the first sensor reacts, or second one reacts.
2) The central processor determines a entry point using the *scoring* obtained from step 1) and Fig. 5 (v). If not corresponding to any through *scoring* shown in the figure, recalculate the *scoring* using the $sensornum(i)$ of the reacted first sensor only, and determine the entry point. Add the $sensornum(i)$ that reacted secondarily to the *scoring* of a exit point.
3) The central processor calculates the *scoring* when the adequate time T_1 goes on after the next sensor reacts. The exit point will be determined using the *scoring* and Fig. 5 (v). If any sensors do not react when the adequate time T_2 goes on after the determination of the entry point, the central processor determines components of a vector using the status of the entry point. If two sensors reacted when the central processor determines the entry point, the processor determines the entry point and exit point in reacting order. If not, it back to step 1).
4) The central processor plots vector using the entry point and exit point, and it back to step 1).

The central processor calculates the time of entry and exit from the time when each sensor react. If plural sensors react, the central processor uses the average of times when plural sensors react. Since T_1 and T_2 depend on the situation, we select $T_1 = 1$[s] and $T_2 = 2$[s] experimentally. In the case of Fig. 5, the central processor generates the vector with the entry point and the exit point as 384 and 12 as shown in Fig. 5 (vi) by the proposed strategy.

B. Simulation of Cellular Motion Detection System

In this simulation, we adopt the following simple dynamics as the basic model of human walking.

$$m\ddot{x} + \mu\dot{x} = F \tag{2}$$

where m, μ and F represent mass of the human, friction and force in the human walking, respectively. x is the position of the human. m and μ are constants which satisfy $57 \leq m \leq 63$ and $4 \leq \mu \leq 8$, respectively. F takes a random continuous variable during $[-20, 20]$ throughout the simulation. We assume that the human moves independently. Note that the human motion itself is not important in this paper; the crucial issue to be concerned is to detect the outlines of such motions by the proposed method.

We consider a square space as a target monitoring space. We install $9 \times 9 = 81$ motion detection cells at the ceiling of the square as shown in Fig. 6. Since each cell monitors area of 1.5[m] $\times$ 1.5[m], the total monitoring area is 13.5[m] $\times$ 13.5[m]. Simulations are carried out by using Matlab and Simulink with Simulink 3D Animation. The original movie from Simulink 3D Animation can be seen on the Web site [6].

Fig. 7 shows the sample human walking motions generated by Eq. (2). There are ten persons walking in the square. The circles mean their entry points and the crosses represent their locations at the edge of the filed or the final time

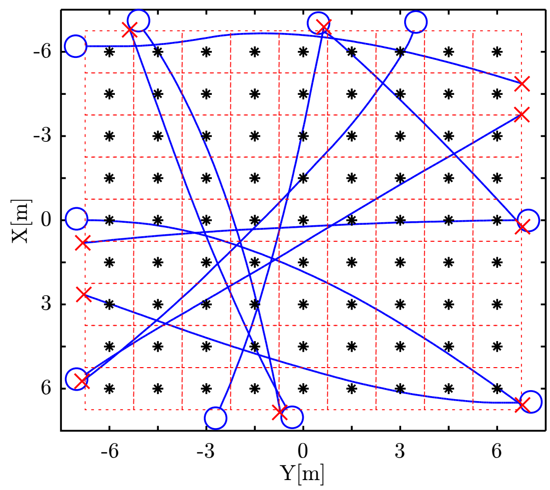

Fig. 7. Human motion in square for 60 seconds.

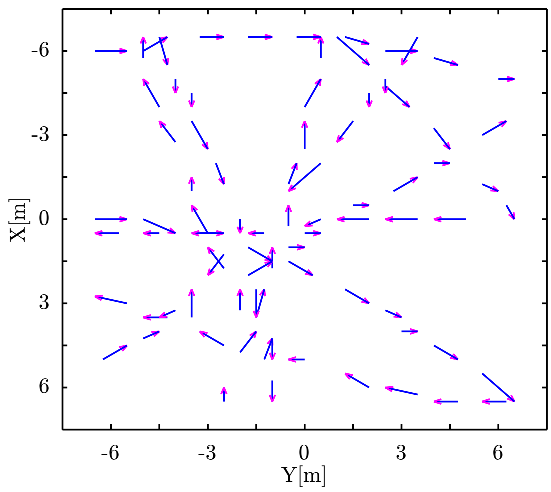

Fig. 8. Vector map from each motion detection cell.

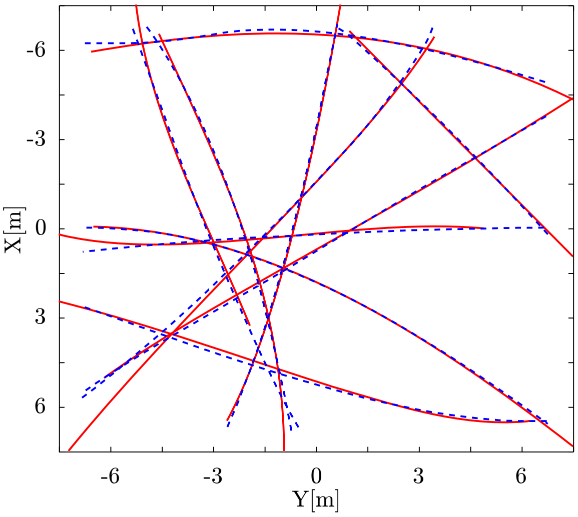

Fig. 9. Estimation for human motion in square.

($t = 60$). In Fig. 7, the mark '*' and dashed lines denotes the center position and monitoring areas of each cell. Each sensor sends the output data to the central processor, when it catches the target point. We can acquire person's walking as vectors using these informations of each cell. Fig. 8 shows a vector map, where each vector is composed by output data from each cell.

Since this vector map Fig. 8 almost coincides with Fig. 7, we can conclude that the generated sample motion is clearly rebuilt by our cellular motion detection system. From the output of the cellular motion detection system, we can easily recognize the situation of the square: how many persons are walking in the area; how fast they are; which direction they are going to. The output of the cellular motion detection system is just the digital signal of each sensor, thus the amount of data is remarkably small comparing with conventional vision systems. Moreover, this cellular motion detection system consists of motion detection cells, we can easily expand the size of monitoring area just by adding the motion detection cells.

III. Trajectory Finding

As described in the former section, the output vector map of the cellular motion detection system is very intuitive for human operators. We can easily imagine the continuous trajectory of the moving objects from the discrete vector fields. However, we need more systematic way to determine the trajectory for machine recognition. Each motion detection cell reports the data of their motion when it finishes the tracking task. The detail is as follows. As described in the former section, when the target objects gets in the range, then the motion detection cell captures the target and follows it until the target is out of the range. At the timing that the target vanishes, the motion detection cell reports the following data to the central processing unit.

- Average 2D velocity vector.

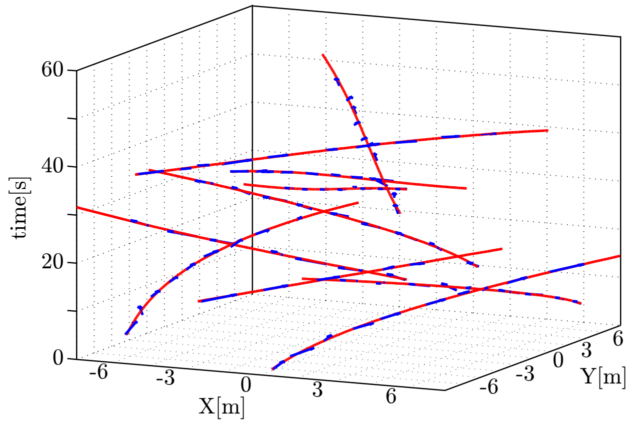

Fig. 10. Estimated trajectory with 2nd order polynomial approximation in 3D representation.

- Average time of entry time and exit time.

This reporting is ad-hoc event, thus it is event driven and asynchronous reporting. The central processing unit rebuilds the trajectories by the following procedure.

A. Initial Setting

The central processing unit has the following data set, which may be null when the whole system starts.

1) n-trajectories $g_1(t), g_2(t), \cdots, g_n(t)$, which are time functions in X-Y plane and they describe the motions of n objects
2) each trajectory has its basement vector set

$$\mathcal{G}_i = \begin{bmatrix} p_i(t_1) & p_i(t_2) & \cdots & p_i(t_m) \\ v_i(t_1) & v_i(t_2) & \cdots & v_i(t_m) \end{bmatrix},$$

where, $v_i(t)$ is the reported velocity vector at time t and $p_i(t)$ is the corresponding position of the reporting cell. From these vectors, the trajectory is derived by polynomial approximation. Namely, the trajectory was determined as an approximate polynomial whose values and gradients are similar to the data set.

B. Ongoing Processing

Let the k-th motion detection cell reports the average vector $v_k = \begin{bmatrix} v_{kx} \\ v_{ky} \end{bmatrix}$ and the time t^*. Let the location of the k-th motion detection cell be $p_k = \begin{bmatrix} p_{kx} \\ p_{ky} \end{bmatrix}$. The central processing unit starts to check the following inequalities to determine whether the new data belong to one of the trajectories.

$$\left| \begin{array}{c} g_{ix}(t^*) - p_{kx} \\ g_{iy}(t^*) - p_{ky} \\ \frac{d}{dt} g_{ix}(t^*) - v_{kx} \\ \frac{d}{dt} g_{iy}(t^*) - v_{ky} \end{array} \right| < \varepsilon, \qquad i = 1, \ldots, n \tag{3}$$

1) YES: If the above inequality is satisfied for j-th trajectory $g_j(t)$, then $g_j(t)$ and its data set $\mathcal{G}_j$ are

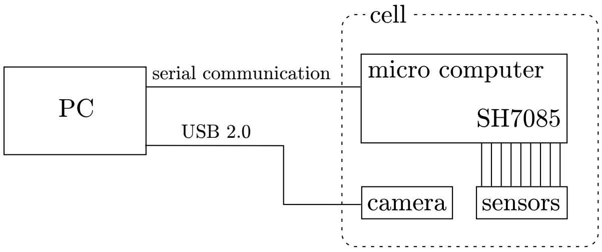

Fig. 11. System configuration.

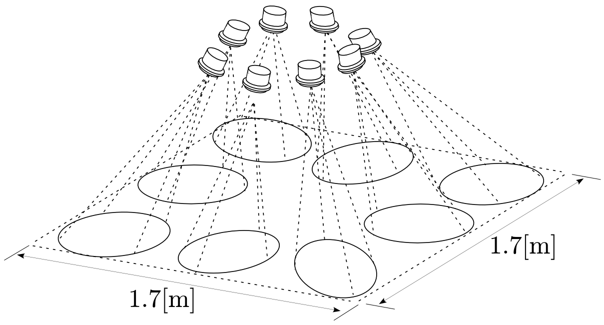

Fig. 12. Field of view of the motion detection cell in the experiment.

revised by the reported data p_k and v_k, by defining $p_j(t^*) = p_k$ and $v_j(t^*) = v_k$,

$$\mathcal{G}_j = \begin{bmatrix} p_j(t_1) & p_j(t_2) & \cdots & p_j(t_m) & p_j(t^*) \\ v_j(t_1) & v_j(t_2) & \cdots & v_j(t_m) & v_j(t^*) \end{bmatrix}.$$

Based on the above expanded basement data set, the revised trajectory $g_j(t)$ is derived by polynomial approximation.

2) NO: If the above inequality is never satisfied for any trajectory $g_i(t)$, $(i = 1, \ldots, n)$, the central processing unit adds one more new trajectory defined as follows.

$$g_{n+1}(t) = \begin{bmatrix} p_{kx} \\ p_{ky} \end{bmatrix} + (t - t^*) \times \begin{bmatrix} v_{kx} \\ v_{ky} \end{bmatrix} \quad (4)$$

And its data is given by defining $p_{n+1}(t^*) = p_k$ and $v_{n+1}(t^*) = v_k$,

$$\mathcal{G}_{n+1} = \begin{bmatrix} p_{n+1}(t^*) \\ v_{n+1}(t^*) \end{bmatrix}.$$

By doing the above procedure, the central processing unit always has new up to date trajectory set described as mathematical functions.

By using the above stated algorithm, we calculate the trajectories of the moving objects of the example shown in the section II. We restrict the class of continuous function to 2nd order polynomial function. The derived trajectories are shown as the solid lines in Figs. 9 and 10. The dashed lines are the actual trajectories of the human motion. Although these lines are not entirely identical, the estimated trajectories can express normal human motions enough.

If we can delete vectors included in the integral sets, from the vector field, the rest is set of noise vector or unidentified vectors. If the density or quantitative amount of such undefined vectors is not negligible, that means something unusual may happen. This criteria may be used for abnormal detection.

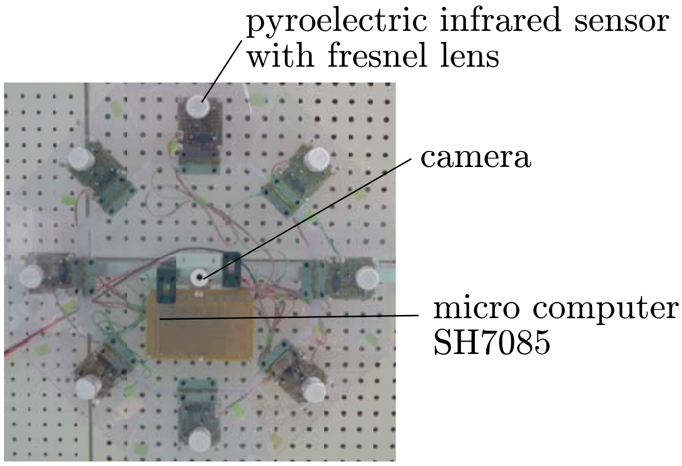

Fig. 13. Exterior view of motion detection cell.

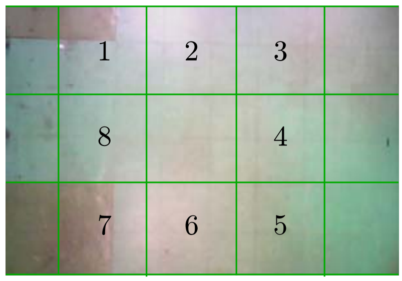

Fig. 14. A view from center of a cell and monitoring area of each sensor.

IV. Experimental Results

In this section, we describe the testbed as a motion detection cell in order to show the feasibility of our proposed method. Firstly, we show the system configuration. An output data of each pyroelectric infrared sensor is send to a PC via a microcomputer SH7085 as shown in Fig. 11. The microcomputer generates a vector components: the entry point, the exit point, and the time of each, from an output data by same manner as explained in the former section. From the these data from the microcomputer, the PC makes out a human motion as the vector.

Next, we show the experiment system. We configure a cell using eight pyroelectric infrared sensors and fresnel lens as RE210 by NIPPON CERAMIC CO., LTD. and MIL-100 by FUJI & CO., respectively. We close the space of sensors up to miniaturize a cell, and incline each sensor to adjust the monitoring area of a cell to about 1.7[m] $\times$ 1.7[m] from the ceiling of 3[m] height as shown in Fig. 12. Although we install the microcomputer SH7085 and the camera which projects human movement on the cell as shown in Fig. 13, the camera is only used in order to confirm human motions.

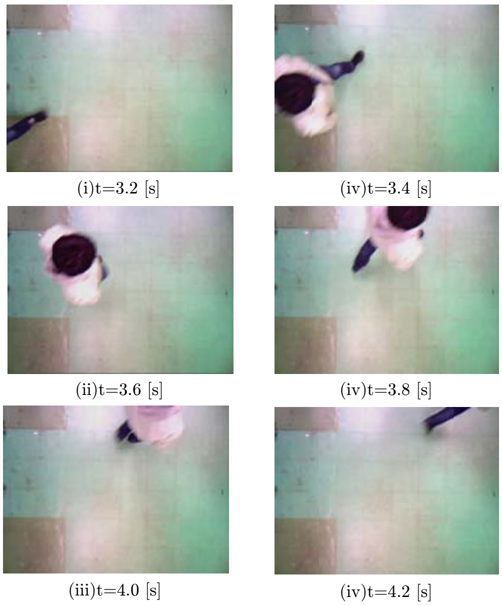

Fig. 15. Actual images sequence.

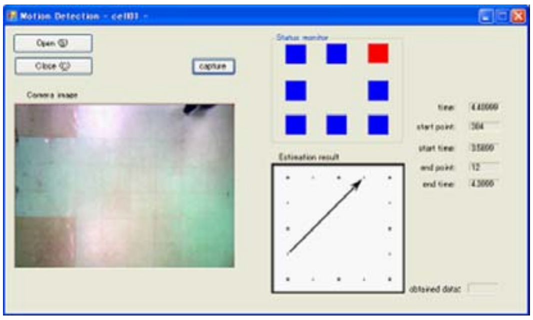

Fig. 16. A window by Windows Forms Application with Visual Studio.

Fig. 14 shows the capture image from the camera and the monitoring areas of each sensor.

Fig. 15 shows the actual image sequences of human walking for 7[s]. In the interval, there is a human moving the area diagonally from the bottom left to the top right. Fig. 16 shows the original software constructed on a personal computer by using Windows Forms Application with Visual Studio. The left monitor projects the capture image from the camera. The top right eight boxes show the status of the each sensor. If a sensor reacts, the corresponding box turns red. Bottom right plot the vector using the data from the microcomputer. The original movie of the software can be seen on the Web site [6]. From the human motion as shown in Fig. 15, we obtain the vector as shown in bottom right of Fig. 16. Although only one motion detection cell is produced in order to confirm the feasibility of our proposed method, we will experiment with a number of cells like a former simulation in our future work.

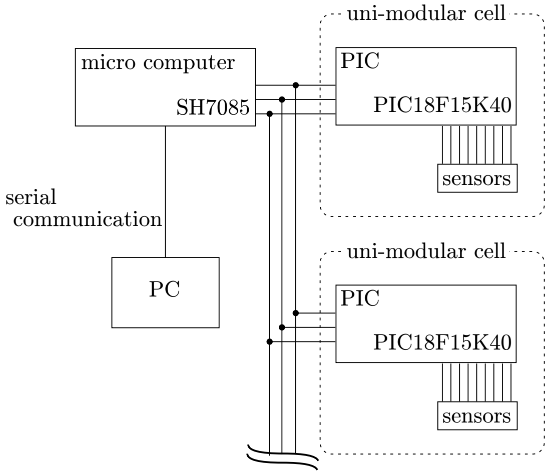

Fig. 17. System configuration with plural cells.

V. Conclusion

This paper has proposed networked cellular motion detection system by using pyroelectric infrared sensors with experimental results. The basic concept stated here may be utilized in various fields of information processing. Especially networked sensing systems are getting popular in coming several years; it must be crucial issue to process huge amount of information. The proposed concept may be a key hint to solve these difficulties. Although only one motion detection cell is used in order to confirm the feasibility of our proposed method in this paper, we will experiment with a number of cells like a former simulation in order to estimate human motion in our future work. Fig. 17 shows the system configuration with plural cells. We insert PIC(Programmable Interface Controller) microcomputer between eight sensors and microcomputer, because there is limited number of ports which communicate with sensors on SH7085. The microcomputer communicate with plural PIC microcomputers with I-squared-C communication as master.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, "Wireless Sensor Networks:a Survey," *Computer Networks*, Vol. 38, No. 4, pp. 393–422, 2002.
[2] C. Heneghan, S. M. Khanna, Å. Flock, M. Ulfendahl, L. Brundin and M. C. Teich, "Investigating the Nonlinear Dynamics of Cellular Motion in the Inner Ear Using the Short-Time Fourier and Continuous Wavelet Transforms," *IEEE Trans. on Signal Processing*, Vol. 42, No. 12, pp. 3335–3352, 1994.
[3] J. G. Nicholls, A. R. Martin, B. G. Wallace and P. A. Fuchs, *From Neuron to Brain* (4th ed.), Sinauer Associates, 2001.
[4] R. Żbikowski, "Fly Like a Fly," *IEEE Spectrum*, Vol. 42, No. 11, pp. 40–45, 2005.
[5] H. Kawai and H. Kobayashi, "Motion Detection with Networked Cellular Vision System for Surveillance," *Proc. of the 32nd Annual Conference of the IEEE Industrial Electronics Society*, pp. 3991–3996, 2006.
[6] http://www2.kanazawa-it.ac.jp/klab/member/2008/kanamaru/IECON 2010video.html