

DESCRIPTION

IMAGE PROCESSING DEVICE, IMAGE PROCESSING METHOD, AND
IMAGE PROCESSING PROGRAM

Technical Field

[0001] The present invention relates to a technique of processing a frame image of a target area shot with a camera, and detecting a plurality of types of objects having different sizes.

Background Art

[0002] Conventionally an image processing device is available, which processes a frame image shot with a camera to detect objects such as a human or a vehicle in the shot frame image and to track such a detected object (tracking). Patent Literature 1, for example, describes a device to measure a traffic flow configured to process a frame image shot at an intersection to count up the number of vehicles traveling in each branching direction (going straight, turning right and turning left).

[0003] Using information such as the number of vehicles counted up at an intersection in each branching direction (going straight, turning right and turning left) and the presence or not of a pedestrian passing through the crosswalk, a traffic light provided at this intersection also is controlled.

Citation List

[0004] Patent Literature 1: Patent Application Publication

No. 2005-190142

Summary of Invention

Technical Problem

[0005] However in order to process a frame image shot with a camera to detect or track an object in the image, the accuracy cannot be ensured unless the image has appropriate resolution for the object to be detected. Meanwhile the appropriate range (the range yielding the sufficient accuracy) of resolution of an image varies with the size of an object to be detected or tracked, the traveling speed of the object and the like.

[0006] In this way, in a system for detecting or tracking of a plurality of types of objects having different sizes and different traveling speeds (e.g., a system to detect and track a plurality of types of objects such as automobiles (four-wheel vehicles), motorbikes (two-wheel vehicles), and pedestrians at an intersection), if the resolution of an image is set at an appropriate range for some type of objects, such resolution of the image may be beyond the appropriate range for other types of objects. Then the system cannot have enough accuracy in detecting and tracking for such types of objects beyond the appropriate range of the image resolution.

[0007] In this way, conventional apparatuses have difficulty in accurately detecting and tracking a plurality of types of objects having different sizes and traveling speeds, for example.

[0008] Note here that the appropriate resolution range of an image for each type of objects has a certain width (i.e.,

the appropriate resolution of an image does not have a certain value).

[0009] Then it is an object the present invention to provide a technique of detecting and tracking a plurality of types of objects having different sizes and traveling speeds, for example, accurately.

Solution to Problem

[0010] To fulfill such an object, an image processing device of the present invention is configured as follows.

[0011] An image input unit receives a frame image of a target area that is shot with a camera as an input. A plurality of processing regions each having a determined type of objects as a target of detection (vehicles, pedestrians and the like) is set on the frame input to the image input unit. Each processing region may or may not overlap with another processing region.

[0012] A resizing processing unit creates a resized image for each of the processing regions, the resized image having resolution that is converted in accordance with the type of the objects as the target of detection.

[0013] Then an object detection unit processes the resized image created by the resizing processing unit to detect an object of the type as the target of detection.

[0014] In this way, instead of directly processing a frame image shot with a camera to detect an object therefrom, a resized image is created for each of the processing regions each having a determined type of objects as a target of detection, the resized image having resolution that is converted in accordance with

the type of the objects as the target of detection, and an object is detected by processing this resized image. Then even when detection and tracking are performed for a plurality of types of objects having different sizes and traveling speeds, for example, the detection and tracking can be performed accurately for each of the types of objects.

[0015] Since there is no need to provide a camera for each type of objects, an increase in the scale or the cost of the system can be suppressed.

[0016] The resizing processing unit may create the resized image having a predetermined number of pixels for each of the processing regions. In this case, the size of a processing region on the frame image may be set so that the resized image to be created has resolution suitable for the type of objects to be detected that is determined for the processing region.

[0017] The resizing processing unit may create an image having resolution that is predetermined for each processing region. In this case, resolution can be set irrespective of the size of the processing region set on the frame image. In this case, however, the number of pixels of resized images to be created for the processing regions will not become uniform.

[0018] The image processing device further may include an object map creation unit that creates an object map for each object detected by the object detection unit, the object map indicating a change in position of the object on the target area over time. This can obtain a change in position of each detected object over time (traveling locus).

[0019] The image input unit may receive frame images shot

with a plurality of cameras having different angles with respect to the target area as an input, and the resizing processing unit and the object detection unit may be provided for each camera. This can prevent an object from being undetected because an image of such an object is not shot behind another object located in front in a frame image shot with some camera.

[0020] Preferably the object map creation unit also is provided for each camera. In this case, the image processing device may further include an integration unit that integrates object maps created by the object map creation units provided for the cameras in units of objects. This allows objects detected from frame images shot with a plurality of cameras to be processed as one object.

[0021] When the target area is an intersection, a type of an object to be detected may be a vehicle (automobiles or motorbikes) or a pedestrian. Then, it may be configured so that an output signal is output, indicating passage of a vehicle or a human through the intersection as the target area on a basis of a detection result of the object detection unit. This allows a traffic light to be controlled in accordance with the amount of traffic.

Advantageous Effects of Invention

[0022] The present invention enables accurate detection and tracking of a plurality of types of objects having different sizes and traveling speeds, for example.

Brief Description of Drawings

[0023]

FIG. 1 is a block diagram illustrating the configuration of a major part of an image processing device.

FIG. 2 illustrates an example of video cameras provided at an intersection.

FIG. 3 is a functional block diagram illustrating the functional configuration of an image processing processor.

FIG. 4 illustrates a frame image of an intersection that is shot with a video camera.

FIG. 5 illustrates a resized image.

FIG. 6 illustrates an object map.

FIG. 7 is a flowchart illustrating the operation of the image processing device.

FIG. 8 is a flowchart illustrating image processing.

Description of Embodiments

[0024] The following describes an image processing device that is one embodiment of the present invention.

[0025] FIG. 1 is a block diagram illustrating the configuration of a major part of the image processing device in this embodiment. The image processing device 1 in this embodiment includes a controller 2, an image input unit 3, an image processing processor 4, and an output unit 5. The image processing device 1 in this embodiment detects and tracks a vehicle and a pedestrian at an intersection. For simplicity of description, this embodiment detects two types of objects, i.e., a vehicle (automobile (four-wheel vehicle) and motorbike (two-wheel vehicle)) and a pedestrian.

[0026] The controller 2 controls the operation of various parts of the main body of the image processing device 1. This unit also performs integration processing (the configuration corresponding to an integration unit specified in this invention) and output-signal generation processing described later.

[0027] The image input unit 3 receives, as an input, a frame image of an intersection that is shot with two video cameras 10 and 11 connected thereto. The video cameras 10 and 11 have a frame rate of a few tens frames/sec. (e.g., 30 frames/sec.).

[0028] The image processing processor 4 includes a first image processing unit 41 that processes a frame image input from the video camera 10 on one side to detect and track a vehicle and a pedestrian as an object of detection, and a second image processing unit 42 that processes a frame image input from the video camera 11 on the other side to detect and track a vehicle and a pedestrian as an object of detection. The details of the image processing processor 4 are described later. This image processing processor 4 corresponds to a computer that executes an image processing method according to the present invention. The image processing processor 4 also corresponds to a computer to allow an image processing program according to the present invention to be executed.

[0029] The output unit 5 outputs a signal based on the result of the processing relating to the detection and tracking of a vehicle and a pedestrian as an object of the detection (output signal generated at the controller 2). For instance, the output unit 5 outputs a passage signal of a vehicle and a pedestrian (passage pulse) to a traffic-light controller (not illustrated)

that controls a traffic light provided at the intersection. Alternatively, the output unit 5 may be configured to count up the number of vehicles in each branching direction (going straight, turning right and turning left) at the intersection for outputting to a device monitoring the amount of traffic at the intersection. That is, the signal (the output signal that the controller 2 generates) that the output unit 5 outputs may be determined depending on the device receiving the signal and the intended use.

[0030] As illustrated in FIG. 2, the two video cameras 10 and 11 are provided so as to shoot images while facing each other at the intersection. Each video camera 10, 11 used may include a wide-angle lens as its lens for shooting so as to cover the intersection entirely as the shooting area. Then the video cameras 10 and 11 include an image pickup device of mega pixels (e.g., about $1,300 \times 1,000$ in length and width). As is evident from FIG. 2, the video cameras 10 and 11 have different angles with respect to the intersection.

[0031] Note here that the two video cameras 10 are disposed so that, when the video camera 10 (or 11) on one side is shooting an image of a large-sized vehicle (e.g., a bus or a lorry), and then an image of a vehicle (normal vehicle), a pedestrian and the like behind such a large-sized vehicle may not be shot, the video camera 11 (or 10) on the other side can shoot an image of such a normal vehicle, pedestrian and the like. In other words, the number of cameras provided at the intersection may be one as long as images of all of the objects located in the shooting area can be shot without hiding them behind other objects. Three

or more of cameras may be provided at the intersection. This embodiment illustrates the case of the two video cameras 10 and 11 provided as described above.

[0032] The image input unit 3 includes a first image input unit (not illustrated) that receives, as an input, a frame image of the intersection that is shot with the video camera 10 on one side and a second image input unit (not illustrated) that receives, as an input, a frame image of the intersection that is shot with the video camera 11 on the other side. The image input unit 3 includes an image memory (not illustrated) that temporarily stores a frame image input.

[0033] FIG. 3 is a functional block diagram illustrating the functional configuration of the image processing processor. The image processing processor 4 includes a first image processing unit 41 and a second image processing unit 42. The first image processing unit 41 and the second image processing unit 42 have the same functional configuration. The first image processing unit 41 processes a frame image of the intersection that is shot with the video camera 10 on one side, and the second image processing unit 42 processes a frame image of the intersection that is shot with the video camera 11 on the other side.

[0034] When one video camera 10 (or 11) is provided at the intersection, the second image processing unit 42 (or the first image processing unit 41) may be omitted. When three or more video cameras are provided at the intersection, then the image processing processor 4 will have image processing units in number corresponding to the number of the video cameras.

[0035] The first image processing unit 41 includes a

resizing processing functional unit 411, a vehicle detection processing functional unit 412, a vehicle tracking processing functional unit 413, a vehicle determination processing functional unit 414, a pedestrian detection processing functional unit 415, a pedestrian tracking processing functional unit 416, and a pedestrian determination processing functional unit 417.

[0036] The resizing processing functional unit 411 performs resizing processing, in which a resized image is created, having resolution that is converted for each processing region set at a frame image input to the first image input unit of the image input unit 3. FIG. 4 illustrates a frame image of the intersection that is shot with the video camera on one side. As illustrated in FIG. 4, a plurality of processing regions is set for the frame image shot with the video camera 10. In the example illustrated in FIG. 4, two processing regions 100 and 101 are regions in which the type of objects to be detected is determined as vehicles (hereinafter called vehicle detection processing regions 100 and 101). Four processing regions 110 to 113 are regions in which the type of objects to be detected is determined as pedestrians (hereinafter called pedestrian detection processing regions 110 to 113).

[0037] As illustrated in FIG. 4, the vehicle detection processing region 100 is the region that covers the inside of the intersection, and the vehicle detection processing region 101 is the region that covers an area close to the stop line on the opposite side of the side where the video camera 10 is provided to shoot the frame image. In the example of FIG. 4, the vehicle detection processing region 100 and the vehicle

detection processing region 101 overlap partially. The four pedestrian detection processing regions 110 to 113 are regions that cover crosswalks provided at the intersection. These four pedestrian detection processing regions 110 to 113 do not overlap with the other pedestrian detection processing regions 110 to 113. The pedestrian detection processing regions 110 to 113, however, overlap with the vehicle detection processing regions 100 and 101 partially.

[0038] The resizing processing functional unit 411 performs resizing processing for each of the vehicle detection processing regions 100 and 101 and the pedestrian detection processing regions 110 to 113 so that images there have a predetermined number of pixels (in this example, a VGA image of 640×480). An image created through this resizing processing is referred to here as a resized image. The resolution for a resized image in the vehicle detection processing regions 100 and 101 is within an appropriate range for vehicle detection. Then the resolution for a resized image in the pedestrian detection processing regions 110 to 113 is within an appropriate range for pedestrian detection. In other words, the vehicle detection processing regions 100 and 101 are configured so that the resolution of a resized image created is set within an appropriate range for vehicle detection, and the pedestrian detection processing regions 110 to 113 are configured so that the resolution of a resized image created is set within an appropriate range for pedestrian detection. This resizing processing may be performed using a known method such as nearest neighbor interpolation, bilinear interpolation or bicubic interpolation. FIG 5(A) illustrates a resized image

for the vehicle detection processing region 100, and FIG. 5(B) illustrates a resized image for the pedestrian detection processing region 110.

[0039] Note here that a resized image for the vehicle detection processing region 100 and the vehicle detection processing region 101 may have resolution within the range suitable for vehicle detection, which does not mean that these two regions have the same resolution for resized images. Similarly, a resized image for the four regions of the pedestrian detection processing regions 110 to 113 may have resolution within the range suitable for pedestrian detection, which does not mean that these four regions have the same resolution for resized images.

[0040] The vehicle detection processing functional unit 412 processes a resized image of each of the vehicle detection processing regions 100 and 101 for detection of vehicles to detect a vehicle in the image. The vehicle detection processing functional unit 412 detects a vehicle in the image using known pattern matching.

[0041] For each vehicle detected by the vehicle detection processing functional unit 412, the vehicle tracking processing functional unit 413 creates an object map for the vehicle, in which the ID of the vehicle, the detected time (time when the frame image was shot), the location (coordinates indicating the location on the intersection), and the status are associated. Since a location on the intersection is used as the location of the vehicle detected on this object map, it enables identification of the vehicle that is located at an overlapping

part of the two vehicle detection processing regions 100 and 101. This means that, when a certain vehicle is detected in the resized image for the vehicle detection processing region 100 as well as in the resized image for the vehicle detection processing region 101, such a vehicle can be detected as one vehicle.

[0042] When the vehicle tracking processing functional unit 413 creates an object map for a vehicle, the vehicle determination processing functional unit 414 determines the status of such a vehicle based on the traveling locus of the vehicle that is obtained from the object map already created for the vehicle. The status of the vehicle determined at this time includes before the entry to the intersection, in the intersection, and leaving the intersection.

[0043] The pedestrian detection processing functional unit 415 processes a resized image of each of the pedestrian detection processing regions 110 to 113 for detection of pedestrians to detect a pedestrian in the image. The pedestrian detection processing functional unit 415 detects a pedestrian in the image using known pattern matching.

[0044] For each pedestrian detected by the pedestrian detection processing functional unit 415, the pedestrian tracking processing functional unit 416 creates an object map for the pedestrian, in which the ID of the pedestrian, the detected time (time when the frame image was shot), and the location (coordinates indicating the location on the intersection) are associated. Since a location on the intersection is used as the location of the pedestrian detected on this object map as well, it enables identification of the pedestrian who is located at an overlapping

part of the plurality of pedestrian detection processing regions 110 to 113.

[0045] In the example of FIG. 4, the pedestrian detection processing regions 110 to 113 do not overlap with each other.

[0046] When the pedestrian tracking processing functional unit 416 creates an object map for a pedestrian, the pedestrian determination processing functional unit 417 determines the status of such a pedestrian based on the traveling locus of the pedestrian that is obtained from the object map already created for the pedestrian. The status of the pedestrian determined at this time includes a pedestrian being outside of the crosswalk (pedestrian who is not walking across the crosswalk but is waiting at the traffic light, for example), a pedestrian passing through the center (pedestrian who is walking across the crosswalk and has reached the center of the crosswalk), and a pedestrian not reaching the center (pedestrian who is walking across the crosswalk but has not reached the center).

[0047] As illustrated in FIG. 3, similarly to the first image processing unit 41 as stated above, the second image processing unit 42 also includes a resizing processing functional unit 421, a vehicle detection processing functional unit 422, a vehicle tracking processing functional unit 423, a vehicle determination processing functional unit 424, a pedestrian detection processing functional unit 425, a pedestrian tracking processing functional unit 426, and a pedestrian determination processing functional unit 427. These resizing processing functional unit 421, vehicle detection processing functional unit 422, vehicle tracking processing functional unit 423,

vehicle determination processing functional unit 424, pedestrian detection processing functional unit 425, pedestrian tracking processing functional unit 426, and pedestrian determination processing functional unit 427 are the same functional blocks as the resizing processing functional unit 411, the vehicle detection processing functional unit 412, the vehicle tracking processing functional unit 413, the vehicle determination processing functional unit 414, the pedestrian detection processing functional unit 415, the pedestrian tracking processing functional unit 416, and the pedestrian determination processing functional unit 417 of the first image processing unit 41 as stated above, respectively, and their descriptions are omitted here.

[0048] Setting is performed separately for the vehicle detection processing regions 100 and 101 as well as the pedestrian detection processing regions 110 to 113 in a frame image shot with the video camera 10 on one side and for those processing regions in a frame image shot with the video camera 11 on the other side. There is no problem when the number of the processing regions set is different between for a frame image shot with the video camera 10 on one side and for a frame image shot with the video camera 11 on the other side.

[0049] Referring now to FIG. 6, the following briefly describes an object map created by the image processing processor 4.

[0050] FIG. 6(A) illustrates an object map for a vehicle. FIG. 6(B) illustrates an object map for a pedestrian. Such an object map is created for each vehicle or each pedestrian detected.

FIG. 6(A) is an object map of a vehicle having the ID "p0001" assigned. As illustrated in FIG. 6(A), this object map contains records registered to be accumulated in chronological order, each record containing the shooting time (shooting time of a frame image that is used to create the resized image), the location of the vehicle on the intersection, and the status of the vehicle that is determined by the vehicle determination processing functional unit 414 that are associated for each resized image of the vehicle detected. Then a change in location on the intersection (traveling locus) over time can be obtained for the corresponding vehicle based on this object map. This shows from which direction the vehicle entered the intersection and in which direction the vehicle passed through the intersection (turn right, turn left or go straight), for example. That is, this shows the ratio of branching of vehicles at the intersection, for example.

[0051] Similarly to the object map for a vehicle as stated above, the same goes for the object map for a pedestrian illustrated in FIG. 6(B). FIG. 6(B) illustrates an object map of a pedestrian having the ID "q0001" assigned. As stated above, this object map also shows a change in location on the intersection over time. That is, this shows in which direction the corresponding pedestrian is walking across the crosswalk.

[0052] The following describes the operation of the image processing device 1 according to this embodiment. FIG. 7 is a flowchart illustrating the operation of this image processing device.

[0053] As described above, the image input unit 3 (the first

image input unit) receives a frame image of the intersection shot with the video camera 10 on one side as an input and the image input unit 3 (the second image input unit) receives a frame image of the intersection shot with the video camera 11 on the other side as an input. The image input unit 3 includes an image memory (not illustrated) that stores a frame image input.

[0054] The image processing device 1 performs image processing of a frame image stored in the image memory (s1). At s1, the first image processing unit 41 processes the frame image shot with the video camera 10 on one side, and the second image processing unit 42 processes the frame image shot with the video camera 11 on the other side. Processing is performed concurrently at the first image processing unit 41 and the second image processing unit 42.

[0055] The image processing performed is substantially the same between the first image processing unit 41 and the second image processing unit 42, which are different only in images to be processed. The following describes the image processing at s1 by way of the example at the first image processing unit 41. FIG. 8 is a flowchart illustrating the image processing at s1.

[0056] The first image processing unit 41 selects a frame image that is stored in the image memory but is not processed, and has an earliest time (the oldest frame image) as a processing target for fetching (s11). The first image processing unit 41 selects a frame image to be processed from frame images shot with the video camera 10 on one side.

[0057] The resizing processing functional unit 411 of the

first image processing unit 41 performs resizing processing for the frame image to be processed that is fetched at s11 (s12). The resizing processing functional unit 411 performs the resizing processing to convert an image to have a predetermined number of pixels (in this example, a VGA image of 640×480) for each of the vehicle detection processing regions 100 and 101 and the pedestrian detection processing regions 110 to 113 illustrated in FIG. 4.

[0058] Then the vehicle detection processing functional unit 412 of the first image processing unit 41 performs processing to detect a vehicle in the shot image for each of the resized images of the vehicle detection processing regions 100 and 101 (s13). At s13, a vehicle in the shot image is detected using a known technique such as pattern matching. Then at s13, coordinates data is acquired for the detected vehicle, indicating the location on the intersection.

[0059] The vehicle tracking processing functional unit 413 of the first image processing unit 41 performs identification processing to extract the vehicle detected at s13 from both of the resized image for the vehicle detection processing region 100 and the resized image for the vehicle detection processing region 101 (s14). The processing at s14 is to associate the images of the vehicle located at the overlapping region of the vehicle detection processing region 100 and the vehicle detection processing region 101 to have a one-to-one correspondence between them. That is, this processing is to detect the vehicles located at the overlapping region of the vehicle detection processing region 100 and the vehicle detection processing region 101 as

one vehicle. This identification processing is performed based on the location of the vehicles on the intersection that is acquired at s13.

[0060] The vehicle tracking processing functional unit 413 creates an object map for each vehicle detected this time (s15). At s15, a record is created for each vehicle detected this time so as to associate the shooting time of the frame image to be processed that is selected at s11 and the location on the intersection acquired at s13. The record created here does not include the status of the vehicle as illustrated in FIG. 6(A). The vehicle tracking processing functional unit 413 also determines whether an ID has been assigned to the vehicle or not for each record created this time. When the ID has not been assigned to the vehicle at this time, this means that the vehicle entered the vehicle detection processing regions 100 and 101 from the outside of the vehicle detection processing regions 100 and 101 from the shooting time of the frame image to be processed last time to the shooting time of the frame image to be processed this time, and so the object map as illustrated in FIG. 6(A) is not created therefor at this time.

[0061] The vehicle tracking processing functional unit 413 assigns an ID to the vehicle that does not have an ID assigned at this time, and creates the object map as illustrated in FIG. 6(A) (object map, in which the record created this time only is registered). On the other hand, when the vehicle already has the ID assigned thereto, an object map is created so as to add the record created this time to the object map of the corresponding vehicle (by adding the record created this time to the object

map of the corresponding vehicle). Determination whether the ID has been assigned or not can be made based on a temporal change between the vehicle position on the intersection acquired at s13 and the vehicle position on the intersection acquired from the object map of the vehicles having an ID assigned thereto at this time.

[0062] Then the vehicle determination processing functional unit 414 of the first image processing unit 41 determines the status of each vehicle detected this time (before the entry to the intersection, in the intersection, and leaving the intersection), and registers such a status at the record that is additionally registered at the object map of the corresponding vehicle created at s15 (s16).

[0063] The pedestrian detection processing functional unit 415 of the first image processing unit 41 performs processing to detect a pedestrian in the image shot for each of the resized images of the pedestrian detection processing regions 110 to 113 (s17). At s17, a pedestrian in the shot image is detected using a known technique such as pattern matching. Then at s17, coordinates data is acquired for the detected pedestrian, indicating the location on the intersection.

[0064] The pedestrian tracking processing functional unit 416 of the first image processing unit 41 performs identification processing to extract the pedestrian detected at s17 from the resized images for the plurality of pedestrian detection processing regions 110 to 113 (s18). The processing at s18 is to associate the images of the pedestrian on the intersection located at the overlapping region of the plurality of pedestrian

detection processing regions 110 to 113 to have a one-to-one correspondence among them. That is, this processing is to detect the pedestrians located at the overlapping region of the plurality of pedestrian detection processing regions 110 to 113 as one pedestrian. This identification processing is performed based on the location of the pedestrians on the intersection that is acquired at s17.

[0065] Since there is no overlapping among the pedestrian detection processing regions 110 to 113 in the example of FIG. 4, the processing at s18 may be omitted.

[0066] The pedestrian tracking processing functional unit 416 creates an object map for each pedestrian detected this time (s19). At s19, a record is created for each pedestrian detected this time so as to associate the shooting time of the frame image to be processed that is selected at s11 and the location on the intersection acquired at s17. The record created here does not include the status of the pedestrian as illustrated in FIG. 6(B). The pedestrian tracking processing functional unit 416 also determines whether an ID has been assigned to the pedestrian or not for each record created this time. When the ID has not been assigned to the pedestrian at this time, this means that the pedestrian entered the pedestrian detection processing regions 110 to 113 from the outside of the pedestrian detection processing regions 110 to 113 from the shooting time of the frame image to be processed last time to the shooting time of the frame image to be processed this time, and so the object map as illustrated in FIG. 6(B) is not created therefor at this time.

[0067] The pedestrian tracking processing functional unit

416 assigns the ID to the pedestrian that does not have an ID assigned at this time, and creates the object map illustrated in FIG. 6(B) (object map, in which the record created this time only is registered). On the other hand, when the pedestrian already has the ID assigned thereto, an object map is created so as to add the record created this time to the object map of the corresponding pedestrian. Determination whether the ID has been assigned or not can be made based on a temporal change between the pedestrian position on the intersection acquired at s17 and the pedestrian position on the intersection acquired from the object map of the pedestrians having an ID assigned thereto at this time.

[0068] Then the pedestrian determination processing functional unit 417 of the first image processing unit 41 determines the status of each pedestrian detected (pedestrian being outside of the crosswalk, pedestrian passing through the center, and pedestrian not reaching the center), and registers such a status at the record that is additionally registered at the object map of the corresponding pedestrian created at s19 (s20).

[0069] The processing from s13 to s16 and the processing from s17 to s20 are substantially the same, which are different only in the types of target of the processing between vehicles and pedestrians. The order of the processing from s13 to s16 and the processing from s17 to s20 may be reversed from the above (following the processing from s17 to s20, the processing from s13 to s16 may be performed). The first image processing unit 41 may have a configuration so as to perform the processing from

s13 to s16 and the processing from s17 to s20 concurrently.

[0070] The resized images for the processing regions 100, 101 and 110 to 113 created at s12 are images having the same number of pixels (in this example, VGA images), they may be processed with a substantially same processing program for the detection of objects (vehicles and pedestrians) for the resized images. Specifically they can be dealt with by changing a parameter for pattern matching or the like, resulting from a difference in types of the objects to be detected (vehicles or pedestrians).

[0071] Referring back to FIG. 7, when finishing the image processing at s1 as stated above, the controller 2 of the image processing device 1 performs integration processing (s2). The integration processing at s2 is the processing to integrate the object map created by the first image processing unit 41 and the object map created by the second image processing unit 42 during the image processing at s1. Specifically, the objects (vehicles and pedestrians) that are detected by both of the first image processing unit 41 and the second image processing unit 42 are detected as one object, and the objects (vehicles and pedestrians) that are detected by any one of the first image processing unit 41 and the second image processing unit only are detected as one object. That is, the object map created by the first image processing unit 41 and the object map created by the second image processing unit 42 are integrated, so that one object map is acquired for each of the objects detected. This enables an object map to be acquired for a vehicle (normal vehicle), a pedestrian and the like as well whose images are

not shot because they are located behind a large-sized vehicle at the time of shooting of the large-sized vehicle by the video camera 10 (or 11) on one side. This can prevent objects (normal vehicles or pedestrians) from being undetected.

[0072] The controller 2 of the image processing device 1 creates an output signal based on the result of the integration processing at s2, and the output unit 5 outputs this output signal (s3), and then the processing returns to s1. This output signal, for example, is output to a traffic-light controller that controls a traffic light provided at the intersection, which is a pulse signal indicating the passage of vehicles and pedestrians. Alternatively, the output signal may be output to a device monitoring the amount of traffic at the intersection, which may be a digital signal indicating the number of vehicles counted up for each branching direction (going straight, turning right and turning left) at the intersection.

[0073] The above description exemplifies the image processing device according to the present invention configured to detect vehicles and pedestrians at the intersection. Instead, vehicles may be differentiated between automobiles (four-wheel vehicles) and motorbikes (two-wheel vehicles). In this case, instead of the vehicle detection processing regions 100 and 101 in FIG. 4, automobile detection processing regions and motorbike detection processing regions may be set at the frame images shot with the video cameras 10 and 11. The type of objects to be detected may include bicycles additionally. In this case, bicycle vehicle detection processing regions may be set at the frame images shot with the video cameras 10 and 11. Then the image

processing processor 4 may include a detection processing functional unit, a tracking processing functional unit, and a determination processing functional unit for each of the types of the objects to be detected and tracked.

[0074] The present invention is not limited to the above examples, and is applicable to devices that detect a plurality of type of objects having different sizes. The shooting area of the video cameras 10 and 11 also is not limited to an intersection, which may be other places such as a platform.

[0075] In the above examples, the resized images created for each of the vehicle detection processing regions 100 and 101 and the pedestrian detection processing regions 110 to 113 are VGA images, and the resized images created for these processing regions may have different number of pixels. Note here that the resolution of a resized image created for each processing region is within an appropriate range for the objects to be detected set for the processing region.

Reference Signs List

[0076]	1	Image processing device
	2	Controller
	3	Image input unit
	4	Image processing processor
	5	Output unit
	10, 11	Video camera
	100, 101	Vehicle detection processing region
	110 to 113	Pedestrian detection processing region
	41	First image processing unit

42	Second image processing unit
411, 421	Resizing processing functional unit
412, 422	Vehicle detection processing functional unit
413, 423	Vehicle tracking processing functional unit
414, 424	Vehicle determination processing functional unit
415, 425	Pedestrian detection processing functional unit
416, 426	Pedestrian tracking processing functional unit
417, 427	Pedestrian determination processing functional unit

CLAIMS

[1] An image processing device, comprising:

an image input unit that receives a frame image of a target area shot with a camera as an input;

a resizing processing unit that creates a resized image created for each of a plurality of processing regions on the frame image input to the image input unit, the processing regions each having a determined type of objects as a target of detection, the resized image having resolution that is converted in accordance with the type of the objects as the target of detection; and

an object detection unit that processes the resized image created by the resizing processing unit to detect an object of the type as the target of detection.

[2] The image processing device according to claim 1, wherein the resizing processing unit creates the resized image having a predetermined number of pixels for each of the processing regions.

[3] The image processing device according to claim 1 or 2, further comprising an object map creation unit that creates an object map for each object detected by the object detection unit, the object map indicating a change in position of the object on the target area over time.

[4] The image processing device according to any one of claims

1 to 3, wherein

the image input unit receives frame images shot with a plurality of cameras having different angles with respect to the target area as an input, and

the resizing processing unit and the object detection unit are provided for each camera.

[5] The image processing device according to any one of claims 1 to 3, wherein

the image input unit receives frame images shot with a plurality of cameras having different angles with respect to the target area as an input,

the resizing processing unit, the object detection unit, and the object map creation unit are provided for each camera, and

the image processing device further comprises an integration unit that integrates object maps created by the object map creation units provided corresponding to the respective cameras in units of objects.

[6] The image processing device according to any one of claims 1 to 5, wherein the target area is an intersection.

[7] The image processing device according to claim 6, wherein in each processing region set on the frame image input to the image input unit, a type of an object to be detected is a vehicle or a human.

[8] The image processing device according to claim 6 or 7, further comprising an output unit that outputs an output signal indicating passage of a vehicle or a human through an intersection as the target area on a basis of a detection result of the object detection unit.

[9] An image processing method executed by a computer that process a frame image of a target area shot with a camera, the frame image being input to an image input unit, comprising:

a resizing processing step of creating a resized image created for each of a plurality of processing regions on the frame image input to the image input unit, the processing regions each having a determined type of objects as a target of detection, the resized image having resolution that is converted in accordance with the type of the objects as the target of detection; and

an object detection step of processing the resized image created at the resizing processing step to detect an object of the type as the target of detection.

[10] An image processing program executed by a computer that process a frame image of a target area shot with a camera, the frame image being input to an image input unit, comprising:

a resizing processing step of creating a resized image created for each of a plurality of processing regions on the frame image input to the image input unit, the processing regions each having a determined type of objects as a target of detection, the resized image having resolution that is converted in

accordance with the type of the objects as the target of detection;
and

an object detection step of processing the resized image created at the resizing processing step to detect an object of the type as the target of detection.