

SYSTEM AND METHOD FOR PARALLEL IMAGE PROCESSING**INVENTORS:**

Dario Berini
20740 Rainsboro Drive
Ashburn, VA 20147 USA
Citizen of: USA

Simion Adrian Moica
1765 Matthews Avenue,
Ottawa, Ontario,
K2C 0X5, Canada
Citizen of: Canada

Prateek Srivastava
47 Wycliffe Street
Ottawa, Ontario,
K2G 5L9, Canada
Citizen of: Canada

Gary Van Beek
3065 Mackey Road.
P.O. Box 262.
North Gower, Ontario,
K0A 2T0, Canada
Citizen of: Canada

Hooman Dadrassan
298 Sandhill Road
Ottawa, Ontario
K2K 2V7
Citizen of: Canada

Bryon Fevens
38 Stinson Ave.
Nepean, ON
K2H 6N3
Citizen of: Canada

ASSIGNEE:

CryptoMetrics, Inc.
160 Michael Cowpland Drive
Ottawa, ON K2M 1P6

Attorney:

Greenberg Traurig, LLP
1750 Tysons Boulevard, 12th Floor
McLean, VA 22102
(703) 749-1300

SYSTEM AND METHOD FOR PARALLEL IMAGE PROCESSING

[0001] This application claims priority from U.S. Provisional Patent Application Serial Number 60/864,840 entitled “System and Method for Processing Images From a Large Number of Disparate Cameras” filed November 8, 2006, which is incorporated herein by reference in its entirety.

BACKGROUND

[0002] There are numerous applications where video images are taken and recorded. Some of these applications involve the recording of video while a transaction is taking place, e.g., at an ATM or at a bank counter. The use of video recording is anticipated to increase significantly in the immediate future, such as in shopping centers, aboard buses and trains, and the like. Digital recording of video takes enormous amounts of recording space despite compression techniques such as MPEG, the use of slow video acquired at several frames per second rather than at a full 30 frames-per-second, and reductions in resolution. As a result, the recording times of digital recorders with multiple video inputs are still limited despite the use of large Giga and Terra Byte storage devices.

[0003] For some time now, there have been market available machine vision cameras and systems that can be programmed to detect certain geometrical objects. In general these objects have a very simple geometry such as nuts, bolts, engine parts, etc. In the case of face finding and processing the current state of the art is a slow process that requires massive computing power and hardware, often resulting in a system too complex to be reliable and manageable. Such a system, since requiring many components to be feasible, is difficult to deploy and scale.

[0004] A real-time processing system built with the current state-of-the-art would be cost prohibitive; as a compromise, system architects of these systems often trade-off complexity for performance. This typically results in small systems processing recorded images. Such systems are slow and incapable of processing images in real time. An improved system for image recognition is highly desirable.

SUMMARY OF THE INVENTION

[0005] In one embodiment, the invention provides a system for processing images, the system including a plurality of image providers configured to transmit images. A plurality of destination processors receives the transmitted images and

transforms the transmitted images to internally useable image data. A plurality of feature object engines finds and identifies in the internally useable image data a plurality of objects. A plurality of object classifier engines index and classify the plurality of objects found by the feature object engines.

[0006] In another embodiment, the invention is a method for processing images. Images are transmitted to a destination processor. The received images are transformed to internally usable data. A plurality of objects is identified in the internally useable image data.

[0007] In yet another embodiment, the invention is an image source comprising an image provider configured to transmit an image to an image enhancement device and an image enhancement device configured to add additional information to the image and to retransmit the image.

DETAILED DESCRIPTION

[0008] In one embodiment, the invention is directed to a system for processing multiple image sources in real time as static data or streams, multiple recorded images as static data or streams, media files, or any combination of the above. The system may classify multiple objects in the image concurrently based on their features and generate feature based indexing results. The system may classify and index the found objects in the images based on their closeness to the feature set used to search. The system may search for multiple, uncorrelated patterns or supersets of features concurrently. For example, the system may be provided with the capability to search concurrently for multiple patterns in the same image at the same time by distributing the processing load among its processing units such that specific units are searching for specific patterns at a given time.

[0009] The system may be implemented such that it is capable of changing its behavior on the fly by field reconfiguration while in operation, by switching in real time to search for a new pattern as soon as the new pattern is available to its engines through the downloading process. The system may be redundant and hot swappable, providing no single point of failure. This functionality may be obtained by duplicating the processing paths, sharing and redistribution of processing loads and load balancing, while also providing hot spares capable of taking over the work of faulty units in the system. Every single unit in the system may be provided with hot swap capability.

[0010] The system may provide various outputs consisting of both image and textual data. The system may be implemented such that it is capable of using specific image patterns as trigger events. This means that once specific patterns have been found in the images, those events can trigger the generation and output, in standardized formats, of, for example: video streams starting from a user-specified point in time relative to the trigger event; static images at user-specified times relative to the trigger event; metadata related to the triggering event such as input image source identification, timestamps and physical location and orientation of the image source, type, position, and size of the found patterns in the image, and other comprehensive standardized tags that facilitate database indexing of the event and its accompanying data.

[0011] Referring first to FIG. 1, in one embodiment, the invention is directed to a system 100 for processing images from at least one of a plurality of image providers 200 and transmitting the processing results to at least one of a plurality of destination processors 300. Referring next to FIG. 2, the image processing system 100 connects to image providers 200 through real-time data stream acquisition interfaces 101 and real-time static data acquisition interfaces 102, transforms the provided data, using at least one of a plurality of image formatter engines 110, into internally useable image data formats, specifies, through a feature object engine 121, at least one of a plurality of specific, uncorrelated and complex feature-based objects types to be found and identified in the image data.

[0012] In one embodiment, the system finds and identifies in the image data, using at least one of a plurality of object finder engines 120, multiple objects of the required type specified by the feature object engine 121, indexes and classifies, using at least one of a plurality of object classifier engines 122, the plurality of objects found by the object finder engine 120, stores image data, using at least one of a plurality of data storage engines 130, related to pre- and post- object finding, and generates alerts. The system may additionally provide a system management engine 150, to the entire system which performs operations such as status report generation, on the fly reconfigurability, code updates, upgrades, and so forth.

[0013] Referring next to FIG. 3, in one embodiment, any one of the plurality of the image providers 200 may provide any type of real-time or recorded, dynamic or static image data streams or files such as: any type of video capture devices 201 such as analog or digital video cameras; any type of video storage devices 202 such as

analog or digital video recorders; any type of static image acquisition devices 203 such as digital photo cameras and scanners; any type of static image storage devices 204 such as digital computer files; or any combination of the above such as other image acquisition systems 205.

[0014] Referring next to FIG. 4, in one embodiment, any one of the plurality of the destination processors 300 may be: any computer-based client application 301 such as software applications; any type of visual transducers 302 such as visual displays; any type of video transducers 303 such as video displays; any type of audio transducers 304 such as telephones; or any combination of the above 305, for example, a field dispatch system.

[0015] In one embodiment, the system may be used for face recognition. The output may be user definable and comprise one or more of the following: a standardized format video stream starting from a user specified point in time relatively to a trigger event; or a static image of standardized format such as a normalized ICAO International Civil Aviation Organization Type 2 and/or Type 3 image. This implies processing and improving the quality of an image in such a way to provide an optimum face image to a face recognition system that is internationally recognized and used for biometric face recognition and travel document other standardized image formats.

[0016] Metadata related to the triggering event may be: input image source identification file or camera; timestamps and physical location and orientation of the image source in the case of a camera it refers to its deployment location – GPS coordinates for outdoor camera - and its orientation – direction of where camera is looking – compass coordinates; position and size of faces found in the image together with primary landmarks, position of eyes and optional secondary landmarks positions of tip of the nose, corners of mouth, chin, etc., orientation of the face pitch, roll, and yaw angles, quality of the face image blur factor, brightness/contrast factors, etc.

[0017] In one embodiment, the system may be based on embedded processor architecture which provides a number of advantages: reduced footprint; reduced power consumption; reduced heat dissipation; automatic, fast system booting; easy and user friendly deployment practically deployable anywhere; modular deployment configuration tailored to user's needs; and scalable and tailored to user's needs (i.e. single to large enterprise camera source.)

[0018] One image source disclosed herein is hereinafter referred to as a “Camera Passport.” A Camera Passport is a small low power device that when attached to deployed cameras provides both preprogrammed and real time camera information data embedded in the video data signal or stream. For example, in the case of the analog cameras, the Camera Passport device is attached as a pass-through on the video data signal path embedding the camera information on the Vertical Blanking Interval, similar to close captioning systems. In fact, the embedded data is compatible with the close captioning format facilitating the visualization of the camera information on any closed-captioning-enabled display set.

[0019] The information data embedded in the video signal relates to, but is not limited to: unique identification number that is user or factory programmable for the camera; optional programmable name and location; optional GPS location data; optional compass orientation data; optional manual and user alerts. Any or all of these embedded data can be used for video routing and/or modification of system behavior.. Video routing commands can be as simple as making a video feed the highest priority and display it on the alerts monitors at the monitoring station.

[0020] The video routing commands may respond to alerts, which may be related to, for example, an officer in distress. An officer in vicinity of a camera can issue an encoded radio distress information signal which is received by the Camera Passport device and forwarded via the video cable to the monitoring station. In the case of a stolen car where cars are quipped with antitheft devices, the device may emit a distress signal which will cause the image from a particular camera, and any subsequent camera that receives that distress signal, to be displayed on the alerts monitors at the monitoring station.

[0021] Fire/smoke in the surveillance area (such as fire/smoke/explosion patterns, which when found will be treated as trigger events) may generate an alert. Motion, for example, in the case of surveillance areas where no motion is allowed at specific periods of time, may also generate an alert. The system may also process more complex objects, such as human faces or moving military vehicles, which have many different and dissimilar appearances. In the case of face finding most of the applications will center on the capture of facial images and only sometimes will require the capture of the “scene”, i.e. the full field of view of the camera. The system may also capture full video that might be useful, as in the case of a hold-ups,

altercations of some kind, etc. where additional motion-containing information may be of benefit.

[0022] The use of face-finding algorithms together with the use of image analysis software to determine the best pose, best quality facial image, facial feature points, eye locations may enable the saving of huge amounts of recording space when only a single, good quality, image is recorded per transaction/incident. Additionally, image processing software to correct images such as de-tilting, color saturation and balance, exposure and sizing as example may save computer processing time and eliminate the need for external processors.

[0023] Even though the above sections refer to facial detection and associated image quality analysis it should be realized that such are only examples of pattern recognition of one kind or another. Note that noise may also be a trigger event, for example, in the case of surveillance areas where no noise is allowed at specific periods of time, when noise is detected it will be treated as a trigger event.

[0024] The invention as taught and described herein may be used in combination with the systems and methods described in U.S. Patent Application No. 11/929,429 entitled "Computerized Biometric Passenger Identification System and Method" filed October 30, 2007, the entire disclosure of which is incorporated herein by reference.

[0025] While the invention has been particularly shown and described with reference to a preferred embodiment thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention.

CLAIMS

We Claim:

1. An autonomous system for processing video streams and/or single images, comprising:
 - a plurality of image providers configured to transmit images;
 - a plurality of destination processors configured to receive the transmitted images and to transform the transmitted images to internally useable image data;
 - a plurality of feature object engines configured to find and identify in the internally useable image data a plurality of objects; and
 - a plurality of object classifier engines configured to index and classify the plurality of objects found by the feature object engines.
2. The system for processing images of claim 1, further comprising a plurality of metadata inserters configured to embed static and/or dynamic information into the image transmission stream.
3. The system for processing images of claim 2, further comprising a plurality of decision engines that store and/or generate actionable activities based on the output of the destination processors and the metadata provided by the metadata inserters.
4. The system for processing images of claim 1, further comprising a plurality of intelligent video multiplexer/routers that use the metadata embedded in the video stream to route the video data to one more downstream destination processors.
5. The system for processing images of claim 1, wherein the plurality of image providers includes at least one image provider selected from the list consisting of: an analog or digital video camera, an analog or digital video recorder, a static image acquisition device, or a static image storage device.

6. The system for processing images of claim 1, wherein the destination processors include at least one image provider selected from the list consisting of: a computer-based client application, a visual transducer, a video transducer, or an audio transducer.
7. The system for processing images of claim 2, wherein the static and/or dynamic information includes static data selected from the list consisting of: name of the camera location, serial number of the camera, image source. or identification file.
8. The system for processing images of claim 2, wherein the static and/or dynamic information includes dynamic data related to an environmental monitoring system.
9. The system for processing images of claim 8, wherein the environmental monitoring system comprises at least one system selected from the set consisting of: fire alarms in the surveillance area, smoke alarms in the surveillance area, noise detector in surveillance area, motion detector in the surveillance area, current physical location of the camera using devices such as GPS system or beacons, current compass bearing where the camera is pointing, timestamp, auxiliary alarm system, lojack in a stolen car, access and proximity sensors, or control systems.
10. The system for processing images of claim 2, wherein the static and/or dynamic information includes dynamic data generated at the request of a person selected from the list consisting of: officer in distress, person at the front door, first responders, panic buttons, voice commands and recognition systems.
11. The system for processing images of claim 4, wherein the intelligent video multiplexer/routers use the metadata to dynamically prioritize the video streams to ensure that higher priority streams are assigned to operational downstream destination processors before assigning the lower priority streams.

12. The system for processing images of claim 11, wherein the system is configured such that the lower priority stream is dropped in the case of downstream destination process unavailability.
13. The system for processing images of claim 4, wherein the intelligent video multiplexer/routers use the static and/or dynamic information to assign a video stream to more than one downstream destination processor to allow simultaneous identification of two or more different kinds of objects in the video stream.
14. The system for processing images of claim 4, wherein the intelligent video multiplexer/routers use the static and/or dynamic information to assign more than one video streams to one downstream destination processor to maximize resource usage.
15. The system for processing images of claim 1, wherein the system is configured to utilize destination processors based on one or more of the following: application specific hardware architecture, reconfigurable hardware architecture, or embedded processors architecture.
16. The system for processing images of claim 15, wherein the system is configured to utilize a flexible and scalable modular deployment configuration tailored to an application's need where destination processors are added as required to deploy additional functionality.
17. The system for processing images of claim 15, wherein destination processors are configured to alter their behavior, and be targeted to find a different class of object.
18. The system for processing images of claim 1, further comprising one or more decision engines which generate alarms in response to the classification of the objects found by one of the feature object engines.

19. The system for processing images of claim 18, wherein the alarm is accompanied by metadata related to the triggering event.
20. The system for processing images of claim 19, wherein the static and/or dynamic information includes data related to the target object, selected from the list consisting of: position of the object, orientation of the object, position and orientation of sub-feature of the object, quality of the image, blur factor, or brightness/contrast factors.
21. The system for processing images of claim 20, wherein, in the case of a face object, the static and/or dynamic information includes data selected from the list consisting of: the position of eyes, the position of tip of the nose or corners of mouth, orientation of the face pitch, roll, and yaw angles.
22. The system for processing images of claim 2, wherein the static and/or dynamic information that was inserted before the image processing, and the metadata generated after image processing, is used to generate an actionable activity selected from the set consisting of: decision to send the data and image for long term storage, send image data to video displays being monitored by operators, perform object recognition, send pertinent information to field dispatch systems, or raise visual/audio alarms.
23. The system for processing images of claim 1, wherein the system is configured for face recognition and outputs a standardized format video stream.
24. The system for processing images of claim 23, wherein the standardized format video stream frames are processed and normalized to ICAO International Civil Aviation Organization static image format, type 2 and/or type 3, ISO International Organization for Standardization format or INCITS International Committee for Information Technology Standards or any other required image standards.

25. The system for processing images of claim 1, wherein at least one of the image providers is an image enhancement device comprising: an image provider configured to transmit an image to an image enhancement device, or an image enhancement device configured to add additional information to the image and to retransmit the image.
26. The system for processing images of claim 25, wherein at least one image provider is a video camera and the additional data is embedded in the camera information on the Vertical Blanking Interval.
27. A method for processing images, comprising the steps of: transmitting images to a plurality of destination processors; receiving transmitted images at the plurality of destination processors; transforming the received images to internally useable image data; and using the plurality of destination processors to identify a plurality of objects in the internally useable image data.
28. The method of claim 27, further comprising the step of classifying the plurality of objects identified in the internally usable image data.
29. The method of claim 27, further comprising the step of generating an alarm in response to classification of the plurality of objects identified in the internally usable image data.
30. The method of claim 27, further comprising the step of enhancing at least one transmitted image with additional information.
31. The method of claim 30, wherein the image transmitted is from a video camera and the additional data is embedded in the camera information on the Vertical Blanking Interval.
32. The method of claim 31, wherein the additional information includes at least one item selected from the list consisting of: a unique identification number, a name and location, optional location data, or compass orientation data.

33. The system for processing images of claim 1, further comprising a system management engine which manages the system and performs system management functions.
34. The system for processing images of claim 33, wherein the system management engine performs one or more operations selected from the list consisting of: status report generation, reconfiguration, code updates and code upgrades.