

Rain Removal Based on Superpixel Spatial-Temporal Matching and Filtering

TECHNICAL FIELD

[0001] The present invention relates to a method of post-processing images and video, and more specifically, to a method of removing rain related distortion from images and video of outdoor environments.

BACKGROUND

[0002] Outdoor cameras are commonly used in automated systems and security. For example, self-driving automobiles typically use cameras to detect and analyze objects on roads. Surveillance cameras are often used to identify individuals in photographs or video. As such, outdoor cameras require accurate feature detection to be effective. Particles in the air, including rain, can degrade or distort images and render cameras ineffective. Rain fall can introduce false frame motion, degrade contrast and visibility and obscure important features that are vital for computer vision algorithms. Accordingly, it is a challenge for cameras to take adequate images/video during inclement weather. Some have proposed methods to remove rain in the post-processing of images and video.

[0003] The visual effects of rain in images and video can be complex. Rain produces an array of non-homogenous refractions and reflections of the light and environment. The small but frequent intensity fluctuations introduced by rain can be caused by several factors such as illumination changes, camera motions and object motions. To remove the rain, it is therefore important to either differentiate these motion causes or to compensate for them.

[0004] Rain removal can be a vital processing procedure for outdoor vision based systems. It is particularly useful in applications that rely on computer vision. Without it, images or video will remain distorted and difficult to view and/or analyze. Such methods typically employ algorithms to identify distortions from rain drops and revert the distortions based on nearby regions on the film. In other words, rain pixels are first detected and then pixels are replaced with an estimated background value. Rain pixels can be detected by analyzing spatial patterns, temporal patterns or a combination of both.

[0005] In the spatial analysis approach, a frame is analyzed for content that might correspond to rain streaks based on geometric properties. This method is described by Kim et al. (J.-H. Kim, C. Lee, J.-Y. Sim, and C.-S. Kim, "Single-image deraining using an adaptive nonlocal means filter," 2013 20th IEEE International Conference on Image Processing (ICIP), Sep. 2013, pp. 914-917). The statistics of the local window around each pixel, such as intensity orientation and relative intensity, is analyzed to classify the pixel. A pixel that presents distortion from rain can be revamped to match other nearby pixels. Although this approach can "de-rain" images without referencing their temporal neighbors, it is too slow to be considered for real time applications.

[0006] The temporal analysis approach works similarly but compares a pixel's value from a particular location over time. Rain streaks that appear over a pixel induce a spike in intensity. Thus, a pixel exhibiting this pattern could be a rain pixel in that particular frame. While this approach offers some benefits over the spatial analysis approach it also has limitations. For example, factors other than rain can change pixel values, most commonly camera motion and the presence of moving objects in the scene.

[0007] Both of the above approaches have limitations with camera movement. With camera movement, a pixel, over a set of frames, no longer corresponds to the same element in the scene; thus, temporal analysis is no longer valid. To compensate for the

moving camera effect, some have applied frame alignment to the set of analyzed frames prior to rain removal. A set of properly aligned frames can appear as though the camera has not moved throughout the duration of the frames.

[0008] However, frame alignment warps all pixels in a frame using a single transform matrix, according to the global consensus of camera motion. Hence, for scenes with a large depth range and captured under most types of camera motion, frame alignment is unable to properly align all content in the scene. Some have proposed using optical flow to perform alignment at the pixel level, but this approach is very slow and not suitable for real-time applications.

[0009] Although these systems and methods have demonstrated some success in removing rain from video, they have shortcomings. They cannot be used in dynamic settings, such as with a moving camera with moving objects in the field of view.

[0010] Accordingly, there is a need for an improved system that can remove rain from video in dynamic scenes.

SUMMARY OF THE INVENTION

[0011] The invention recognizes that there exists a need for an improved system that can remove rain from video in dynamic scenes.

[0012] The following summary is provided to facilitate an understanding of some of the innovative features unique to the disclosed embodiment and is not intended to be a full description. A full appreciation of the various aspects of the embodiments disclosed herein can be gained by taking into consideration the entire specification, claims, drawings, and abstract as a whole.

[0013] In an embodiment of the present invention, there is provided a method of removing distortion caused by rain or airborne particulates in an image or series of

images. A first image (for rain removal) is segmented into superpixels. A second image of the same scene is also segmented into superpixels. The superpixels can be defined based on perceptually meaningful shapes in each of the scenes. In a preferred method, the first and second images are each segmented into 288 superpixels. The first and second image can be successive frames in a series of images such as a video. Differences between them can be caused by camera movement and/or movement of objects within the scene.

[0014] Each superpixel in the first image is analyzed for the presence of distortion caused by rain or airborne particulates. Regions with distortion can be replaced with regions from corresponding superpixels in the second image. The distortion can be detected by comparing a superpixel in the first image to a corresponding superpixel of the second image.

[0015] In another embodiment, there is provided a method of removing distortions on images comprising the steps of (a) taking a first image, (b) taking at least one additional image of the same scene, (c) dividing each image into superpixels so that the images are segmented into perceptually meaningful shapes based scene contents, (d) analyzing each superpixel in the first image for the presence of one or more distortions, (e) comparing a superpixel with one or more distortions to a corresponding superpixel from the at least one additional image of the same scene and (f) replacing the one or more distortions with content from the at least one additional image of the same scene.

[0016] The superpixels from the additional image can be used to analyze superpixels in the first image for the presence of distortions. The first image can be one of a series of images in a video. The additional images of the same scene can be preceding or subsequent frames in a video. Distortions can be detected by comparing a superpixel region in a first image to that of a superpixel in the same location of the at least one additional image of the same scene.

[0017] In another embodiment, there is provided a method of removing rain from

images or video comprising the steps of (a) obtaining a first image from which to remove rain, (b) obtaining one or more additional images of the same scene, (c) applying frame geometrical alignment to the first image and the one or more additional images to minimize scene content disparity, (d) applying superpixel segmentation to segment the first image into perceptually meaningful shapes based on scene contents, (e) detecting one or more regions distorted by rain in the first image by analyzing each superpixel, (f) removing the one or more regions distorted by rain and (g) replacing the one or more regions distorted by rain with a regions from corresponding superpixels of the one or more additional images of the same scene. The method may include the additional step of creating a spatial-temporal buffer for each superpixel to compensate for scene content motion and depth variation. The method can also include the step of performing matching and filtering operations on images for rain removal. Rain can be detected by comparing each superpixel to one or more counterparts (i.e. corresponding superpixels) from neighbor frames in a spatio-temporal buffer.

INTRODUCTION

[0018] A first aspect of the invention is a system and method for removing distortions caused by rain from still images and video.

[0019] A second aspect of the invention is a system and method for removing distortions caused by rain from video with dynamic settings, such as with a moving camera and moving objects in the field of view.

[0020] A third aspect of the invention is a system and method that utilizes frame alignment at the superpixel level to remove distortions caused by rain from still images and video.

[0021] A fourth aspect of the invention is a system and method that utilizes superpixel segmentation to remove distortions caused by rain from still images and video.

[0022] A fifth aspect of the invention is a system and method of classifying rain by comparing the intensity of a target superpixel to the intensity of reference superpixels.

Brief Description of Figures

[0023] The drawings described herein are for illustrative purposes only of selected embodiments and not all possible implementations, and are not intended to limit the scope of the present disclosure.

[0024] The summary above, as well as the following detailed description of illustrative embodiments, is better understood when read in conjunction with the appended drawings. For the purpose of illustrating the disclosure, exemplary constructions of the disclosure are shown in the drawings. However, the disclosure is not limited to specific methods and instrumentalities disclosed herein. Moreover, the drawings are not to scale.

[0025] FIG. 1A depicts a series of images that are combined using conventional methods of temporal analysis.

[0026] FIG. 1B depicts the geometric alignment that occurs with conventional methods.

[0027] FIG. 2A depicts a scene that has been segmented with rectangular patches.

[0028] FIG. 2B depicts a scene that has been segmented with superpixels.

[0029] FIG. 3 depicts the spatio-temporal buffer of the invention.

[0030] FIG. 4 depicts an example of a template T_{rd} and a matched reference patch $I_{ref(t)}$ within the spatio-temporal buffer.

[0031] FIG. 5A depicts Mask M_{SP} , which indicates the region of a superpixel

wherein dark values indicate the superpixel region.

[0032] FIG. 5B depicts Mask M_{RAIN} , which indicates the rain pixels in the corresponding region of M_{SP} wherein light values indicate rain pixels.

[0033] FIG. 5C depicts Mask M_{match} , which is used for template matching in Rain Recovery.

DETAILED DESCRIPTION OF THE INVENTION

Definitions

[0034] While the invention is primarily described for improving the quality of images and video, it is understood that the invention is not so limited and can be used to assist with other endeavors that require removal of rain or other matter from images or video. Other applications include, for example, include using the invention in video surveillance, post-production editing of photographs/film, as well as industrial and medical imaging.

[0035] Reference in this specification to "one embodiment/aspect" or "an embodiment/aspect" means that a particular feature, structure, or characteristic described in connection with the embodiment/aspect is included in at least one embodiment/aspect of the disclosure. The use of the phrase "in one embodiment/aspect" or "in another embodiment/aspect" in various places in the specification are not necessarily all referring to the same embodiment/aspect, nor are separate or alternative embodiments/aspects mutually exclusive of other embodiments/aspects. Moreover, various features are described which may be exhibited by some embodiments/aspects and not by others. Similarly, various requirements are described which may be requirements for some embodiments/aspects but not other embodiments/aspects. Embodiment and aspect can be in certain instances be used interchangeably.

[0036] The terms used in this specification generally have their ordinary meanings in the art, within the context of the disclosure, and in the specific context where each term is used. Certain terms that are used to describe the disclosure are discussed below, or elsewhere in the specification, to provide additional guidance to the practitioner regarding the description of the disclosure. For convenience, certain terms may be highlighted, for example using italics and/or quotation marks. The use of highlighting has no influence on the scope and meaning of a term; the scope and meaning of a term is the same, in the same context, whether or not it is highlighted. It will be appreciated that the same thing can be said in more than one way.

[0037] Consequently, alternative language and synonyms may be used for any one or more of the terms discussed herein. Nor is any special significance to be placed upon whether or not a term is elaborated or discussed herein. Synonyms for certain terms are provided. A recital of one or more synonyms does not exclude the use of other synonyms. The use of examples anywhere in this specification including examples of any terms discussed herein is illustrative only, and is not intended to further limit the scope and meaning of the disclosure or of any exemplified term. Likewise, the disclosure is not limited to various embodiments given in this specification.

[0038] Without intent to further limit the scope of the disclosure, examples of instruments, apparatus, methods and their related results according to the embodiments of the present disclosure are given below. Note that titles or subtitles may be used in the examples for convenience of a reader, which in no way should limit the scope of the disclosure. Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure pertains. In the case of conflict, the present document, including definitions, will control.

[0039] The term “affine transformation” refers to any transformation that preserves

collinearity (i.e., all points lying on a line initially still lie on a line after transformation) and ratios of distances (e.g., the midpoint of a line segment remains the midpoint after transformation).

[0040] The term “Good-Features-To-track” refers to a method of extracting visual features (e.g. corners, textured areas) and “tracking” them over multiple frames.

[0041] The term “homography transformation” refers to a geo-referencing technique based on geometric and mathematic concepts, known as "homogeneous coordinates" and "projective planes."

[0042] The term “Lucas-Kanade tracker” refers to a technique which can provide an estimate of the movement of interesting features in successive images of a scene.

[0043] The term “M-estimator SAmple Consensus” or “MSAC” refers to a modification of the RANSAC algorithm wherein inliers are scored according to their fitness while outliers are given a constant weight.

[0044] The term “pixel” refers to a minute area of illumination on a display screen, one of many from which an image is composed.

[0045] The term “super-pixel” refers to a polygonal (or other shaped) part of a digital image larger than a normal pixel that is rendered similar color and brightness.

[0046] The term “RANSAC” refers to “Random Sample Consensus” and is a commonly used algorithm in computer vision. It is an iterative method to estimate parameters of a mathematical model from a set of observed data that contains outliers, when outliers are to be accorded no influence on the values of the estimates.

[0047] The term “RGB value” or “Red, Green, Blue value” refers to a color space that is defined by the three chromaticities of the red, green, and blue additive primaries, and can produce any chromaticity that is the triangle defined by those primary colors.

[0048] The term “SLIC superpixel segmentation method” or “Simple Linear Iterative Clustering” refers to a method that clusters pixels in the combined five-dimensional color and image plane space to efficiently generate compact, nearly uniform superpixels.

[0049] The term “Sobel operator,” “Sobel–Feldman operator” or “Sobel filter” refers to an operation commonly used in image processing and computer vision, particularly within edge detection algorithms where it creates an image emphasizing edges.

[0050] The term “Speeded-Up Robust Feature” or “SURF” refers to a local feature detector and descriptor that can be used for tasks such as object recognition, image registration, classification or 3D reconstruction.

[0051] The term “YCbCr image” or “YCbCr image” refers to a family of color spaces used as a part of the color image pipeline in video and digital photography systems.

[0052] Other technical terms used herein have their ordinary meaning in the art that they are used, as exemplified by a variety of technical dictionaries. The particular values and configurations discussed in these non-limiting examples can be varied and are cited merely to illustrate at least one embodiment and are not intended to limit the scope thereof.

Description of Preferred Embodiments

[0053] Conventional systems of rain removal typically break up each image into individual patches or pixels (i.e. image segmentation). Each pixel is analyzed to detect distortions caused by rain drops. If present, the system can adjust the distortion based on aligning the same (or similar) image to determine the proper color and lighting of the pixel. Conventional techniques can be suitable for stationary images but are generally ineffective with video of dynamic scenes, particular with a moving camera.

[0054] For example, with a moving camera, it can be difficult (or impossible) to align multiple images. FIG. 1A and 1B demonstrate the geometrical alignment between two frames. Obvious misalignments exist at different distances. This is encountered with a moving camera and/or moving objects within the field of view.

[0055] In FIG. 1A, the left rectangle frame is aligned with the right rectangle frame. When the two images are superimposed, overlapping is evident as shown in FIG. 1B. The buildings on the left are well-aligned. However, for the lamp post and bench on the left and the car and buildings on the right, obvious mis-alignment appears.

[0056] The present invention utilizes the superpixel (SP) technology. The SP concept is based on grouping pixels into perceptually meaningful atomic regions, in contrast to the rigid rectangular structure that is not adaptive to the scene content. The inventions utilizes superpixel technology with frame alignment to overcome the limitations of conventional techniques.

[0057] FIG. 2A depicts a conventional method of image segmentation that uses rectangular patches. In FIG. 2B, a scene has been segmented into regions using superpixels. It is apparent that the boundaries of SP often coincide with those of the scene contents. The SPs are more adaptive in shape than rectangles and can segment meaningful regions, which are more likely to have uniform scene depth. And as demonstrated in experiments, after unit-wise processing, SPs are much less likely to produce blocky or regional artifacts.

A. Frame Alignment

[0058] Frame alignment is the first step in a preferred use of the invention. Given a video, or a sequence of input frames, rain removal is applied individually to each frame. The currently processed frame is designated the "target frame." In each target frame, each superpixel individually is de-rained. The currently processed superpixel is

designated the “target SP.” To de-rain a target frame, it is first aligned with its temporal neighbors, which are called neighbor frames, to get a set of N_F frames. In a preferred implementation, $N_F = 5$ is set; for target frame t_0 , frames $t_0 - 2$, $t_0 - 1$, $t_0 + 1$ and $t_0 + 2$ to t_0 are aligned. The aim is to first align the scene content at a coarse scale before aligning them at the superpixel scale.

[0059] The target frame is then segmented into SPs, and each SP segment is de-rained individually. The target frame is de-rained in two steps — rain detection and rain recovery. For rain detection, each SP is compared with similar patches in neighbor frames to detect rain pixels. For rain recovery, the rain pixels in each SP are replaced by pixels estimated from similar patches from the target frame and neighbor frames.

[0060] In the first step, a set of neighbor frames is aligned to the target frame. To align a neighbor frame to the target frame, the transform matrix that transforms the neighbor frame to the target frame is estimated, and then the neighbor frame according to this transform matrix is warped. The transform matrix can be a homography transform or an affine transform. In a preferred implementation, the homography transform is used.

[0061] Any transform matrix estimation method can be used, as long as it can robustly estimate the homography or affine transform matrix. In our implementation, the transform matrix $T_{t,t+1}$ that warps frame F_t to frame F_{t+1} is computed as follows. First, a set of Good-Features-To-Track (GFTT) corner feature points in F_t is detected. These points are tracked using the Lucas-Kanade tracker to obtain the corresponding points in frame F_{t+1} . With the two sets of points, RANSAC is used to robustly estimate $T_{t,t+1}$.

[0062] An alternative implementation would be to compute a set of Speeded-Up Robust Feature (SURF) points and their respective descriptors for the F_t and F_{t+1} . The two sets of points are then matched to obtain their correspondence. With the two sets of points, M-estimator Sample Consensus (MSAC) is used to robustly estimate $T_{t,t+1}$.

B. Supapixel Segmentation

[0063] Image segmentation is the process of partitioning a digital image into multiple segments. Numerous image/video processing applications process the data after decomposing larger frames into smaller regions. By doing so, local statistical properties of the visual signal can be better exploited. Examples can be found in image denoising, contrast enhancement, and image/video compression. For most cases, rectangular decomposition are applied as it is easy to implement.

[0064] As discussed, the concept of superpixel (SP) is to group pixels into perceptually meaningful atomic regions, in contrast to a rigid rectangular structure which is not adaptive to the scene content. FIG. 2B depicts a scene that has been segmented into regions using superpixels.

[0065] In a preferred implementation, the SLIC superpixel segmentation method is used, which is an iterative regional pixel clustering algorithm that clusters each pixel to an initiated center grid according to the their respective normalized spatial and color distances. (See, e.g. R. Achanta et al., "Slic superpixels compared to state-of-the-art superpixel methods," IDDD Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 11, pp. 2274-2282, 2012). The number of SPs created is a parameter that can be adjusted depending on the target application and the scene captured by the input video. In a preferred method, the number of SPs in each frame is set to 288, which works well for videos taken from a car-mounted camera. However, other SP segmentation algorithms can also be used.

C. Rain Detection

[0066] At this point, a set of aligned frames and a SP segmentation of the target frame is obtained and subsequently the rain frame F_{rain} is to be computed. For each SP in the target frame, rain is detected by comparing each SP to its counterparts (i.e. corresponding SP) from neighbor frames in a spatio-temporal buffer. These

counterparts, which can be called reference SPs, are SPs with the same or very similar content as the target SP. A spatio-temporal buffer, which is created for each SP, is a set of image regions in the aligned frames that are centered over the target SP. This is graphically depicted in FIG. 3. This buffer specifies the regions in the aligned frames where the reference SPs are to be searched. Reducing the spatial size of the buffer improves the computational speed of the search process. However, it also reduces the chances of finding the optimal reference SPs. In our implementation with five aligned frames, the target SP is located in frame three of the buffer.

[0067] Searching for reference SPs can be done by template matching of the target SP with each neighbor frame in the spatio-temporal buffer. As showed in FIG. 4, an enlarged bounding box of the target SP may be used as the template instead of just the SP region. This can improve matching for SPs that do not contain strong features. The image patch in each neighbor frame of the spatio-temporal buffer that is most similar to the template, T_{rd} may be searched for.

$$I_{ref}(t) = \min_{x,y} ||T_{rd} - I(x,y,t)||_2$$

where I is an image patch the size of template T_{rd} at position (x, y) and frame t within the spatio-temporal buffer. Each I_{ref} , is called a reference patch and contains a reference SP. Note that $I_{ref}(t)$ for $t = 3$ is not computed because the comparison to the scene content from neighboring frames is only desired.

[0068] Because rain streaks increase the intensity of the affected pixels, rain may be classified by comparing the intensity of the target SP to the intensity of the reference SPs. Firstly, the binary intermediate masks $M_{int}(t)$ is computed to find pixels with lower intensity in $I_{ref}(t)$, i.e.,

$$M_{int}(t) = (T_{rd} - I_{ref}(t)) \geq \epsilon_{rain}$$

Rain pixels in the target SP are expected to have higher intensity than their corresponding pixels in all reference SPs. However, rain streaks can re-occur over the same pixels in the reference SPs. Hence, pixels may be classified as rain pixels if they have lower intensity in either three or four reference patches, i.e.,

$$M_{\text{rain}} = \sum_{t \neq 3} M_{\text{int}}(t) \geq 3$$

The values in M_{rain} for the target SP are then copied to F_{rain} . This rain detection step is repeated for every SP in the target frame. Thus, the rain frame is obtained.

$$F_{\text{rain}}(x,y) = \begin{cases} 1 & (x,y) \text{ is a rain pixel} \\ 0 & \text{otherwise} \end{cases}$$

[0069] It is possible for edges of background scenes to be misclassified as rain. These false rain pixels are unclassified with the following procedure. Firstly, the RGB values of the target frame are converted into a YCbCr image. The first order horizontal and vertical derivatives of chroma channels Cb and Cr are extracted using the Sobel operator. The four gradient images are summed together to form a gradient magnitude image F_{grad} . For each pixel (x, y) , if $F_{\text{rain}}(x, y) = 1$ and $F_{\text{grad}}(x, y) \geq \epsilon_{\text{grad}}$, then $F_{\text{rain}}(x,y)$ is set to 0.

[0070] In a suggested implementation, the size of the spatio-temporal buffer is set to 30 pixels x 30 pixels x 5 frames. Template T_{rd} is set to 1.5 times the size of the bounding box of the target SP. ϵ_{rain} is set to 0.012 while ϵ_{grad} is set to 0.2.

D. Rain Recovery

[0071] After computing the rain frame F_{rain} , the rain pixels are recovered by estimating their value from similar samples. As in rain detection, rain recovery is done individually for each SP segment, and pixel values are estimated from reference SPs found from the spatio-temporal buffer. The spatio-temporal buffer is created for the

target SP the same way as described above.

[0072] Unlike in rain detection, reference SPs are searched for using masked template matching. The template used in rain recovery, T_{rr} , is the region of the target frame specified by the bounding box of the target SP. A matching mask M_{match} is also specified, which is binary, is the size of T_{rr} , and indicates the non-rain pixels of the target SP in T_{rr} . FIG. 5A depicts a mask M_{SP} , the region of a SP. FIG. 5B depicts Mask M_{rain} , rain pixels in the corresponding region of M_{SP} . Light values indicate rain pixels. FIG. 5C depicts Mask M_{match} , which is used for template matching in rain recovery. Dark values in the figure of M_{match} correspond to a value 1, and the white values correspond to a value of 0.

[0073] Image patches are searched for in the spatio-temporal buffer that are most similar to T_{rr} , subject to M_{match} . A matching error $E(x, y, t)$ of patch $I(x, y, t)$ is defined in the spatiotemporal buffer as

$$E(x,y,t) = \frac{\| (T_{rr} - I(x,y,t)) \odot M_{match} \|_2}{\sum M_{match}}$$

where $\odot$ denotes pixel-wise multiplication, $\sum M_{match}$ is the number of masked pixels in $\sum M_{match}$, and $E(x, y, t)$ is the average matching error of masked pixels. The reference patch set $I_{ref} = \{I\}$ contains the patches with the lowest $E(x, y, t)$ and with $E(x, y, t) \leq \epsilon_{err}$. The maximum number of reference patches is set to N_{ref} . It is possible that fewer than N_{ref} patches are obtained. In the case where I_{ref} is empty, the condition $E(x, y, t) \leq \epsilon_{err}$ is ignored and N_{ref} patches with the lowest $E(x, y, t)$ in I_{ref} are included.

The estimated rain pixel values are averaged from I_{ref} , i.e.,

$$I_{reco} = \frac{1}{|I_{ref}|} \sum I_k, \quad I_k \in I_{ref}$$

The de-rained patch is calculated as

$$I_{\text{derain}} = I_{\text{reco}} \odot M_{\text{rain}} + T_{\text{rr}} \odot \bar{M}_{\text{Rain}}$$

where $\bar{M}_{\text{Rain}}$ is the complement of rain mask M_{Rain} and $\odot$ denotes pixel-wise multiplication.

[0074] The values in I_{derain} corresponding to the target SP are copied to the final de-rained frame. This rain recovery step is repeated for every SP in the target frame to get the final de-rained frame. In a preferred implementation, N_{ref} is set to 10 and ϵ_{err} is set to 0.01.

WORKING EXAMPLE

Removal of Rain and Related Distortions from Video

[0075] With a moving camera, image translation, rotation, and scaling can occur from frame to frame. This makes rain detection a difficult challenge because it requires scene contents between different frames to be perfectly aligned. Regardless of movement, rain is removed according to the following steps.

A. Frame Alignment

[0076] In our proposed method, content alignments are performed on each SP level. The concept of Super-Pixel (SP) is to group pixels into perceptually meaningful atomic regions instead of a rigid rectangular structure. The SPs are more adaptive to shape than rectangles and can segment meaningful regions, which are more likely to have uniform scene depth, therefore, misalignments of SPs between frames are minimal to none.

B. Superpixel Segmentation

[0077] The presence of dynamic content in a scene, (e.g., moving cars) can cause or intensify fluctuations that can be mistaken for rain. In previous methods, dynamic scene content is identified and segmented out to be excluded from rain removal.

However, this step does not remove rain that covers dynamic scene content.

[0078] In the framework described herein, SP segmentation can naturally segment dynamic scene content from static scene content. Because each SP is aligned individually, dynamic content can be aligned just as robustly as static content. Thus dynamic content need not be differentiated from nor treated differently from static content; rain streaks that covers both static content and dynamic content are removed in the same manner. Thus, unlike other methods, the framework described herein does not require extra consideration for dynamic scene content.

C. Rain Detection

[0079] Because rain streaks increase the intensity of the affected pixels, rain may be classified by comparing the intensity of the target SP to the intensity of the reference SPs.

D. Rain Recovery

[0080] After detecting the rain, the rain pixels are recovered by estimating their value from similar samples. As in rain detection, rain recovery is done individually for each SP segment, and pixel values are estimated from reference SPs found from the spatio-temporal buffer.

[0081] It will be appreciated that variations of the above disclosed and other features and functions, or alternatives thereof, may be combined into other systems or applications. Also, various unforeseen or unanticipated alternatives, modifications, variations or improvements therein may be subsequently made by those skilled in the art which are also intended to be encompassed by the following claims.

[0082] Although embodiments of the current disclosure have been described comprehensively, in considerable detail to cover the possible aspects, those skilled in the art would recognize that other versions of the disclosure are also possible.

What is claimed is:

1. A method of removing distortion caused by rain or airborne particulates in images,
wherein a first image is segmented into superpixels,
wherein a second image of the same scene is segmented into superpixels,
wherein each superpixel in the first image is analyzed for the presence of distortion caused by rain or airborne particulates, and
wherein regions with distortion caused by rain or airborne particulates are replaced with regions from corresponding superpixels in the second image.
2. The method of claim 1, wherein distortion caused by rain or airborne particulates is detected by comparing a superpixel in the first image to a corresponding superpixel from the second image.
3. The method of claim 1, wherein superpixels are defined based on perceptually meaningful shapes in scenes.
4. The method of claim 1, wherein the first image is one of a series of images in a video.
5. The method of claim 1, wherein the second image is previous or subsequent frame of the first image in a video.
6. The method of claim 1, wherein at least one of camera movement and movement of objects in a scene cause differences between the first image and second image.
7. The method of claim 1, wherein the first and second images are each segmented into 288 superpixels.

8. A method of removing distortions on images comprising the steps of:
 - a) taking a first image;
 - b) taking at least one additional image of the same scene;
 - c) dividing each image into superpixels so that the images are segmented into perceptually meaningful shapes based scene contents;
 - d) analyzing each superpixel in the first image for the presence of one or more distortions;
 - e) comparing a superpixel with one or more distortions to a corresponding superpixel from the at least one additional image of the same scene; and
 - f) replacing the one or more distortions with content from the at least one additional image of the same scene.
9. The method of claim 8 wherein superpixels from the at least on additional image are used in analyzing superpixels in the first image for the presence of one or more distortions.
10. The method of claim 8, wherein the first image is one of a series of images in a video.
11. The method of claim 8, wherein the at least one additional image of the same scene is a preceding or subsequent frame of the first image in a video.
12. The method of claim 8, wherein the distortions are caused by rain or particulate matter in the air.
13. The method of claim 8, wherein distortions are detected by comparing a superpixel region in a first image to that of a corresponding superpixel of the at least one additional image of the same scene.
14. The method of claim 8, wherein at least one of camera movement and

movement of objects in a scene cause differences between the first image and the at least one additional images of the same scene.

15. The method of claim 8, wherein the number of superpixels in each image is 288.

16. A method of removing rain from images or video comprising the steps of:

- a) obtaining a first image from which to remove rain;
- b) obtaining one or more additional images of the same scene;
- c) applying frame geometrical alignment to the first image and the one or more additional images to minimize scene content disparity;
- d) applying superpixel segmentation to segment the first image into perceptually meaningful shapes based on scene contents;
- e) detecting one or more regions distorted by rain in the first image by analyzing each superpixel;
- f) removing the one or more regions distorted by rain; and
- g) replacing the one or more regions distorted by rain with a regions from corresponding superpixels of the one or more additional images of the same scene.

17. The method of claim 16, including the additional step of performing matching and filtering operations on images for rain removal.

18. The method of claim 16, including the additional step of creating a spatial-temporal buffer for each superpixel to compensate for scene content motion and depth variation.

19. The method of claim 18, wherein rain is detected by comparing each superpixel to one or more counterparts from neighbor frames in the spatio-temporal buffer.

20. The method of claim 16, wherein at least one of camera movement and movement of objects in a scene cause differences between the first image and the one or more additional images of the same scene.

21. The method of claim 16, wherein the number of superpixels in each image is 288.