A Computer Vision Approach for Reducing Energy Losses

Ahmad Khan[1] (ahmadkhan@nwfpuet.edu.pk)

Muhammad Uzair[2] (uzair@ciitwah.edu.pk)

[1]University of Engineering & Technology Peshawar (Mardan Campus) Pakistan,
[2]COMSATS Institute of IT Wah Cantt. Pakistan.

***Abstract*— Saving electrical energy is one of the major requisites of the countries around the world, especially third world countries. In this paper we present a computer vision system for reducing electrical energy waste. The system senses the environment using computer vision and turns-off electrical appliances when they run needlessly. The absence of people is determined by detecting and tracking people in places like offices, classrooms or halls. To increase accuracy, the concept of data fusion is employed. Simple electronics sensors are used to cope with situations where only camera decision is not sufficient. We have tested our system in university campus and in the absence of people, it is capable of intelligently turning off electrical appliances to avoid waste of electrical energy: resulting in energy saving.**

I. INTRODUCTION

Electricity is a major energy source and countries around the world pay special attention on saving it, especially third world countries. In recent times, Pakistan is passing through serious electric power crises. One factor responsible for this crisis is the waste of electricity by home and office users. According to our observation in university campuses; the labs, class rooms, halls and offices are busy places and use electricity in one form or another during working hours. However; during break times or after office hours they become empty. In these times, some or all of the electric appliances still run needlessly. Similarly, at times employees or students leave offices, laboratories or halls forgetting to switch off electric appliances. In these cases electricity is needlessly consumed, sometimes for days. Situations like these contribute to a big extent in the prevailing energy crisis.

Computer vision can be a cost effective solution for minimizing such energy losses. Keeping in mind such aforementioned problem, we have designed a simple, efficient and inexpensive system which is capable to robustly switch-off the electrical appliance (light, fans, air-conditions etc.) in free hours or after working hours when the labs, halls, offices or class rooms are empty. An inexpensive camera along with efficient computer vision algorithm is a very effective way to detect presences or absences of peoples in such places.

The paper is organized as follows: Section two explains the proposed system, its components and working. Section three presents experimental results and section four concludes the paper.

II. PROPOSED METHOD

The outline of the proposed system is illustrated in Figure 1.

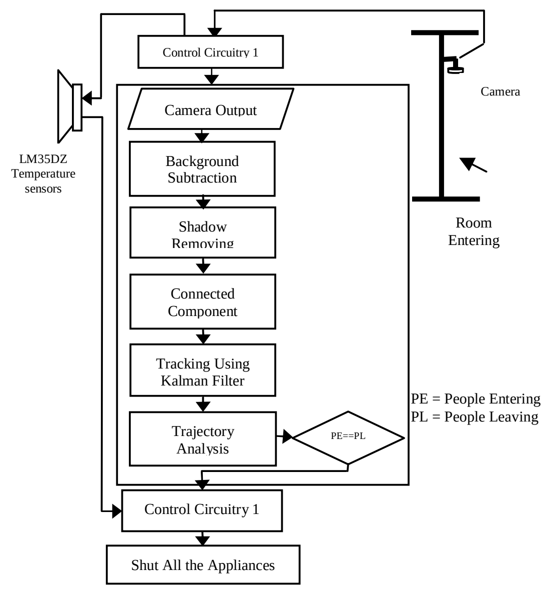

Figure 1. System Overview

Our method is based on decision fusion. Decision of turning off the appliances is based on data taken from still camera as well as from temperature sensors.

The camera detects and tracks moving objects using background subtraction and Kalman filter based tracking algorithms. Tracking results are analyzed to assess the number of entering and/or exiting people. In addition the results of camera are fused with electronic sensors in order to maintain the accuracy of the system in situation when images are not available from the camera. For example in cases when all lights are off and fans or air-conditions are running needlessly.

A. *Object Tracking and Trajectory Analysis*

For people detection and tracking we used Multimodal Mean Adaptive Background Subtraction algorithm [2]. Moving people are then tracked using Kalman Filter algorithm. The resulting trajectories are then analyzed to assess the number of people entering or leaving the vicinity.

- Backgound Subtraction

Background Subtraction is the first step in most of the applications where moving objects is to be detected. In this work we used a Multimodal Mean (MM) algorithm [2] for its real time performance and accuracy.

Let the pixel in a frame obtained from a video scene at time t is I_t, and $I_{t.x}$ represent its x-component, where x is red, green or blue color. For each pixel maintain a set of four cells (*K=4*) are maintained in the background. Each cell B_i is represented as three running sums for each color component $S_{i,t.x}$ and a count $C_{i,t}$ of how many times a matching pixel value has been observed in *t* frames. At any given frame *t*, the mean for any color component computed as follow.

$$\mu_{i,t.x} = S_{i,t.x} / C_{i,t} \tag{1}$$

A pixel I_t is considered to be a background pixel if it satisfy the condition given in equation 2, otherwise it is considered to be a foreground pixel.

$$\left(\hat{x} \left| I_{t,x} - \mu_{i,t-1.x} \right| \leq E_x \right) \wedge \left(C_i > T_{FG} \right) \tag{2}$$

Where E_x is threshold and T_{FG} is another small threshold number of times a pixel value can be seen and still considered to be foreground. In our system we use 30 for E_x is 30 and 3 for T_{FG}. When a pixel I_t is matched to any cell B_i, the cell is updated by adding each color component to the corresponding running sum $S_{i,t.x}$ and incrementing the count $C_{i,t}$. As the background gradually changes, due to lighting fluctuation, the running averages will adapt as well. In addition, to enable long-term adaptation of the background model, all cells are repeatedly *decimated* by halving both the sum and the count every *d* frames. When I_t matches a cell B_i, the cell is updated as follows:

$$S_{i,t.x} = \left(S_{i,t-1.x} + I_{t,x} \right) / 2^b \tag{3}$$

$$C_{i,t} = \left(C_{i,t-1} + 1 \right) / 2^b \tag{4}$$

Where $b = 1$ if t mod $d = 0$, and $b=0$, otherwise.

When a pixel I_t does not match cells at that pixel position, it is declared to be foreground. Then a cell is selected to be replaced according to the cell's overall count $C_{i,t}$ and a recency $R_{i,t}$, which measures how often the background cell's mean matched a pixel in a recent window of frames. A sliding window maintained by a pair of counts $(r_{i,t}, s_{i,t})$ in each cell B_i. At first $r_{i,t}$, starts at 0 and is incremented whenever B_i is matched, and is reset every *w* frames. The second $s_{i,t}$, simply holds the maximum value of $r_{i,t}$ computed in the previous window and $R_{i,t}$ is addition of $r_{i,t}$ and $s_{i,t}$. The cell to be replaced is selected from the subset of cells seen least recently, i.e., cells whose recency $R_{i,t} < w/K$. From this set, the cell with the minimum count $C_{i,t}$ is selected for replacement. If all cells have a recency $R_{i,t} > w/K$, then the cell with lowest $C_{i,t}$ is replaced. In our experiments, we chose $w = 32$.

- Shadow Detection and Elimination

After the input frame is segment into background and foreground, we define two classes for the foreground pixels, shadow or moving object, based on a threshold value. For the classification purpose we assumed that color of the shadow pixel remains unchanged and only the intensity becomes lowered. The intensity of the current pixel is computed and compared with the background pixel. It is considered to be shadow pixel if satisfied the given threshold other wise it is considered to be foreground pixel. The algorithm is given in [3].

- Connected Component Analysis

After object detection and shadow elimination the next step is to label the connected pixels for identification. Various algorithms exist for labeling like chain code or Fourier descriptor. We used region growing method for this purpose due to its speed and simplicity.

Region growing method is based on 8-connected neighbors, i.e. we move 3x3 pixel windows and check connected pixels in its neighborhood and label it if they satisfy the given conditions.

- Tracking

The main objective of tracking in our system is to trace the trajectory produced by moving people. The obtained trajectories are then analyzed for the presence of people in the vicinity. We used predictive Kalman Filter [4] to track the object based on its feature such as gray-scale mean, center position, corner features and bounding box.

The camera is placed in a position such as shown in Fig. 1. This arrangement has the advantage of avoiding the occlusion problem.

The Kalman filter is used estimate the next state the moving object.

$$x_k = A x_{k-1} + w_{k-1} \tag{5}$$

Equation 5 is used for calculating next state of the object, where x_k is the next priori state, A is the transition matrix or the system internal state matrix, x_{k-1} is the previous state matrix, w_{k-1} is the previous Gaussian Noise matrix.

After calculating the priori matrix then we find how much measurement we done in the position, which is calculating by using the following formula.

$$z_k = H x_k + v_k \tag{6}$$

Where z_k the measurement matrix of size 2x1 is, is the priori state, v_k is the measurement white noise and H relates the state to the measurement x_k. In practice might change with each time step or measurement, but here we assume it is constant.

After calculating the priori and measurement states we then calculate the predicted state which is the corrected state, by using the following formula.

$$\hat{x}_k = x_k + K\left(z_k - H x_k \right) \tag{7}$$

Where K is the Kalman gain and based on this we store

- Trajectory Analysis

The obtained trajectory or paths of each object corresponding to its label are then analyzed for checking the status of moving object, entering or leaving. At this moment simple directional information is used to check the status of the moving object. For both statuses we maintain two counters C_1 and C_2 respectively. Both the statuses are checked online for every frame. The vicinity is empty if both the counters have equal value. The equality condition is monitored for a fixed amount of time (25 minutes we choose for system). If there is no activity in the room for this much time the area is considered to be empty and the appliance turn off system is activated to turn of the needlessly running electrical appliances.

B. Improving Decisions Based on Temperature Sensor.

In order for our system to be effective in situations where images from camera are not available for decisions, the control circuitry-1 multiplexes the control to the temperature sensor. This situation is detected by calculating the average grey level of the input frames. A close to zero average grey level means that the lights are off. In this case the temperature sensors LM35DZ starting sensing. If the temperature is below certain threshold, our system switch-off all the appliances based on our assumption that at night the temperature falls by some degrees.

Microcontroller based control circuitry-2 is used to turn off the appliances based on the output obtained from either tracking phase or from sensor.

III. RESULTS

The developed system is tested in our campus labs and halls using P4 PC having 3.2 Intel processor and 1GB of RAM under Windows OS and using Microsoft Visual C++ 6.0 compiler with OpenCV.

Table 1 shows a one week summery of our measured electrical energy consumption in Digital Signal Processing Lab in our campus.

Table 1
Energy Consumption per week

Electrical Appliance	Quantity	Consumption in one week (KWH)
Fans	15	60
Tube Lights	10	20
Air Conditions	1	50
PCs	15	70

Figure 2 shows the comparison of energy consumption with our system installed and without our systems installed. We measured the electricity consumption of the lab for 13 weeks without our system installed and with our system installed It can be seen that our system is capable of saving ample amount of electrical energy by turning off the appliances needlessly running. Our system thus saves 1.49% of electrical energy per year in this laboratory. Similar results can be achieved in our campus faculty halls and offices.

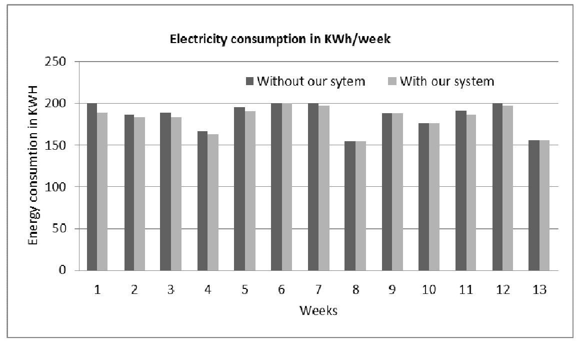

Figure 2. Energy consumption with and without our system installed

IV. CONCLUSION

In this paper we designed an efficient system for preventing energy losses using camera rather than using RFID sensors. The system accuracy depends on the tracking trajectories of moving objects for which we used statistical Kalman filters. In addition we used data fusion; besides taking data from camera we also take data from LM35DZ sensors in case when images from the camera are not available.

REFERENCES

[1] C. Stauffer, W. Grimson, “Learning Patterns of Activity using Real-time Tracking”, *IEEE Transactions on Pattern Analysis and Machine Intelligence, vol 22, no. 8, August 2000.*

[2] S. Apewokin, B. Valentine, L. Wills, A. Gentile, “Multimodal Mean Adaptive Backgrounding for Embedded Real-time Video Surveillance”, *Proc. IEEE CVPR, pp. 1-6, 2007.*

[3] Tang, Z. Miao, “Fast Background Subtraction and Shadow Elimination Using Improved Gaussian Mixture Modal”, *IEEE International workshop on Haptic, Audio and Visual Environments and Games, HAVE, pp. 38-41, 2007.*

[4] Greg Welch, Gary Bishop, “An Introduction to the Kalman Filter”, *University of North Carolina at Chapel Hill, Department of Computer Science. Chapel Hill, NC 27599-3175.*

[5] S.M. Smith and M. Brady, “SUSAN - a New Approach to Low Level Image Processing”, *International Journal of Computer Vision, Vol. 23(1), 45-78, 1997.*

[6] E. Persoon and K. S. Fu, “Shape Discriminations Using Fourier Descriptors”, *IEEE Trans. Systems, Man, and Cybernetics, 7(1977) 170-179.*

[7] QiZang and R. Klette, “Object Classification and Tracking in Video Surveillance”, *10th International Conference on CAIP, vol. 2756, pp. 198-205, 2003.*