

AUTOMATIC CAMERA TRACKING

Field of the Invention

5 The present invention relates to surveillance or video camera monitoring system. Typically, the camera system could automatically check for the presence of some pre-defined moving object, and upon such detection, zoom in for a close-up view, and track and follow the moving object, including making a video recording of the tracked object.

10

It's applications include situations in a classroom or lecture hall environment where the camera system, without the need of an operator could automatically zoom in, follow and track the lecturer or speaker including making a video recording of him as he moves around the classroom or lecture hall.

15

It could also be used for security surveillance applications within or outside building surrounding, where in the presence of moving objects (or suspicious persons or vehicles) appearing in the scene to be monitored, the camera could zoom in for a close up view, track and monitor the moving objects. In certain situations, multiple cameras
20 is required in order to cover all the scenes to be surveyed or monitored.

The same system could also be used for traffic monitoring where in the presence of either a speed or traffic red-light violation, the camera could zoom in, track the specific moving object including making a video recording of the object that had
25 caused the violation.

Background of the Invention

Current system in used, for example, as in security surveillance would have the
30 camera signals wired to a control room, where all the camera scenes could be remotely monitored by security personnel, including remotely controlling the cameras to check for suspicious scene or personnel.

Camera systems in used for lecture hall or seminar hall situation, would employ a camera man to manually control, zoom in, track and follow the lecturer or speaker including making the corresponding video recording.

- 5 For certain applications, it is required to have the automatic camera tracking of the object without the need of a human to control the camera to track and follow the moving object.

Existing state of art for automatic camera tracking system as in US. Pat. No.
10 5,572,317 employs an Infra-red (IR) transmitting device and IR receiving device to aid in tracking. A single camera is used and the concept is to tag the target object with an IR transmitting device, and to mount the IR receiving device on the camera system. As long as the IR receiving device is facing and in direct line-of-sight of the IR transmitting device, detection occurs. Therefore as the target object moves around
15 the scene, the camera system will track and follow the target object by turning the camera such that the IR receiving device mounted on the camera system receives the strongest IR signal from the IR transmitting device that is tagged to the moving object.

20 U.S. Pat. No. 5,525,967 requires two transceiver units where one is the base transceiver unit and another is the mobile transceiver unit that will be worn or tagged to the object to be tracked. The two transceivers will communicate through radio waves.

25 The problem with existing state of art system is, the object needs to be tagged with a transmitting device, which in actual classroom situation, the lecturer has to wear the IR device on the front and back of his body. This is to ensure that he could still be detected when he turns around to face the white board, with his back facing the camera.

30

However, if the lecturer were to turn side-ways, or if he were to move too fast to be out of detection angle of the IR receiving device, detection would be lost, and the

camera system would not know where the target object is, till he moves back into the detection area of the IR receiving device.

It is rather troublesome for the object to be tagged with such a transmitting device.

5 And many times, it is not possible to tag the object before hand, for example in traffic monitoring system where the system is looking out for a speeding or red-light traffic violation.

Also, for such existing system, you could only limit tracking to a single object based
10 on the tagged transmitting device.

The above problems could be solved if there is a camera system that could automatically locate and track the moving target, based on the video content of the camera view without the need of a tagged transmitting device.

15

Summary of the Invention

The invention discloses using a two camera system, the first is a fixed video camera (normally a wide angle camera) that covers the whole scene to be monitored, and a
20 second camera capable of pan, tilt, zoom (called a PTZ camera). The video signals from the fixed video camera can be fed to a computer or CPU and the video image signals can be analysed. As this camera is fixed, the scene is always stationary other than objects that moved. In such a case, it is relatively easy to detect the presence of moving objects. And the characteristics of each of these moving objects could be
25 further analysed for its position, colour, shape, size, motion value (or speed) and motion direction. Based on one or more pre-determined characteristics (for example to track a person with a red shirt, could mean to track a red moving object), the camera could be controlled to zoom in for a close-up view, track and follow a certain pre-determined object. Corresponding video recordings can also be made.

30

As processing of the image is done on the fixed wide angle camera, the system could identify all the moving objects in the whole scene to be monitored. And when two or

more objects were needed to be tracked that may be out of the zoomed camera view, the system could control the PTZ camera to zoom out accordingly such that these objects are visible within the PTZ camera view. This feature is not possible with a single camera system, as the single camera may be zooming into one specific object and the other object is not within its view.

The second aspect of the invention is to extend the two camera system to multiple camera system where there could be multiple fixed cameras to cover a bigger or different views and multiple tracking cameras to track different objects.

The third aspect of the invention is have an audio input device (such as a microphone or directional sound sensor) and analysing the characteristics of this sound to determine if the PTZ camera could be directed to point to the source of the sound.

The fourth aspect of the invention is to have a low power operation mode of the system. In the situation where low power operation is required, the whole system could be powered down or put into sleep mode, and periodically based on some schedule, the input monitoring device is turned on and the image is analysed. Only when the detected object characteristics met a certain pre-set criteria, would the whole system be turned on for full operation. In the sleep mode, the audio input could also be used as the trigger for turning on the system.

The camera system could also be simplified to a one camera tracking system, where the PTZ camera will do the dual function of gathering the video signals to be analysed by the computer, and at the same time, also be used for tracking the moving object. The limitation for such a simplified system is, there is no way to track object that is out of the camera view, though it may still remain within the scene to be monitored. And tracking would also be lost, if the moving object suddenly move too fast to be out of the PTZ camera view. The tracking would resume when the moving object move back within the camera view.

Other aspects of the invention are defined in the claims.

Brief Description of the Drawings

Figure 1 shows the front view of the camera setup.

5

Figure 2 shows the right side view of the camera setup.

Figure 3 shows the left side view of the camera setup.

10 Figure 4 shows the direction of the camera movement and the connection of the cables from and to the camera setup.

Figure 5 illustrates the communication between the cameras, pan-tilt head and the CPU.

15

Figure 6 illustrates the definition of overall area and selected region.

Figure 7 shows the flow chart of the tracking process with RETIRE mode switched off.

20

Figure 8 shows the flow chart of the tracking process with RETIRE mode switched on.

Figure 9 shows the flow chart of the detection process.

25

Figure 10 illustrates the detection of horizontal edges and the calculation of horizontal center.

Figure 11 illustrates the detection of vertical edges and the calculation of vertical center.

30

Figure 12 illustrates the conversion of signal from the camera to the CPU.

Figure 13 illustrates the control of microphone by the system.

Figure 14 shows the flow chart of the SLEEP and WAKE-UP mode operations.

5

Description of Preferred Embodiments and Best Mode

The operation of the two camera system would be explained here. The explanation can be extended to a multiple camera systems.

10

For the two camera system, the operation would be in the following steps:

- a) A fixed wide-angle video camera (called the primary camera) is used to capture the whole scene to be monitored.
- 15 b) A second video camera capable of pan, tilt and zoom (PTZ) functions herein called the PTZ camera or secondary camera, that is used to zoom in and track the moving object.
- 20 c) The input video image signals from the fixed wide-angle camera is fed to a computer or C.P.U to be analysed. The CPU would analyse and process the video image signals and detect for moving objects within the scene. The CPU would analyse and extract one or more of the object's characteristics such as position, colour, shape, size motion value and motion direction.
- 25 d) Based on some user selectable pre-determined criteria of the characteristics, the PTZ camera could be controlled to zoom in, track and follow a certain specific moving object when it appears. The video image signals from the fixed camera and the PTZ camera could correspondingly be recorded.
- 30 e) If more objects appear in the scene, though not within the view of the tracking PTZ camera, these objects would be still captured by the wide angle camera and analysed in the same way. If the additional object is also of interest, based on pre-determined

criteria, the PTZ camera could be controlled to zoom out in such a way as to capture the view of both the original object as well as the new object. Such criteria could be pre-configured and set by the system.

- 5 For the simplified single camera system, the PTZ camera would do the dual function of gathering the video signals of the scene and at the same time be used to track the moving object.

The operation of the multiple camera system is similar to that of the two camera
10 system where the input video signals could be from more than one fixed cameras to cover more views, and where more than one PTZ cameras could be used to track each specific objects.

A typical setup of the camera tracking system is shown in Figure 5. In the setup, 2
15 cameras are used. The first camera is the fixed camera known as the primary camera 5 while the second camera is the PTZ camera known as the secondary camera 1. The primary camera will be fixed and is usually placed directly along the same vertical axis as the secondary camera (see Figure 1). This primary camera is stationary and a low-cost camera could be used. The secondary camera could be better in terms of
20 quality and has a pan-tilt head 3 attached to the camera in order for the camera to pan and tilt and follow the movement of the target objects and to capture the images of the objects. The secondary camera is also equipped with a zoom lens 2 for it to capture objects from a distance. All the devices are placed on a base 6.

- 25 For effective alignment and result, the primary camera should be placed above or under the secondary camera. If the primary camera is not placed directly above or under the secondary camera and is placed somewhere else, adjustment would have to be made to compensate for the off-center placement of the primary camera.

30 There are four modes of operation for the camera, namely, pan, tilt, zoom and focus. A pan operation involving a camera means that the camera is able to pivot about the horizontal axis. A tilt operation means that the camera could rotate about the vertical

axis. A combination of pan and tilt operations will result in the camera moving left, right, up, down, and diagonal directions (see Figure 4). A zoom operation is when the camera produces images with the object getting bigger or smaller depending on the zooming mode used. A focus operation is when the object under observation become clearer or blurrer. The camera is capable of any of these operations or a combination of these operations.

Tracking of object will be started once the start command is given to the CPU to start capturing images and perform image processing (Figure 7). Initialization is done before the actual capturing 16. The system will first keep a copy of the first or current image in the memory 17 as the reference frame. The system will then continuously capture images 18. These later images are not stored. The latest image to be captured is the current frame.

The primary camera will capture the entire area 14. This captured image is called the overall area image. The dimension of this area is w_{overall} and h_{overall} for the width and height respectively. From this overall area image, a predefined area 15 could be use for the tracking. In a sense, this is restricting the tracking to a specific area instead of to the entire area captured by the primary camera. The width and height of this selected region is w_{region} and h_{region} respectively.

Figure 12 shows the process from the output of the camera 58 and the conversion of the analog signal 59 of the images to digital signal 61 through a analog-to-digital converter 60 that is required by the CPU 62 for further processing. After processing, the CPU will decide what command to issue to the pan-tilt head 63.

The first stage is to isolate the object from the background and this is described below.

In the detection of object, the current frame will be compared with the reference frame 19.

A binary frame 34 is setup which is made up of the values 0 and 1 (see Figure 9). A value of "1" will indicate than there is a difference between the reference frame and the current frame at the position (x,y). A value of "0" indicates that there is no difference between the reference frame and the current frame at the position (x,y).

- 5 This difference is set based on some threshold value. See explanation in subsequent paragraph. And (x,y) is the position in the frame with x representing the position in the horizontal position and y representing the position in the vertical position. A value in the reference frame at location (x,y) is represented by R(x,y). Same goes for the current frame, C(x,y), and binary frame, B(x,y).

10

The initialization stages also included the setting of detection width 35, detection height 36 and the setting of threshold of difference 37.

- Module 38, 39 and 40 checks for the availability of pels and getting the pel values from the reference frame and the current frame. A value of "1" in the binary frame is set if the following condition is met 41:

$$|R(x,y) - C(x,y)| > T$$

- 20 Where T is a pre-set threshold.

$$B(x,y) = 1 \text{ when } |R(x,y) - C(x,y)| > T$$

$$B(x,y) = 0 \text{ otherwise}$$

- 25 The shape and size of the object is defined in the binary frame. The surrounding of the object is marked as "0" 43 while the objects are marked as "1" 42.

The second stage is to find the edges of the object.

- 30 The left edge of the object is detected by finding the first occurrence of "1" pattern starting from the left of the binary frame 44 - 46. The same process is repeated for the right edge 47-49.

Same process apply to the search for top edge 51-53 and bottom edge 54-56.

In order to improve the accuracy of detection, instead of detecting the first occurrence of "1", a group of "1" is detected. Examples are listed here.

During experiments, it was concluded that the detection of 3 or 4 pels yielded the most appealing result.

10.

0	1	0
1	0	1
0	1	0

$$B(x,y) \&\& B(x+2,y) \&\& B(x+1,y-1) \&\& B(x+1,y+1) = 1$$

15

0	1	0
1	0	1
0	0	0

$$B(x,y) \&\& B(x+2,y) \&\& B(x+1,y-1) = 1$$

0	1	0
1	0	0
0	1	0

20

$$B(x,y) \&\& B(x+1,y-1) \&\& B(x+1,y+1) = 1$$

0	0	0
1	0	1
0	1	0

$$B(x,y) \&\& B(x+2,y) \&\& B(x+1,y+1) = 1$$

0	1	0
0	0	1
0	1	0

$$B(x,y-1) \&\& B(x,y+1) \&\& B(x+1,y) = 1$$

Result from the edge finding:

- 10 Left edge = x_L
 Right edge = x_R
 Top edge = y_T
 Bottom edge = y_B

- 15 With the left and right edges detected, the horizontal center 50 of the object could be calculated as follow:

$$\text{Horizontal center, } C_H = x_L + (x_L + x_R)/2$$

- 20 The vertical center 57 is calculated using the same formula:

$$\text{Vertical center, } C_V = y_T + (y_T + y_B)/2$$

- 25 Once the horizontal center and/or vertical center are found, the CPU can then proceed to the next stage.

The third stage is to send commands to the pan-tilt head to move the camera in the direction of the objects.

- 30 In order to prevent the movement of the camera in the event of small movement of the object and to improve the steadiness of the captured images, previous position of the camera is used to determine if the new position warrants the movement of the camera.

If current position, P_c , satisfy the following condition:

$$|P_c - P_p| < T_p$$

5

Where T_p is a preset threshold and P_p is the position for previous moved position.

If the above condition is met, the CPU will send commands to the pan-tilt head to move the camera to the position detected.

10

The camera system will also be equipped with an automatic SLEEP mode. When the target is out of the tracking range for a certain long period, it is to be assumed that the recording session is over. The system will go into SLEEP mode where the video camera will be turned off and the pan-tilt device will be turned off to reduce power consumption. This function could be turned off by the user if so desired.

15

When in SLEEP mode, the camera could be set to check for the target periodically. The user has to set the period where the camera will come out of the SLEEP mode and scan the tracking region to determine if the target is in it. If such target is detected, the system will turn on-line again and start tracking/recording. Again, the user can set this WAKE-UP mode and designate the period of WAKE-UP and SLEEP.

20

While in the SLEEP mode, the system could also be turn on by the presence of audio. Using a microphone or some other audio sensors as input device, upon detection of presence of audio signals, or when such audio levels exceed a certain minimum threshold, the system could be triggered to turn on.

25

This invention can be connected to a device whereby it will start the automatic recording of the audio from an attached microphone to the target or a microphone placed near to the target. The microphone will be powered on and off together with the invention.

30

The user can pass command to this invention either by sending command straight from the control panel attached to the CPU or through an external remote control device with such capabilities as but not limited to infra-red (IR). All the control on the control panel will be duplicated on the remote control device if so desired by the user.

5

The control panel will have access to such controls as the zoom, focus, speed, pan, tilt, power on and off, preset location, and etc.

Communication between the CPU and pan-tilt device will be through such standards as
10 but not limited to RS-232, RS-244, and ETHERNET.

Analog images from the cameras will be converted to digital signals by passing through an analog-to-digital converter before feeding the digital signal to the CPU. If a digital camera is used, the digital signal could be fed to the CPU through such
15 standards as IEEE1394.

This invention will have the ability to detect noise and to remove them if they exist in order to improve the quality of tracking and detection. Noise from the A/D converter could be filtered out by using threshold method in the CPU. The detection and
20 tracking could then be done on the filtered images. The filtered images will have distinctively reduce the occurrence of false detection of the targets.

The invention could perform self-diagnosis on the components of the system to determine if any part of the system has failed. Diagnostic such as the determination if
25 the pan-tilt function has stalled or the break-down of the camera could be reported to the user for further action. The diagnostic report could also be channelled to the vendor for servicing of the unit.

There will be a RESET button when the camera will completely reinitialize itself in
30 the event that certain unexpected malfunctions occurs. The function associated with this button is to power-off all components and wait for certain period before power-on all the components again.

There will also be HOME button where the camera will be positioned to its original predefined position (usually placed at the center of the horizontal axis and a preset vertical position). It is possible to for the system to treat the object located as this
5 position as the pre-defined object to be tracked. This could be user configurable. Also, in the situation where the camera could no locate the targets, this button could tell the camera system to start from this original position.

10 In order for the secondary camera to track accurately, the secondary camera will have to be calibrated to the primary camera. This could be done by the self-calibration function. It could be set to calibrate every time it is powered on or to calibrate within a certain period as defined by the user.

15 A display unit such as a monitor could be placed near the user for that person to assess the successful tracking of the system. If the users would want to make adjustment to the system, the accompanied remote control or control panel could be made available for adjustment to suit the requirements of the users.

20 If the invention is to be used in a dim or dark environment, light such as spot-light or torch light could be attached to the secondary camera to increase the brightness of the targets.

The user could specify the area where the tracking is to take place. By setting the limit perimeter, the users could limit the pan and tilt operation of the camera to within the
25 area of interest.

Apart from recording pictures, the invention could also snap still pictures and store in the hard disks of the CPU. With the functions of the camera (like zooming and focusing), the user could snap pictures of any sizes and distance. These pictures could
30 later be used as thumbnails or other means of archiving the stored video.

The camera has an OFFSET mode whereby the users could set the offset of the targets from the center of the images. This is usually done so as to capture the background of the target in situation where the background is important (e.g. the content of the blackboard behind a teacher). The offset could be specify in terms of height, width or
5 whether to place the offset on the left, right, top or bottom of the targets.

The user could also specify the size of the target. By supplying this information, this invention will automatically zoom in or out to comply with the size required. If during the recording session, the target move nearer to the system, the camera will zoom out
10 to reduce the target size. On the other hand, if the target move far away from the system, the camera will zoom in to increase the size of the target.

An OVERRIDE command is available to the user. This command will allow the user to seize control of the tracking device and the users could send commands to directly
15 control the movement of the camera.

To reduce the complexity of calculation, a sub-sample of the images could be taken. A sampling of 2 in the horizontal and vertical direction will cut the calculation to $\frac{1}{4}$ of the original if full sampling is used. By reducing the complexity of the calculation,
20 one can improve the responsiveness of the tracking system, or a lower CPU requirement could be used.

If the target is in the tracking region when the tracking is turned on, the image of the target will be in the reference frame and when a comparison of the current frame to
25 the reference frame is done, there will be two objects in the final detection, one of the current object position and the other one of the "ghost" object left behind because of the initialization. To solve this problem, a RETIRE mode is introduced. This mode will remove the target in the reference frame by removing the corresponding pels in the reference frame and thus eliminating the "ghost" object in the final detection. A
30 predefined period of the activation of the RETIRE function will be used.

The invention described here would allow the scene to be monitored and upon the detection of certain specific object, the camera system could zoom in for close up view, including tracking the moving object and making video recording, without the need of an operator. The system could be used in a classroom, to automatically track
5 and make a video recording of the lecturer as he moves around the classroom. It could also be used for traffic monitoring, to zoom in, track and make a video recording of vehicles that had caused a speed or traffic red light violation.

**CLAIMS:**

- (1) A system for the automatic tracking of moving objects, comprising :
a fixed video camera wherein the camera view of the said fixed video camera
5 covers the scene to be monitored;
a second video camera capable of pan, tilt and zoom functions (referred to as the PTZ camera);
a means for receiving input video image signals from the said fixed video camera;
10 a means for processing the said input video image signals to determine the presence of moving objects;
a means for determining the characteristics of each said moving objects based on the said processing, wherein the said characteristics comprise one or more of position, colour, shape, size, motion value and motion direction;
15 a means for making comparison of each said characteristics based on a set of pre-defined criteria;
a means for making indication of a specific object to be tracked based on the said comparison; and
a means for controlling the pan, tilt, zoom of the said PTZ camera to track the
20 said specific object.
- (2) The system according to claim 1, further comprising:
a means for making said indication of additional objects to be tracked based on the said comparison; and
25 a means for controlling the pan, tilt, zoom of the said PTZ camera for camera view of the said PTZ camera to include both the said specific object and the said additional objects.
- (3) A system according to claim 1 or 2, wherein the said fixed video camera
30 comprises a plurality of fixed video cameras whose combined camera view covers the scene to be monitored; and wherein the said PTZ camera comprises a plurality of PTZ cameras; and further wherein the said means for receiving input video image signals



comprises the means for receiving input video image signals from the said plurality of fixed video cameras; and yet further wherein the said means for controlling the pan, tilt, zoom comprises the means for controlling the pan, tilt, zoom of the said plurality of PTZ cameras for each said PTZ camera to be tracking each said specific object.

5

(4) A system according to claim 1, 2 or 3, further comprising a means for sending the video image signals from the said PTZ camera and the said fixed video camera to a TV monitor.

10

(5) A system according to claim 4, further comprising a means for recording the video image signals from the said PTZ camera and the said fixed video camera to a storage media.

15

(6) A system according to claim 5, further comprising an audio device or a microphone near the target; and a means for recording the audio to be recorded.

20

(7) A system according to claim 6, further comprising a means to power down the system to sleep mode, wherein during the said sleep mode, the said system could be periodically power-up based on a pre-defined schedule to check for the presence of moving objects; and
a means to start the said recording to the said storage media when the said characteristic of the said object correspond to the said pre-defined criteria.

25

(8) A system according to claim 7, further comprising:
a home button wherein upon the press or activation of the said home button, the said PTZ camera would be set to focus at its preset position; and
a means to obtain the characteristics of the object at the preset position; and
wherein the said characteristics of the said object would be used as the pre-defined criteria of the object to be tracked.

30

(9) A system for the automatic tracking of objects comprising :
a means for receiving input audio signals from the scene to be monitored;



- a means for processing the said input audio signals;
a means for determining the characteristics of the said audio signals;
a means for making comparison of each said characteristics based on a set of pre-defined criteria;
- 5 a means for making indication of the objects to be tracked based on the said comparison; and
 a means for controlling the pan, tilt, zoom of a PTZ camera to track the said objects.
- 10 (10) A system for the automatic tracking of moving objects based on a single camera comprising :
- a means for receiving input video image signals from a video camera wherein the said video camera is capable of pan, tilt and zoom functions, (referred to as the PTZ camera);
- 15 a means for processing the said input video image signals for determining the presence of moving objects;
- a means for determining the characteristics of each said moving objects based on the said processing wherein the said characteristics comprises one or more information such as position, colour, shape, size, motion value and motion direction;
- 20 a means for making comparison of each said characteristics based on a set of pre-defined criteria;
- a means for making indication of a specific object to be tracked based on the said comparison; and
 a means for controlling the pan, tilt, zoom of the said PTZ camera to focus and
- 25 track the said specific object.
- (11) A system according to claim 10, further comprising a means to send the said video image signals to a TV monitor.
- 30 (12) A system according to claim 11, further comprising a means to send the said video image signals for recording to a storage media.



(13) A system according to claim 12, further comprising:

a home button wherein upon the press or activation of the said home button, the said PTZ camera would be set to focus at its preset position; and

5 a means to obtain the characteristics of the object at the preset position wherein the said characteristics of the said object would be used as the pre-defined criteria of the object to be tracked.

(14) A system according to claim 8 or 13, further comprising:

10 a means for the system to perform self diagnosis of the functional operation of the said system; and

a means to record the results of the said self diagnosis.

(15) A method for automatically tracking moving objects, comprising the steps of: receiving input video signals from a fixed camera;

15 processing said signals to identify the presence of moving objects; and

controlling a second camera, having pan, tilt and zoom functions, to track an identified moving object.

(16) A system for the automatic tracking of moving objects, comprising:

20 a fixed video camera having a view of the scene to be covered;

a second video camera capable of pan, tilt and zoom functions; and

a processor for receiving and processing video signals from said fixed camera to identify the presence of a moving object therefrom, and to cause control of said second camera to track an identified moving object.

25