(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 December 2004 (29.12.2004)

PCT

(10) International Publication Number
WO 2004/114648 A2

(51) International Patent Classification[7]: H04N

(21) International Application Number:
PCT/US2004/019722

(22) International Filing Date: 21 June 2004 (21.06.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/479,950 19 June 2003 (19.06.2003) US

(71) Applicant *(for all designated States except US)*: **SARNOFF CORPORATION** [US/US]; 201 Washington Road, Princeton, New Jersey 08543 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **SAMARASEKERA, Supun** [US/US]; 24 Blackstone Drive, Princeton, New Jersey 08540 (US). **KUMAR, Rakesh** [IN/US]; 966 Ridge Road, Monmouth Junction, New Jersey 08852 (US). **HANNA, Keith** [GB/US]; 30 Scott Avenue, Princeton Junction, New Jersey 08544 (US). **SAWHNEY, Harpreet** [US/US]; 17 Melville Road, West Windsor, New Jersey 08550 (US). **ARPA, Aydin** [TR/US]; 1815 Ravens Crest, Plainsboro, New Jersey 08536 (US). **AGGARWAL, Manoj** [IN/US]; 4817 Ravens Crest Drive, Plainsboro, New Jersey 08536 (US). **PARAGANO, Vincent** [US/US]; 71 Bunker Hill, Lawrenceville, New Jersey 08648 (US).

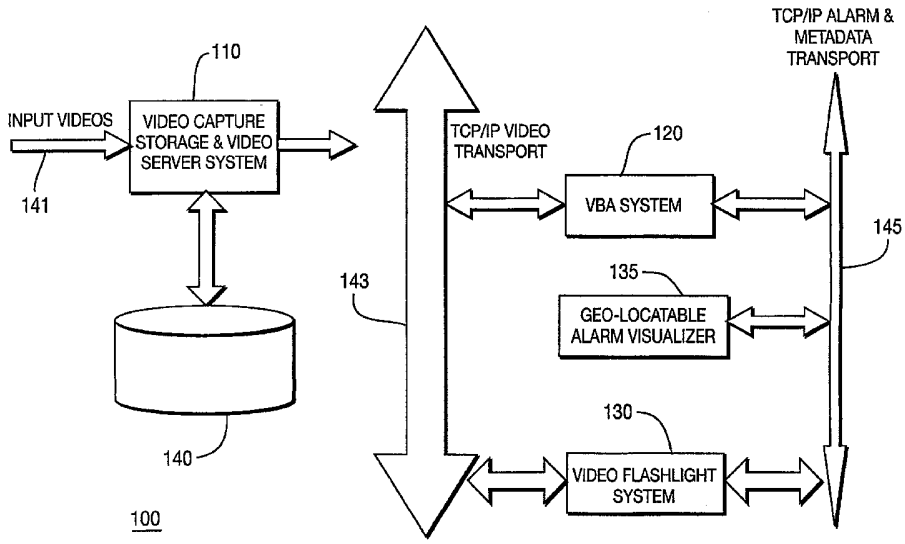
(74) Agent: **TONG, Kin-Wah**; Moser, Patterson & Sheridan, LLP, 595 Shrewsbury Avenue, Shrewsbury, New Jersey 07701 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,

[Continued on next page]

(54) Title: METHOD AND APPARATUS FOR PROVIDING A SCALABLE MULTI-CAMERA DISTRIBUTED VIDEO PROCESSING AND VISUALIZATION SURVEILLANCE SYSTEM



(57) Abstract: A scalable architecture for providing real-time multi-camera distributed video processing and visualization. An exemplary system comprises at least one video capture and storage system for capturing and storing a plurality of input videos, at least one vision based alarm system for detecting and reporting alarm situations or events, and at least one video rendering system (e.g., a video flashlight system) for displaying an alarm situation in a context that speeds up comprehension and response. One advantage of the present architecture is that these systems are all scalable, such that additional sensors (e.g., cameras, motion sensors, infrared sensors, chemical sensors, biological sensors, temperature sensors and like) can be added in large numbers without overwhelming the ability of security forces to comprehend the alarm situation.

WO 2004/114648 A2

ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

METHOD AND APPARATUS FOR PROVIDING A SCALABLE MULTI-CAMERA DISTRIBUTED VIDEO PROCESSING AND VISUALIZATION SURVEILLANCE SYSTEM

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims benefit of United States provisional patent application serial number 60/479,950, filed June 19, 2003, which is herein incorporated by reference.

BACKGROUND OF THE INVENTION

Field of the Invention

[0002] Embodiments of the present invention generally relate to image processing. Specifically, the present invention provides a scalable architecture for providing real-time multi-camera distributed video processing and visualization.

Description of the Related Art

[0003] Security forces at complex, sensitive installations like airports, refineries, military bases, nuclear power plants, train and bus stations, and public facilities such as stadiums, shopping malls, office buildings, are often hampered by 1970's-era security systems that do little more than show disjointed closed circuit TV pictures and the status of access points. A typical surveillance display, for example, is 16 videos of a scene shown in a 4 by 4 grid on a monitor. As the magnitude and severity of threats has escalated, the need to respond rapidly and more effectively to more complicated and dangerous tactical situations has become apparent. Simply installing more cameras, monitors and sensors will quickly overwhelm the ability of security forces to comprehend the situation and take appropriate actions.

[0004] The challenge is particularly daunting for sites that the Government must protect and defend. Merely asking personnel to be even more vigilant cannot reasonably guard enormous areas, ranging from army, air and naval bases to

extensive stretches of border. In addition, as troops deploy, new security personnel (e.g., reserves) may be utilized who are less familiar with the facility.

[0005] Therefore, there is a need for a method and apparatus for providing a scalable architecture for providing real-time multi-camera distributed video processing and visualization that can present an alarm situation to the attention of a security force in a context that speeds up comprehension and response.

SUMMARY OF THE INVENTION

[0006] In one embodiment, the present invention generally provides a scalable architecture for providing real-time multi-camera distributed video processing and visualization. An exemplary system comprises at least one video capture and storage system for capturing and storing a plurality of input videos, at least one vision based alarm system for detecting and reporting alarm situations or events, and at least one video rendering system (e.g., a video flashlight system) for displaying an alarm situation in a context that speeds up comprehension and response. One advantage of the present architecture is that these systems are all scalable, such that additional sensors (e.g., cameras, motion sensors, infrared sensors, chemical sensors, biological sensors, temperature sensors and like) can be added in large numbers without overwhelming the ability of security forces to comprehend the alarm situation.

[0007] To illustrate, the present invention outlines a highly scalable video rendering system, e.g., the Video Flashlight™ system that integrates key algorithms for remote immersive monitoring of a monitored site, area or scene using a blanket of video cameras. The security guard may monitor the monitored site or area using a live model, e.g., a 2D or 3D model, which is constantly being updated from different directions using multiple video streams. The monitored site or area can be monitored remotely from any virtual viewpoint. The observer can see the entire scene from far and get a bird's eye view or can fly/zoom in and see activity of interest up close. In one embodiment, a 3D-site model is constructed of the monitored site or area and used as glue for combining the multiple video streams. Each video stream is overlaid on top of the video model using the recovered camera pose. The

background 3D model and the recovered 3D geometry of foreground objects is used to generate virtual views of the scene and the various video streams are overlaid on top of it.

[0008] Coupling a vision based alarm system further enhances the surveillance capability of the overall system. Various alarm detection methods (e.g., methods that detect objects being left behind, methods that detect motion, methods that detect movement of objects against a preferred flow, methods that detect a perimeter breach, methods that count the number of objects and the like) can be deployed in the vision based alarm system. Upon detection of potential alarm situations, the vision based alarm system will report the alarm situations where the security guard will then employ the video rendering system to quickly view and assess the alarm situation.

[0009] Namely, the present invention provides tools that act as force multipliers, raising the effectiveness of security personnel by integrating sensor inputs, bringing potential threats to guards' attention, and presenting information in a context that speeds comprehension and response, and reduces the need for extensive training. When security forces can understand the tactical situation more quickly, they are better able to focus on the threat and take the necessary actions to prevent an attack or reduce its consequences.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] So that the manner in which the above recited features of the present invention can be understood in detail, a more particular description of the invention, briefly summarized above, may be had by reference to embodiments, some of which are illustrated in the appended drawings. It is to be noted, however, that the appended drawings illustrate only typical embodiments of this invention and are therefore not to be considered limiting of its scope, for the invention may admit to other equally effective embodiments.

[0011] Figure 1 illustrates an overall architecture of a scalable architecture for providing real-time multi-camera distributed video processing and visualization of the present invention;

[0012] Figure 2 illustrates a scalable system for providing real-time multi-camera distributed video processing and visualization of the present invention;

[0013] Figure 3 illustrates a plurality of software modules deployed within the video rendering or video flashlight system of the present invention;

[0014] Figure 4 illustrates a plurality of software modules deployed within the vision alert system of the present invention;

[0015] Figure 5 illustrates an illustrative system of the present invention using digital video streaming; and

[0016] Figure 6 illustrates an illustrative system of the present invention using analog video streaming.

[0017] To facilitate understanding, identical reference numerals have been used, wherever possible, to designate identical elements that are common to the figures.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

[0018] Figure 1 illustrates an overall architecture of a scalable architecture 100 for providing real-time multi-camera distributed video processing and visualization of the present invention. In one embodiment, an overall system may comprise at least one video capture storage and video server system 110, a vision based alarm (VBA) system 120 and a video rendering system, e.g., a video flashlight system 130 and a geo-locatable alarm visualizer 135.

[0019] In operation, a plurality of input videos 141 are received and captured by the video capture storage and video server system 110. In one embodiment, the input videos are time-stamped and stored in storage 140. The input videos are also provided to the vision based alarm (VBA) system 120 and the video rendering system 130 via a network transport 143, e.g., a TCP/IP video transport. In turn, a separate optional network transport 145, e.g., a TCP/IP alarm and metadata transport can be employed for forwarding and receiving alarm and metadata information. This second network transport increases robustness and provides a fault-tolerant architecture. However, the use of a separate transport is optional and is application specific. Thus, it is possible to implement the TCP/IP video transport and the TCP/IP alarm and metadata transport as a single transport.

[0020] In one embodiment, the geo-locatable alarm visualizer 135 operates to receive alarm signals, e.g., from the VBAs and associated meta-data, e.g.,

camera coordinates, or other sensor data associated with each alarm signal. To illustrate, if a VBA generates an alarm signal to indicate an alarm condition, the alarm signal may comprise a plurality of meta data, e.g., the type of alarm condition (e.g., motion detected within a monitored area), the camera coordinates of one or more cameras that are currently trained on the monitored area, other sensor metadata (e.g., detecting an infrared signal in the monitored area by an infrared sensor, detecting the opening of a door leading into the monitored area by a contact sensor). Using the alarm and metadata, the geo-locatable alarm visualizer 135 can integrate all the data and then generate a single view with the proper pose that will allow security personnel to quickly view and assess the alarm situations. For example, the geo-locatable alarm visualizer 135 may render annotated alarm icons, e.g., a colored box around an area or an object, on the alarm visualizer display. Additionally, the geo-locatable alarm visualiser can be used to control the viewpoint of the Video Flashlight system by a mouseclick on an alarm region, or by automatic analysis of the alarm and metadata information.

[0021] It should be noted that although the geo-locatable alarm visualizer 135 is illustrated as a separate module, it is not so limited. Namely, the geo-locatable alarm visualizer 135 can be implemented in conjunction with the VBA system or the video rendering system. In one embodiment disclosed below, the geo-locatable alarm visualizer 135 is implemented in conjunction with the video rendering system 130.

[0022] Effective video security and surveillance applications of the present invention need to handle hundreds and thousands of cameras with real-time intelligent processing, alarm and contextual video visualization, storage and archiving functions integrated in a system. The present invention is a scaleable real-time processing system that is unique in the sense that tens to hundreds to thousands of videos are continuously captured, stored, analyzed and processed in real-time, alerts and alarms are generated with no latency, and alarms and videos can be visualized with an integrated display of videos and 3D models and 2D iconized maps. The display management of thousands of cameras is managed by the use of a video switcher that selects which camera feeds to display at any one time, given the pose of the required viewpoint and the pose

of all the cameras. In one embodiment, the Video Flashlights / Vision-based Alarms (VF-VBA) system can typically process 1 Gbps to 1 Terra bits per sec. pixel data from tens of cameras to thousands of cameras using an end-to-end modular and scaleable architecture.

[0023] In one embodiment, as the number of cameras is increased, the present architecture allows deployment of a plurality of VBA systems. The VBA systems can be centrally located or distributed, e.g., deployed locally to support a set of cameras or even deployed within a single camera. Thus, each VBA or each of the video cameras may implement one or more smart image processing methods that allow it to detect moving and new objects in the scene and to recover their 3D geometry and pose with respect to the world model. The smart video processing can be programmed for detecting different suspicious behaviors. For instance, it can be programmed to detect left-behind objects in a scene, to detect if moving objects (people, vehicle) are present in a locale or are moving in the wrong or non-preferred direction, to count people passing through a zone and so on. These detected objects can be highlighted on the 3D model and used as a cue to the operator to direct his viewpoint. The system can also automatically move to a virtual viewpoint that best highlights the alarm activity.

[0024] Figure 2 illustrates a scalable system 200 of the present invention for providing real-time multi-camera distributed video processing and visualization. Specifically, Figure 2 illustrates an exemplary hardware implementation of the present system. However, since Figure 2 is only provided as an example, it should not be interpreted to limit the present invention in any way because many different hardware implementations are possible in view of the present disclosure or in response to different application requirements.

[0025] The scalable system 200 comprises at least one video capture storage and video server system 110, a vision based alarm (VBA) system or PC 120, at least one video rendering system, e.g., a video flashlight system or PC 130, a plurality of sensors, e.g., fixed cameras, pan tilt and zoom (PTZ) cameras, or other sensors 205, various network related components such as adapters and switches and input/output devices 250 such as monitors.

[0026] In one embodiment, the video capture storage and video server system 110 comprises a video distribution amplifier 212, one or more QUAD processors

214 and a digital video recorder (DVR) 216. In operation, video signals from cameras, e.g., fixed cameras and PTZ cameras are amplified by the video distribution amplifier 212 to ensure robustness of the video signal and to provide multiple distribution capability. In one embodiment, up to 32 video signals can be received and amplified, where up to 32 video signals can be distributed to the video flashlight PC and to the VBA PC 120 simultaneously.

[0027] In turn, the amplified signals are forwarded to QUAD processors 214 where the 32 video signals are reduced to 8 video signals. In one embodiment, four signals are reduced to one video signal, where the resulting signal may be a video signal having a lower resolution. In turn, the 8 signals are received and recorded by the DVR 216. It should be noted that the videos to the DVR 216 can be recorded and/or simply passes through the DVR to the video flashlight PC 130.

[0028] It should be noted that the use of the QUAD processors and the DVR is application specific and should not be deemed as a limitation to the present invention. For example, if a system is totally digital, then the QUAD processors and the DVR can be omitted altogether. In other words, if the video stream is already in digital format, then it can be directed red to the video flashlight PC 130.

[0029] The video flashlight PC 130 comprises a processor 234, a memory 236 and various input/output devices 232, e.g., video capture cards, USB port, network RJ45 port, serial port and the like. The video flashlight PC 130 receives the various video signals and is able to render one or more of the input videos over a model, e.g., a 2D or a 3D model of a monitor area. Thus, a user is provided by a real time view of a monitored area. Examples of a video rendering system or video flashlight system capable of applying a plurality of videos over a 2D and 3D model are disclosed in US patent applications entitled "Method and Apparatus For Providing Immersive Surveillance" with serial number 10/202,546, filed 7/24/02 with docket SAR 14626 and entitled "Method and Apparatus For Placing Sensors Using 3D Models" with serial number 10/779,444, filed 2/13/04 with docket SAR 14953, which are both herein incorporated by reference.

[0030] The vision alert PC or VBA 130 comprises a processor 224, a memory 226 and various input/output devices 222, e.g., video capture cards, Modular Input Output (MIO) cards, network RJ45 port, and the like. The vision alert PC 120 receives the various video signals and is able to one or more alarm or suspicious conditions. Specifically, the vision alert PC employs one or more detection methods (e.g., methods that detect objects being left behind, methods that detect motion, methods that detect movement of objects against a preferred flow, methods that detect a perimeter breach, methods that count the number of objects and the like). The specific deployment of a particular detection method is application specific, e.g., detecting a large truck in a parking lot reserved for cars may be an alarm condition, detecting a person entering a point reserved for exit only may be an alarm condition, detecting entry of an area after working hours may be an alarm condition, detecting a stationary object greater than a specified time duration within a secured area may be an alarm condition and so on.

[0031] Upon detection of potential alarm situations, the vision based alarm system 120 will report the alarm situations, e.g., logging the events into a file and/or forwarding an alarm signal to the video flashlight PC 130. In turn, a security guard will then employ the video rendering system to quickly view and assess the alarm situation.

[0032] Thus, a network switch 246 is in communication with the DVR 216, the video flashlight PC 130, and the vision based alarm system 120. This allows the control of the DVR to pass through current videos or to display previously captured videos in accordance with an alarm conditions or simply in response to a viewing preference of a security guard at any given moment.

[0033] Similarly, the system 200 employs an adapter 242 that allows the video flashlight PC 130 to control the cameras. For example, the PTZ cameras can be operated to present videos of a particular pose selected by a user. Similarly, the selected PTZ values can also be provided to a matrix switcher 244 where the selected pose will be displayed on one or more primary display monitors. In one embodiment, the matrix switcher 244 is able to select four out of 12 video inputs to be displayed. Thus, in addition to a render video stream provided by

the video flashlight PC, one can also see the full resolution videos as captured the cameras as well.

[0034] In one embodiment, various sensors 205 are optionally deployed. These sensors may comprise motion sensors, infrared sensors, chemical sensors, biological sensors, temperature sensors and like. These sensors are in communications with MIO cards on the vision alert PC 120. These additional sensors provide additional information or confirmation of an alarm condition detected by the vision alert PC 120.

[0035] Finally, an optional uninterruptible power supply (UPS) is also deployed. This additional device is intended to provide robustness to the overall system, where the loss of power will not interrupt the security function provided by the present surveillance system.

[0036] Figure 3 illustrates a plurality of software modules deployed within the video rendering system or video flashlight PC 130. The video flashlight PC 130 employs three software modules or applications: a 3-D video viewer or rendering application 310, a system monitor application 320, and an alarm visualizer application 330. Although the present invention is described illustratively with various software modules or sub-modules, the present invention is not so limited. Namely, the functions performed by these modules can be deployed in any number of modules depending on specific implementation requirements.

[0037] The 3-D video viewer or rendering application 310 comprises a plurality of software components or sub-modules: a video capture component 312, a rendering engine component 313, a 3-D viewer (GUI) 314, a command receiver component 315, a DVR control component 316, a PTZ control component 317, and a matrix switcher component 318. In operation, videos are received and captured by the video capture component 312. In addition to its capturing function, the video capture component 312 also time stamps the videos for synchronization purposes. Namely, since the module operates on a plurality of video streams, e.g., applying a plurality of video streams over a 3-D model, it is necessary to synchronized them for processing.

[0038] The rendering engine 313 is the engine that overlays a plurality of video streams over a model. Generally, the model is a 3-D model. However, there

might be situations where a 2-D or adaptive 3D model can be applied as well depending on the application. The 2-D model can be a plan layout of a building, for example. Video is shown in the vicinity of the camera location, and not necessarily overlaid on the model. In the adaptive 3D model, video is shown overlaid on the 3D model when the viewer views the scene from a viewing angle or pose that is similar to that of the camera, but is shown in the vicinity of the camera location if the viewing angle or pose is very dissimilar to that of the camera.

[0039] The 3-D viewer (GUI) 314 serves as the graphical user interface to allow control of various viewing functions. To illustrate, the 3-D viewer (GUI) 314 controls what videos will be captured by the video capture component 312. For example, if the user provides input indicative of a viewing preference pointing in the easterly direction, then videos from the westerly direction are not captured.

[0040] Additionally, the 3-D viewer (GUI) 314 controls what pose will be rendered by the rendering engine 313 by forwarding pose information (e.g., pose values) to the rendering engine 313. The 3-D viewer (GUI) 314 also controls the DVR 216 and PTZ cameras 205 via the DVR control component 316 and the PTZ control component 317, respectively. Namely, the user can select a recorded video stream in the DVR via the DVR control component 316 and control the pan, tilt and zoom functions of a PTZ camera via the PTZ control component 317. For example, a user can click on the 3-D model (e.g., in x,y,z coordinates) and the proper PTZ values will be generated, e.g., by a PTZ pose generation module and sent to the relevant PTZ cameras.

[0041] The commands receiver component 315 serves as a port to the alarm visualizer application 330, where a user clicking on the alarm browser 332 will cause the commands receiver component 315 to interact with the rendering engine component 313 to display the proper view. Additionally, if necessary, the commands receiver component 315 may also obtain one or more stored video streams in the DVR to generate the desired view if an older alarm condition is being recalled and viewed.

[0042] Finally, the 3-D viewer (GUI) 314 interacts with the matrix switcher control component 318 to obtain full resolution videos. Namely, the user can obtain the full resolution video from a camera output directly.

[0043] The alarm visualizer application 330 comprises a plurality of software components or sub-modules: an alarm browser (GUI) 332, an alarm status storage update engine component 334, an alarm status receiver component 336, an alarm status processor component 338 and an alarm status display engine component 339. The alarm browser (GUI) 332 serves as a graphical user interface to allow the user to select the viewing of various potential alarm conditions.

[0044] The alarm status receiver component 336 receives status for an alarm condition, e.g., as received by a VBA system or from an alarm database. The alarm status processor component 338 serves to mark whether an alarm is acknowledged and cleared or responded and so on. In turn, alarm status display engine component 339 will display the alarm conditions, e.g., in a color scheme where acknowledged alarm conditions are shown in a green color and unacknowledged alarm conditions are shown in a red color and so on. Finally, the alarm status storage update engine 334 is tasked with updating a system alarms database 340, e.g., updating the status of alarm conditions that have been acknowledged or responded. The alarm status storage update engine 334 may also update the alarm status on the vision alert PC as well.

[0045] In one embodiment, the system alarms database 340 is distributed among all the vision alert PCs 120. The system alarms database 340 may contain various alarm condition information, e.g., which vision alert PC reported an alarm condition, the type of alarm condition reported, the time and date of the alarm condition, health of any PCs within the system, and so on.

[0046] The system monitor application 320 comprises a plurality of software components or sub-modules: a system monitor (GUI) 322, a health status information receiver component 324, a health status information processor component 326 and a health status alarms storage engine component 328. In operation, the system monitor (GUI) 322 serves as a graphical user interface to monitor the health of a plurality of vision alert PCs 120. For example, the user can click on a particular vision alert PC to determine its health.

[0047] The health status information receiver component 324 operates to ping the vision alert PCs, e.g., periodically to determine whether the vision alert PCs are in good health, e.g., whether it is operating normally and so on. If an error

is detected, the health status information receiver component 324 reports an error for the pertinent vision alert PC.

[0048] In turn, the health status information processor component 326 is tasked with making a decision on the status of the error. For example, it can simply log the error via the health status alarm storage engine 328 and/or trigger various functions, e.g., direct the attention of the user that a vision alert PC is off line, schedule a maintenance request, and so on.

[0049] Finally, the video flashlight system 130 also employs a time synch module 342, e.g., a TARDIS time synch server. The purpose of this module is to ensure that all components within the overall system have the same time. Namely, the video flashlight PC and the vision alert PC must be time synchronized. This time consistency serves to ensure that alarm conditions are properly reported in time and that time stamped videos are properly stored and retrieved.

[0050] Figure 4 illustrates a plurality of software modules deployed within the vision alert system 120 of the present invention. The vision alert system 120 employs a vision alert application 410 that comprises a video capture component 411, a video alarms processing engine component 412, a configuration (GUI) 413, a processing (GUI) 414, a system health monitoring engine component 415, a video alarms presentation engine component 416, a video alarms information storage engine component 417 and a video alarms AVI storage engine component 418.

[0051] In operation, videos are received and captured by the video capture component 412. In addition to its capturing function, the video capture component 412 also time stamps the videos for synchronization purposes.

[0052] The video alarms processing engine component 412 is the module that employs one or more alarm detection methods that detect the alarm conditions. Namely, alarm detection methods such as methods that detect objects being left behind, methods that detect motion, methods that detect movement of objects against a preferred flow, methods that detect a perimeter breach, methods that count the number of objects and the like can be deployed in the video alarms processing engine component 412. The methods that will be selected and/or the thresholds set for each alarm detection method can be

configured using the configuration (GUI) component 413. In fact, configuration of which videos will be captured is also controlled by the configuration (GUI) component 413 as well.

[0053] The vision alert PC 120 employs one or more network transport, e.g., HTPP and ODBC channels for communications with other devices, e.g., the video flashlight system 130, a distributed database and so on. Thus, the system health monitoring engine component 415 serves to monitor the overall health of the vision alert PC and to respond to pinging from the system monitor application 320 via a network channel. For example, if the system health monitoring engine component 415 determines that one or more of its functions have failed, then it may report it as an alarm condition on the alarms information database 422.

[0054] The video alarms presentation engine component 416 serves to present an alarm condition over a network channel, e.g., via an IIS web server 420. The alarm condition can be forwarded to a video flashlight system 130. Additionally, the detection of an alarm condition will also cause the video alarms information storage engine 417 to log the alarm condition in the alarm information database 422. Additionally, the video alarms AVI storage engine 418 will also store a clip of the pertinent videos associated with the detected alarm condition on the AVI storage file 424 so that it can be retrieved later upon request.

[0055] In one embodiment, the processing (GUI) component can be accessed to retrieve the stored video clips that is stored in the AVI storage file. The forwarding of the stored video clip can be implemented manually, e.g., upon request by a user clicking on the alarm browser 332, or performed automatically, where certain types of important alarm conditions (e.g., perimeter breach) are such that the video clips are delivered automatically to the video flashlight system for viewing.

[0056] Finally, the video flashlight system 120 also employs a time synch module 426, e.g., a TARDIS time synch server. The purpose of this module is to ensure that all components within the overall system have the same time. Namely, the video flashlight PC and the vision alert PC must be time synchronized. This time consistency serves to ensure that alarm conditions are

properly reported in time and that time stamped videos are properly stored and retrieved.

[0057] The CORBA is a 3rd party networks communications program on top of which we have built functions that we use for sending real-time tracking positions, PTZ pose information across the network.

[0058] Figure 5 illustrates an illustrative system 500 of the present invention using digital video streaming, whereas Figure 6 illustrates an illustrative system 600 of the present invention using analog video streaming. These illustrative systems are examples of the general scalable architecture as disclosed above. Namely, the present architecture allows a system to easily scale up the number of sensors, video capture/compress stations, vision based alert stations, and video rendering stations (e.g., video flashlight rendering systems or dedicated alarm rendering systems). Namely, the present invention provides tools that act as force multipliers, raising the effectiveness of security personnel by integrating sensor inputs, bringing potential threats to guards' attention, and presenting information in a context that speeds comprehension and response, and reduces the need for extensive training. When security forces can understand the tactical situation more quickly, they are better able to focus on the threat and take the necessary actions to prevent an attack or reduce its consequences.

[0059] It should be understood that the various modules, components or applications as discussed above can be implemented as a physical device or subsystem that is coupled to a CPU through a communication channel. Alternatively, these modules, components or applications can be represented by one or more software applications (or even a combination of software and hardware, e.g., using application specific integrated circuits (ASIC)), where the software is loaded from a storage medium (e.g., a magnetic or optical drive or diskette) and operated by the CPU in the memory of the computer. As such, these modules, components or applications (including associated data structures) of the present invention can be stored on a computer readable medium or carrier, e.g., RAM memory, magnetic or optical drive or diskette and the like.

[0060] Although the present invention is disclosed within the context of a vision alert system, various embodiments of video rendering can be implemented that

are not in response to an alarm condition. For example, it is possible to deploy a very large number of cameras along a perimeter such that the video flashlight system is configured to provide a continuous real time "bird's eye view", "walking view" or more generically "virtual tour view" of the perimeter of a monitored area. For example, this configuration is equivalent to a bird flying along the perimeter of the monitored area and looking down. As such, as the view passes from one portion of the perimeter to another portion, the video flashlight system will automatically access the relevant videos from the relevant cameras (e.g., a subset of a total number of available videos) to overlay onto the model while ignoring other videos from other cameras. In other words, the subset of videos will be updated continuously as the view shifts continuously. Thus, it is possible to greatly increase the number of cameras without overwhelming the attention of the security staff.

[0061] While the foregoing is directed to embodiments of the present invention, other and further embodiments of the invention may be devised without departing from the basic scope thereof, and the scope thereof is determined by the claims that follow.

Claims:

1. A method for monitoring at least one scene, comprising:
receiving a plurality of input videos; and
applying selectively a subset of said plurality of input videos to a model of the at least one scene in response to a pose parameter.

2. The method of claim 1, wherein said pose parameter is selected in accordance with an alarm signal.

3. The method of claim 2, wherein said alarm signal is generated by an alarm detection method.

4. The method of claim 1, further comprising:
receiving signals from at least one sensor deployed within the at least one scene.

5. The method of claim 2, further comprising:
highlighting a portion of the at least one scene to indicate a location associate with said alarm signal.

6. The method of claim 1, wherein said subset of said plurality of input videos is continuously updated to provide a continuous virtual view of the at least one scene.

7. The method of claim 1, wherein said plurality of input videos are provided by a plurality of cameras, wherein at least one of said cameras has pan, tilt and zoom (PTZ) capability.

8. The method of claim 1, wherein said subset of said plurality of input videos is overlaid over said model.

9. An apparatus for monitoring at least one scene, comprising:

a plurality of cameras for providing a plurality of input videos; and

a video rendering system for applying selectively a subset of said plurality of input videos to a model of the at least one scene in response to a pose parameter.

10. The apparatus of claim 9, further comprising at least one vision based alarm system for generating an alarm signal, wherein said pose parameter is selected in accordance with said alarm signal.

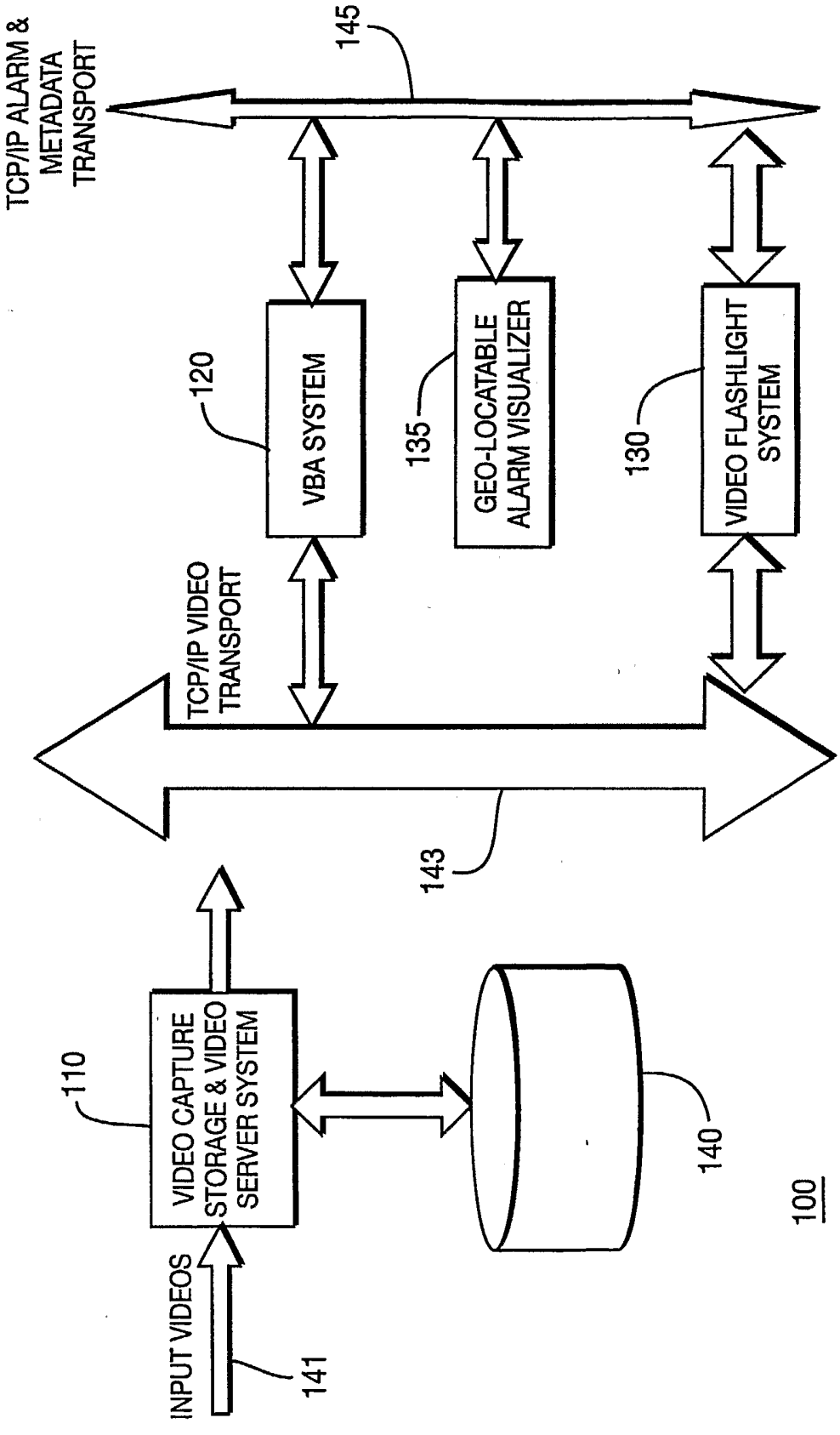

TCP/IP ALARM & METADATA TRANSPORT
145
VBA SYSTEM
120
GEO-LOCATABLE ALARM VISUALIZER
135
VIDEO FLASHLIGHT SYSTEM
130
TCP/IP VIDEO TRANSPORT
143
VIDEO CAPTURE STORAGE & VIDEO SERVER SYSTEM
110
140
INPUT VIDEOS
141
100

FIG. 1

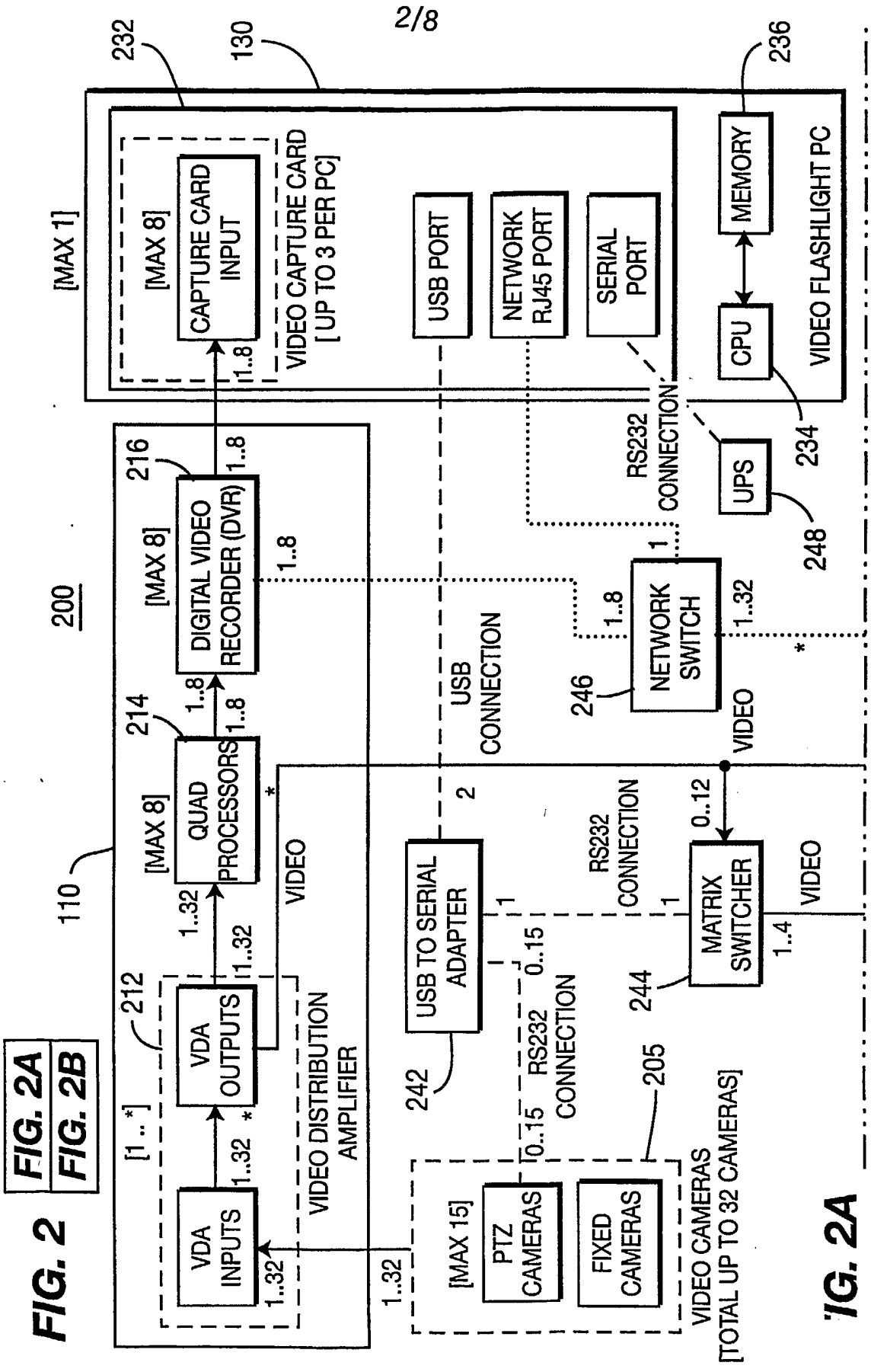


FIG. 2
FIG. 2A
FIG. 2B
200
110
[MAX 1]
232
130
236
[MAX 8]
CAPTURE CARD INPUT
1..8
VIDEO CAPTURE CARD [UP TO 3 PER PC]
USB PORT
NETWORK RJ45 PORT
SERIAL PORT
MEMORY
CPU
VIDEO FLASHLIGHT PC
RS232 CONNECTION
234
UPS
248
216
[MAX 8]
DIGITAL VIDEO RECORDER (DVR)
1..8
1..8
214
[MAX 8]
QUAD PROCESSORS
1..8
1..8
USB CONNECTION
246
NETWORK SWITCH
1
1..8
1..32
VIDEO
*
212
[1 .. *]
VDA OUTPUTS
1..32
*
1..32
VIDEO
VIDEO DISTRIBUTION AMPLIFIER
USB TO SERIAL ADAPTER
2
242
1
0..15
RS232 CONNECTION
RS232 CONNECTION
1
MATRIX SWITCHER
0..12
1..4
VIDEO
244
205
VDA INPUTS
1..32
1..32
[MAX 15]
PTZ CAMERAS
FIXED CAMERAS
VIDEO CAMERAS [TOTAL UP TO 32 CAMERAS]

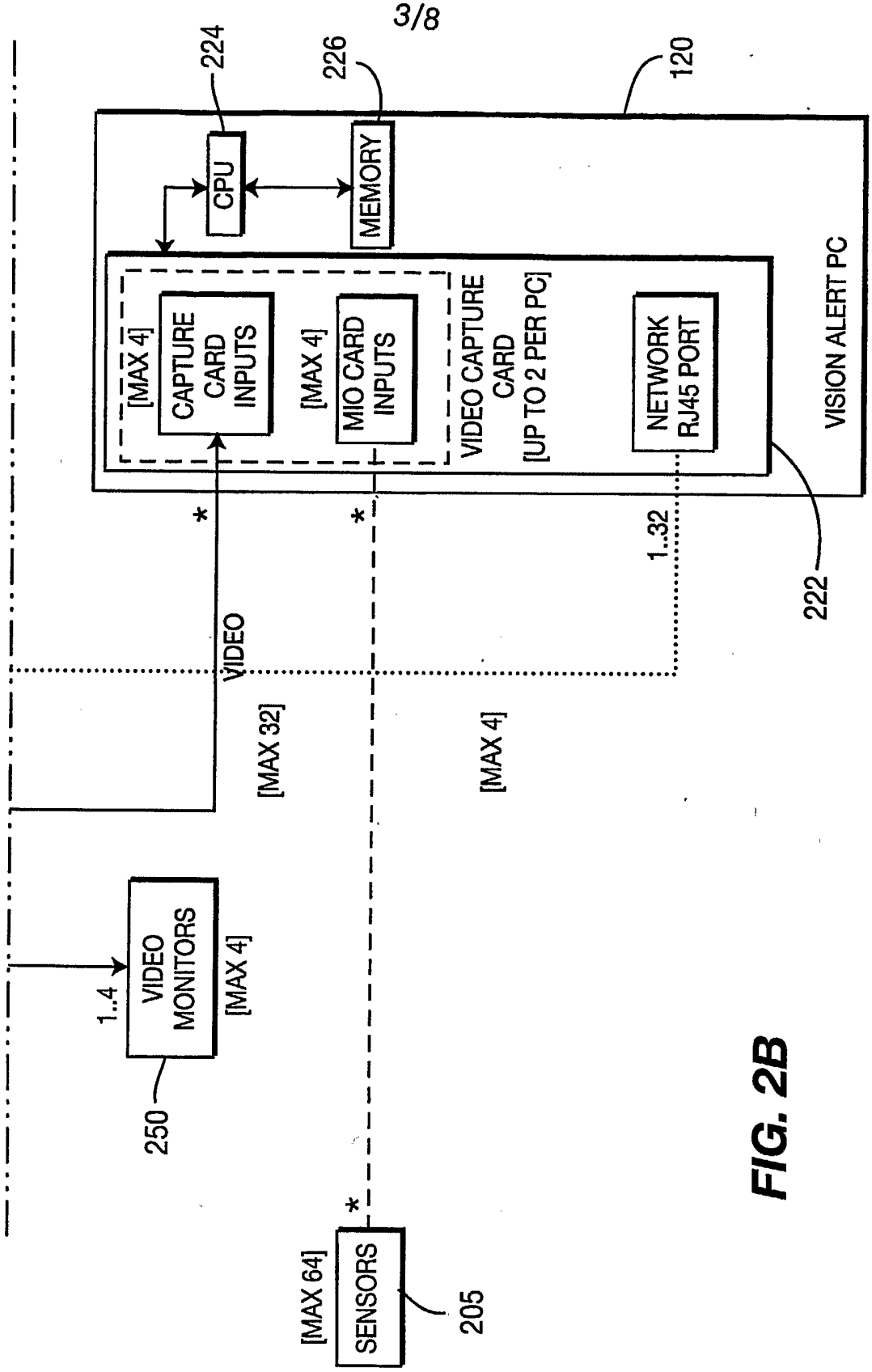
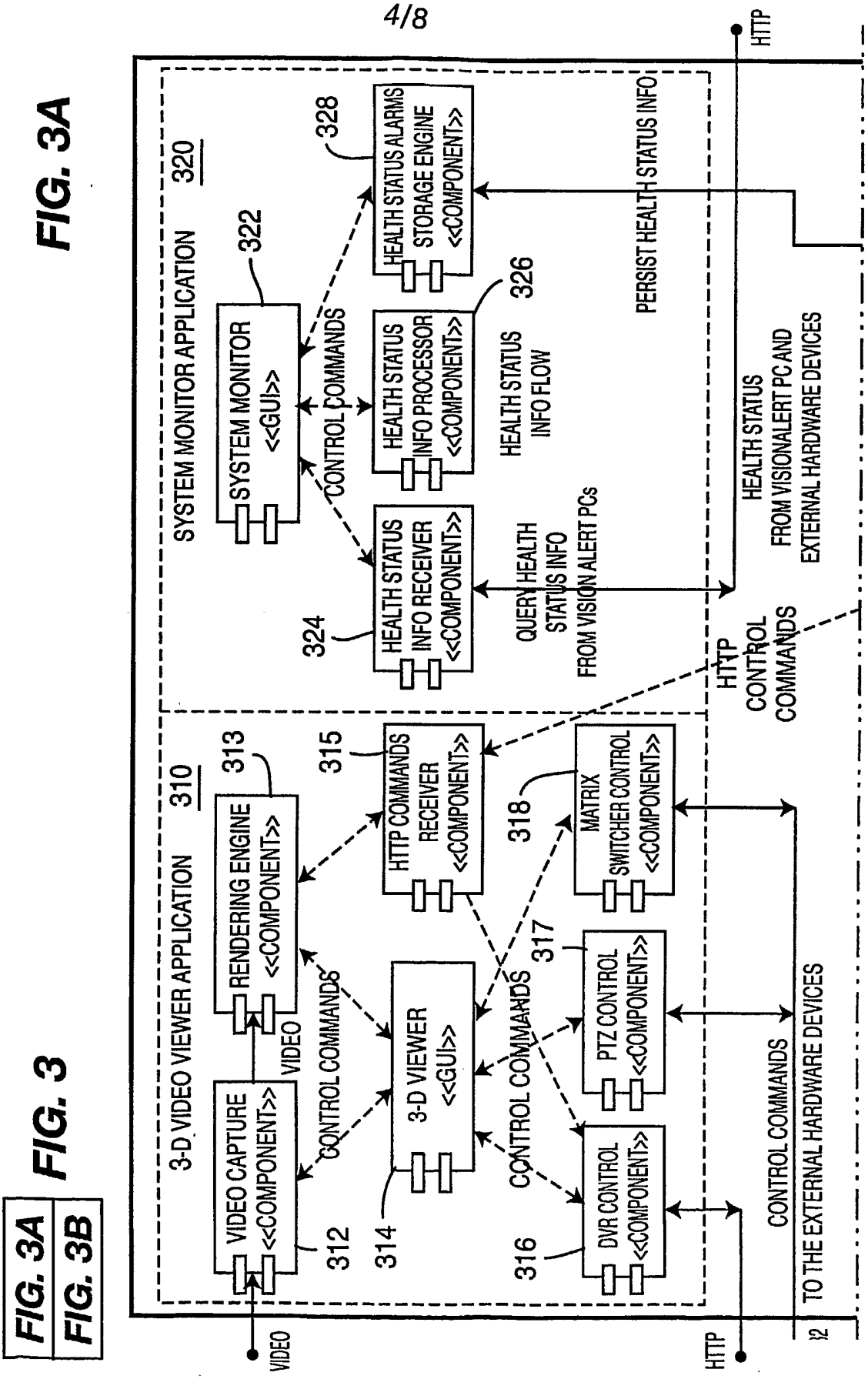

224
226
120
CPU
MEMORY
[MAX 4]
CAPTURE CARD INPUTS
[MAX 4]
MIO CARD INPUTS
VIDEO CAPTURE CARD [UP TO 2 PER PC]
NETWORK RJ45 PORT
VISION ALERT PC
222
1..32
VIDEO
[MAX 32]
[MAX 4]
1..4
VIDEO MONITORS
[MAX 4]
250
[MAX 64]
SENSORS
205

FIG. 2B

FIG. 3A

FIG. 3

FIG. 3A
FIG. 3B



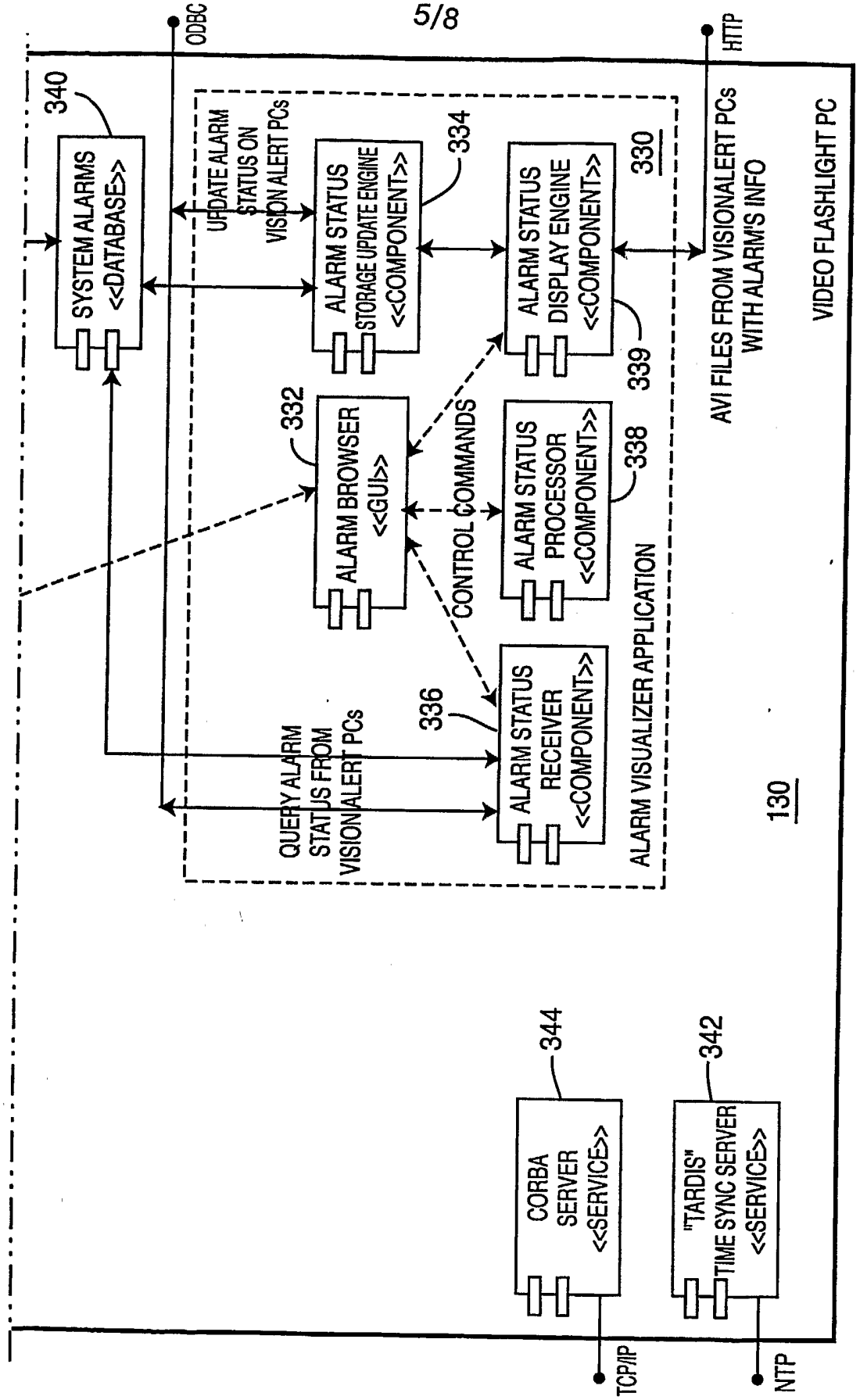

ODBC
HTTP
340
SYSTEM ALARMS
<<DATABASE>>
UPDATE ALARM STATUS ON VISION ALERT PCs
ALARM STATUS STORAGE UPDATE ENGINE
<<COMPONENT>>
334
ALARM STATUS DISPLAY ENGINE
<<COMPONENT>>
330
339
AVI FILES FROM VISIONALERT PCs WITH ALARM'S INFO
VIDEO FLASHLIGHT PC
332
ALARM BROWSER
<<GUI>>
CONTROL COMMANDS
ALARM STATUS PROCESSOR
<<COMPONENT>>
338
ALARM VISUALIZER APPLICATION
336
ALARM STATUS RECEIVER
<<COMPONENT>>
QUERY ALARM STATUS FROM VISIONALERT PCs
130
344
CORBA SERVER
<<SERVICE>>
342
"TARDIS" TIME SYNC SERVER
<<SERVICE>>
TCP/IP
NTP

FIG. 3B

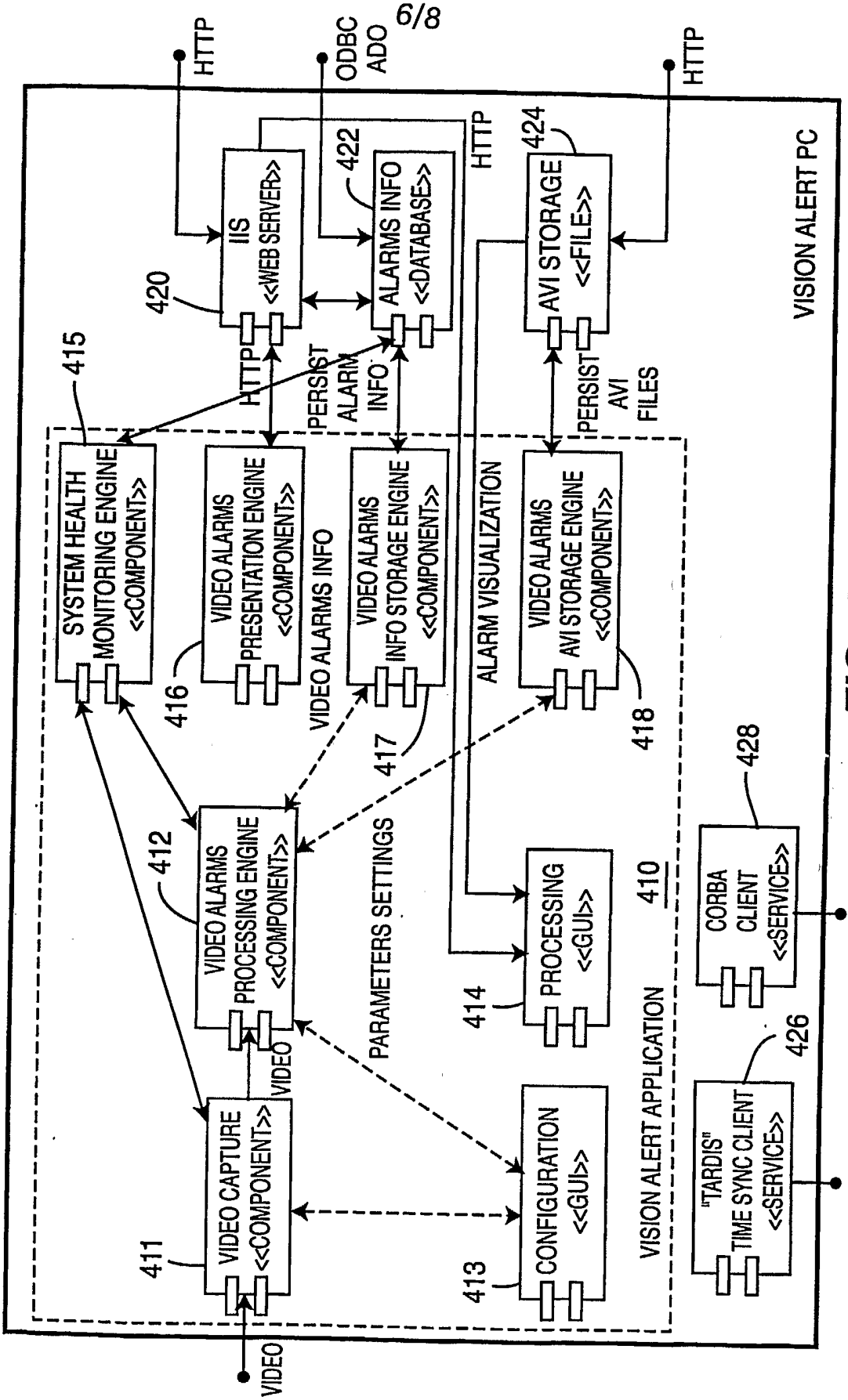

HTTP
ODBC ADO
HTTP
VISION ALERT PC
IIS <<WEB SERVER>>
ALARMS INFO <<DATABASE>>
AVI STORAGE <<FILE>>
420
422
424
HTTP
PERSIST ALARM INFO
PERSIST AVI FILES
415
SYSTEM HEALTH MONITORING ENGINE <<COMPONENT>>
VIDEO ALARMS PRESENTATION ENGINE <<COMPONENT>>
VIDEO ALARMS INFO STORAGE ENGINE <<COMPONENT>>
VIDEO ALARMS AVI STORAGE ENGINE <<COMPONENT>>
416
VIDEO ALARMS INFO
417
ALARM VISUALIZATION
418
412
VIDEO ALARMS PROCESSING ENGINE <<COMPONENT>>
PARAMETERS SETTINGS
410
VIDEO
414
PROCESSING <<GUI>>
CORBA CLIENT <<SERVICE>>
428
TCP/IP
VIDEO CAPTURE <<COMPONENT>>
411
CONFIGURATION <<GUI>>
413
VISION ALERT APPLICATION
"TARDIS" TIME SYNC CLIENT <<SERVICE>>
426
NTP
VIDEO

FIG. 4

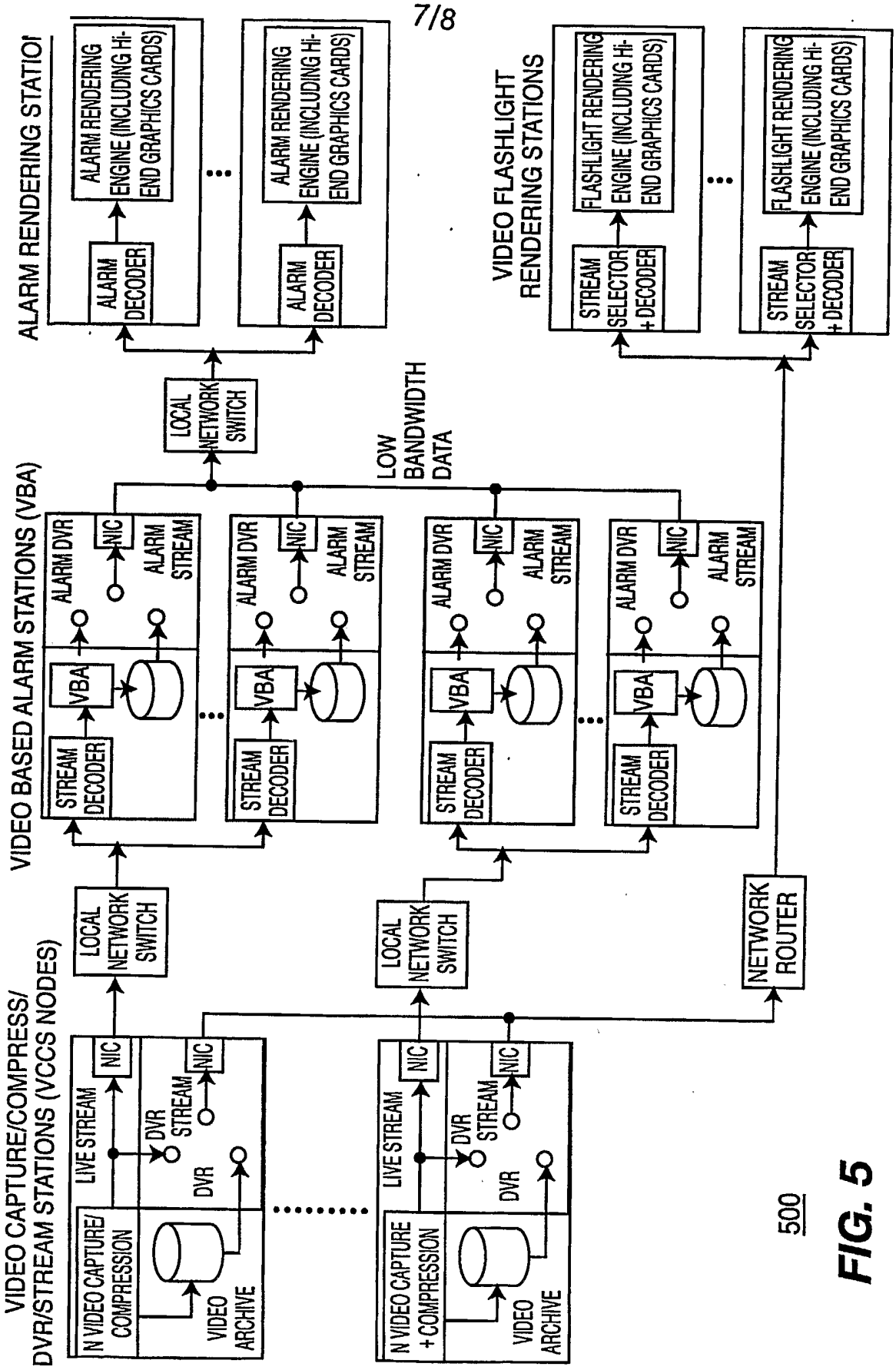

ALARM RENDERING STATIOI
ALARM RENDERING ENGINE (INCLUDING Hi-END GRAPHICS CARDS)
ALARM DECODER
VIDEO FLASHLIGHT RENDERING STATIONS
FLASHLIGHT RENDERING ENGINE (INCLUDING Hi-END GRAPHICS CARDS)
STREAM SELECTOR + DECODER
LOCAL NETWORK SWITCH
LOW BANDWIDTH DATA
VIDEO BASED ALARM STATIONS (VBA)
ALARM DVR
NIC
ALARM STREAM
VBA
STREAM DECODER
NETWORK ROUTER
VIDEO CAPTURE/COMPRESS/ DVR/STREAM STATIONS (VCCS NODES)
LIVE STREAM
DVR STREAM
DVR
N VIDEO CAPTURE/ COMPRESSION
N VIDEO CAPTURE + COMPRESSION
VIDEO ARCHIVE
500

FIG. 5

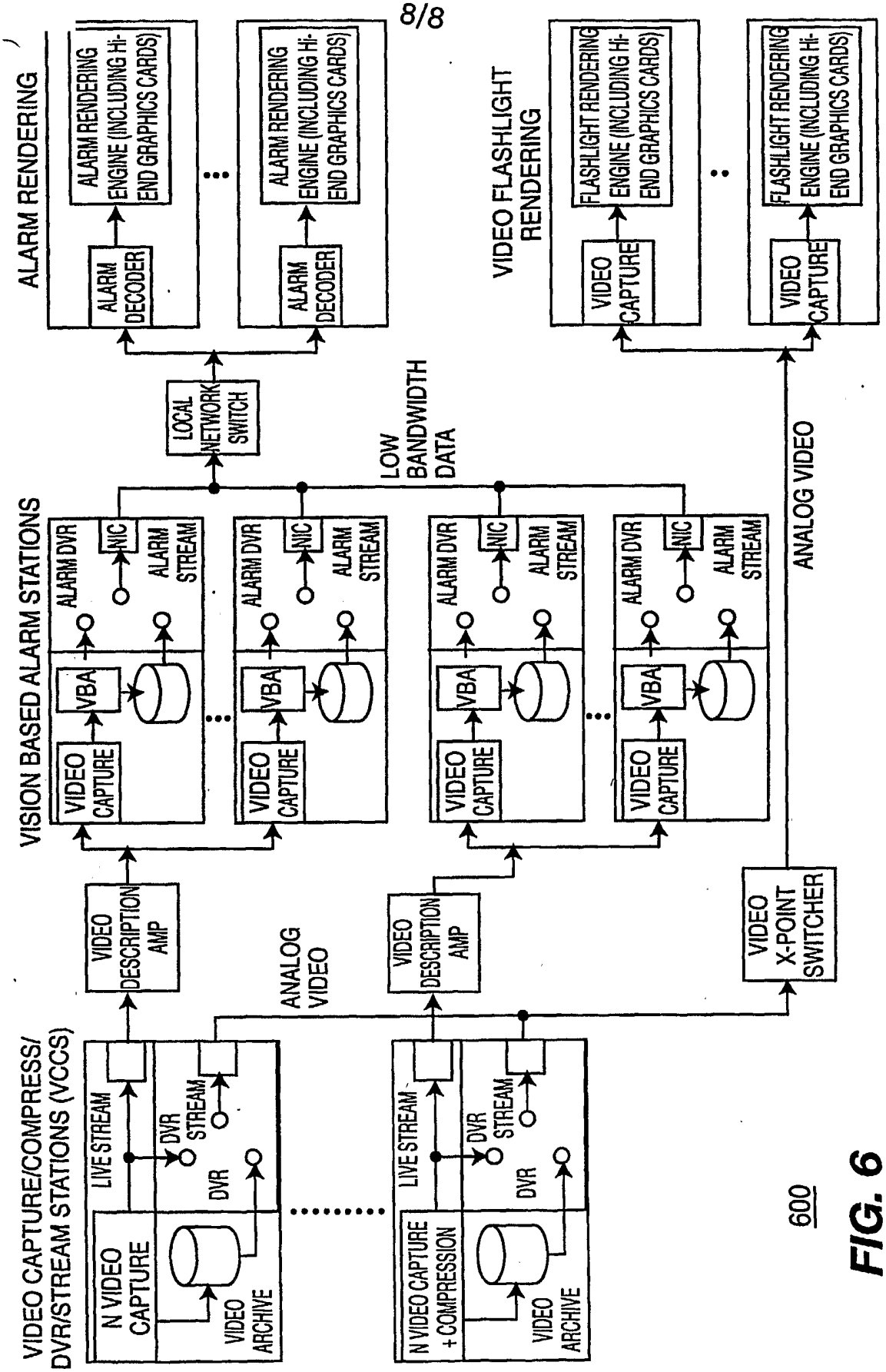

ALARM RENDERING
ALARM RENDERING ENGINE (INCLUDING HI-END GRAPHICS CARDS)
ALARM DECODER
VIDEO FLASHLIGHT RENDERING
FLASHLIGHT RENDERING ENGINE (INCLUDING HI-END GRAPHICS CARDS)
VIDEO CAPTURE
LOCAL NETWORK SWITCH
LOW BANDWIDTH DATA
ANALOG VIDEO
VISION BASED ALARM STATIONS
ALARM DVR
NIC
ALARM STREAM
VBA
VIDEO CAPTURE
VIDEO DESCRIPTION AMP
VIDEO X-POINT SWITCHER
VIDEO CAPTURE/COMPRESS/ DVR/STREAM STATIONS (VCCS)
N VIDEO CAPTURE
N VIDEO CAPTURE + COMPRESSION
LIVE STREAM
DVR STREAM
DVR
VIDEO ARCHIVE
600

FIG. 6