Title: CAMERAS AND DEPTH ESTIMATION OF IMAGES ACQUIRED IN A DISTORTING MEDIUM

Abstract: The invention provides a method for depth estimation in image or video obtained from a distorting medium. In the method, a pixel blurriness map is calculated. A rough depth map is then determined from the pixel blurriness map while assuming depth in a small local patch is uniform. The rough depth map is then refined. The method can be implemented in imaging systems, such as cameras or imaging processing computers, and the distorting medium can be an underwater medium, haze, fog, low lighting, or a sandstorm, for example. Preferred embodiments determined the blurriness map by calculating a difference between an original image and multi-scale Gaussian-filtered images to estimate the pixel blurriness map.
CAMERAS AND DEPTH ESTIMATION
OF IMAGES ACQUIRED IN A DISTORTING MEDIUM

PRIORITY CLAIM AND REFERENCE TO RELATED APPLICATION

[001] The application claims priority under 35 U.S.C. §119 and all applicable statutes and treaties from prior provisional application serial number 62/220,306, which was filed September 18, 2015.

STATEMENT OF GOVERNMENT INTEREST

[002] The invention was made with support under grant number CCF-1160832 awarded by the National Science Foundation. The government has certain rights in the invention.

FIELD

[003] Fields of the invention include camera systems and image reconstruction, particularly for images and videos acquired in a distorting medium. Particular applications of the invention include underwater photography and video and reconstruction of images and videos acquired underwater.
BACKGROUND

[004] Underwater images or videos often suffer from visible degradation and color distortion due to the propagated light attenuated with distance from the camera, primarily resulting from absorption and scattering effects. Acquired images and videos often suffer from poor contrast and details of images and video frames are often lost, even after efforts toward reconstruction of the images or videos. Similarly, other distorting media, such as fog, haze and other liquids created visible degradation and color distortion from absorption and scattering effects.

[005] Prior methods for reconstruction of underwater images estimate depth based upon particular color channels. Some previous methods use the image formation model (IFM) to estimate the background light and the transmission map based on depth estimation. These methods still provide images with poor contrast. Some of these methods use a dark channel prior (DCP) to derive scene depth and remove haze in natural terrestrial images by calculating the amount of spatially homogeneous haze using the darkest channel in the scene. H. Yang, P. Chen, C. Huang, Y. Zhuang and Y. Shiau, "Low complexity underwater image enhancement based on dark channel prior," Int. Conf. Innov. in Bio-inspired Comput. and App. (IBICA), pp. 17-20, Dec. 2011; L. Chao and M. Wang, "Removal of water scattering," in Proc. IEEE Int. Conf. Comput. Engin. and Tech. (ICCET) vol. 2, pp. 35-39 (Apr. 2010). With the DCP approach, points in the scene closer to the camera have a shorter path over which scattering occurs, and dark pixels therefore experience less brightening from scattered light. This observation is used to estimate depth.


The inventors have determined that light absorption and different lighting conditions existing in underwater images and other distorting media can create exceptions to those priors. When such exceptions exist, the resulting image restoration based on the priors can be poor.

SUMMARY OF THE INVENTION

An embodiment of the invention is a method for depth estimation in image or video obtained from a distorting medium. In the method, a pixel blurriness map is calculated. A rough depth map is then determined from the pixel blurriness map while assuming depth in a small local patch is uniform. The rough depth map is then refined. The method can be implemented in imaging systems, such as cameras or imaging processing computers, and the distorting medium can be an underwater medium, haze, fog, low lighting, or a sandstorm, for example. Preferred embodiments determined the blurriness map by calculating a difference between an original image and multi-scale Gaussian-filtered images to estimate the pixel blurriness map.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 illustrates a preferred embodiment system for depth estimation of images acquired in a distorting medium;
FIG. 2 provides pseudo code for selection of separate BL (background light) for separate color channels that is used in a preferred embodiment depth estimation method of the invention; and

FIG. 3 shows a set of weighted combination curves illustrating (from left to right) curves plotted using different \( s = 8, s = 16, \) and \( s = 32 \) for calculating the BL estimate.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Preferred embodiments of the invention include camera systems that acquire still images and/or video. The systems include software stored on a non-transient medium for reconstructing images and videos to enhance the images and videos from distortion caused by a distorting surrounding medium. Systems and methods of the invention are applicable, for example, to underwater photography and video acquisition, and to photography and video acquisition from other distorting media, such as fog, haze, dimly lit images and sandstorm images. Methods of the invention can be performed within cameras or video acquisition equipment, and can also be performed by image/video processing and adjustment software on non-transient medium executed by a computer.

Preferred methods for depth estimation of distorted medium images in video use image blurriness to estimate scene depth. A preferred embodiment leverages a recognition of the inventors that larger scene depth causes more object blurriness for distorted medium images, e.g. underwater images. Preferred methods estimate background light first, and then perform depth estimation and transmission map estimation based on the estimated background light so as to adaptively produce the more accurate scene depth and transmission map for an underwater image. The inventors have determined that background light is a key individual factor to decide the
effectiveness of the depth estimation methods based on color channels for underwater images. This depth estimation can enhance image reconstruction.

[0014] Additional preferred methods for reconstruction provide a unified medium depth estimation based on image blurriness and light absorption. The extended method provides unified underwater depth estimation based on image blurriness and light absorption that works well in different lighting conditions and hues of water. Unlike other methods, preferred methods of the invention perform depth estimation based on background light estimation so as to adaptively produce the more accurate scene depth for underwater image restoration.

[0015] Preferred methods of the invention have been shown to work well in different lighting conditions and hues of water. Preferred methods use background light estimation to estimate scene depth and can therefore adaptively produce an accurate scene depth for image restoration, of images acquired in a distorting medium such as water.

[0016] A preferred depth estimation method makes a pixel blurriness estimation. A difference between an original and a multi-scale Gaussian-filtered image is calculated to estimate the pixel blurriness map. A rough depth map is generated by applying a maximum filter to the pixel blurriness map while assuming depth in a small local patch is uniform. Closing by morphological reconstruction and guided filtering are then used to refine the depth map.

[0017] Those knowledgeable in the art will appreciate that embodiments of the present invention lend themselves well to practice in the form of computer program products. Accordingly, it will be appreciated that embodiments of the present invention may comprise computer program products comprising computer executable instructions stored on a non-transitory computer readable medium that, when executed, cause a computer to undertake methods according to the present invention, or a computer configured to
carry out such methods. The executable instructions may comprise computer
program language instructions that have been compiled into a machine-
readable format. The non-transitory computer-readable medium may
comprise, by way of example, a magnetic, optical, signal-based, and/or
circuitry medium useful for storing data. The instructions may be
downloaded entirely or in part from a networked computer. Also, it will be
appreciated that the term "computer" as used herein is intended to broadly
refer to any machine capable of reading and executing recorded instructions.

It will also be understood that results of methods of the present invention
may be displayed on one or more monitors or displays (e.g., as text, graphics,
charts, code, etc.), printed on suitable media, stored in appropriate memory
or storage, etc.

[0018] Preferred embodiments are discussed with respect to the RGB color space.
Artisans will appreciate that the invention is applicable to arbitrary color
spaces.

[0019] Preferred embodiments use image blurriness to estimate scene depth. One
technique for estimating blurriness is provided in S. Bae and F. Durand,
pp. 571-579 (2007). This is a scale edge detector to estimate pixel blurriness,
and then defocus map is generated by using edge aware interpolation. See,
A. Levin, D. Lischinski, and Y. Weiss, "Colorization using optimization,"
ACM Trans, on Graphics 23(3), pp. 689-694 (2004). In constructing the
defocus map, the blurriness at non-edge pixels is interpolated and propagated
by neighbouring edge pixels based on the similarities of luminance. If
matting Laplacian [A. Levin, D. Lischinski, and Y. Weiss, "A Closed-Form
vol. 30, no. 2, pp. 228 - 242, (Feb. 2008)] is applied to perform the
interpolation, noise and low contrast may cause incorrect blurriness
propagation, especially for underwater image. The edge-aware

Preferred embodiments of the invention will now be discussed with respect to the drawings and with respect to experiments that have been conducted to demonstrate the invention. The drawings may include schematic representations, which will be understood by artisans in view of the general knowledge in the art and the description that follows. Features may be exaggerated in the drawings for emphasis, and features may not be to scale. Artisans will understand broader aspects of the invention from the experiments and experimental data.

FIG. 1 illustrates a preferred embodiment system of the invention. A camera 10 obtains images or video. The camera 10 is an underwater camera in preferred embodiments, and includes a sealed housing 12. The camera further includes memory 14 for storing images, and a processor 16 for controlling camera functions. The processor 16 also can run software that processes images or video that is obtained prior to and after storing, and can control transfer of images via wired or wireless connection to another devices, such as a computer 18. The computer 18 can include image processing/adjustment software. Either or both of the processor 16 and the computer 18 can perform present methods for estimating depth and image restoration. The camera 10 also includes a lens 20, light meter 22 and operator controls 24.

The camera 10 is illustrated in a distorting medium of water 26 that separates the lens 20 from a scene 28. This makes the observed intensity $I$ a distorted version of the scene radiance $J$. Specifically,

$$I(x) = J(x)t(x) + B^c(1 - t(x)), \quad (1)$$
where $I_i$ is the observed intensity of the input image at pixel $x$, $J_i$ is the scene radiance, $B^c_i$ is the background light (BL), and $t$ is the transmission medium map that describes the portion of the scene radiance that is not scattered or absorbed and reaches the lens 20 of camera 10. In this model, the transmission map is the depth map, i.e., smaller $t$ means the scene point is farther from the camera.

Prior DCP methods are focused upon outdoor scenes taken above ground and recover the scene radiance $J$ with a statistical observation that about three quarters of the pixels in non-sky parts of outdoor images have zero values in the dark channel. This can be written as follows (taken from K. He et. al, "Single image haze removal using dark channel prior," IEE Trans. Pattern Anal. Mach. Intell., vol. 33, no 12, pp. 2341-53 (Dec. 2011).

$$\min_{y \in \Omega(x)} \left\{ \min_{c \in \{R,G,B\}} J^c(y) \right\} = 0, \quad (2)$$

where $\Omega(x)$ is the local patch centered at $x$, and $J^c_i$ is the intensity of the scene radiance in channel $c$, where $c$ is one of the R, G, B channels. The minimum operator of Eq. 2 can be applied to both sides of Eq. 1, which produces:

$$t(x) = 1 - \min_{y \in \Omega(x)} \left\{ \min_{c \in \{R,G,B\}} \frac{I^c(y)}{B^c} \right\}. \quad (3)$$


The dark channel of the input image can be defined as $l_{dark}(x) = \min_{y \in \Omega(x)} \{ \min_c I^c(y) \}$. To estimate the BL $\overline{B}^c$, the top 0.1% brightest pixels in $l_{dark}$ are selected. The corresponding intensities in the input image are
averaged. This effectively obtains BL from the 0.1% most distant scene points. If \( S_{0.1} \) is the set of the 0.1% brightest pixels, then BL can be calculated as

\[
\overline{B^c} = \frac{1}{|S_{0.1}|} \sum_{x \in S_{0.1}} I^c(x), \quad c \in \{r, g, b\}, \quad (4)
\]

Then, the scene radiance \( J \) can be recovered by placing \( I, t, \) and \( \beta \) into Eq. 1. as follows:

\[
J(x) = \frac{I(x) - \overline{B^c}}{\max(I(x), r_0)} + \overline{B^c}, \quad (5)
\]

where \( r_0 \) is typically set ranging from 0.3 to 0.5. The value is chosen empirically to increase the exposure of \( J \) for display. This prior DCP method is effective in some cases but fails in others. The failure case can produce images missing a great amount of detail, especially when a foreground object (close object) has very bright features, which are mistakenly treated as being distant.

Preferred methods of the invention avoid such issues. In a preferred method, a blurriness estimate is obtained. The blurriness estimate can be determined from prior technique for blurriness, including those mentioned above. Preferred methods include multi-scale low-pass filtering to determine blurriness and multi-scale Gaussian-filtered images to determine blurriness. Another approach is to measure blurriness in the frequency domain. The frequency domain approach. See, Kanjar and Masilamani, "Image Sharpness Measure for Blurred Images in Frequency Domain," International Conference on Design and Manufacturing, Volume 64, Pages 149-158 (2013).

The preferred Gaussian-filtered images will be discussed as a detailed example. In the example, a difference between an original and multi-scale Gaussian-filtered images is calculated to estimate a pixel blurriness map. A rough depth map is then generated by applying a maximum filter to the pixel blurriness map while assuming depth in a small local patch is uniform. The
rough depth map is refined, such as by guided filtering or image matting. See, K. He, "Guided Image Filtering," IEEE Trans. Pattern. Anal. Mach. Intell., vol. 35, no. 6, pp. 228-242 (Oct. 2012); Levin et al, "A Closed-Form Solution to Natural Image Matting," IEEE Trans Pattern Anal Mach Intell. 30(2):228-42 (2008). Experiments compared the present method to the dark channel prior methods (DCP) maximum intensity prior (MIP) discussed in the background. Qualitative and quantitative measures were evaluated in different underwater lighting conditions. The experiments showed that the depth estimation based upon blurriness worked well for a variety of images, and provided better results with reasonable computation costs, whereas the prior methods sometimes fail to distinguish between background and foreground and sometimes cause color distortion and poor global contrast.

The pixel blurriness map can be calculated as follows. $G^{k,\sigma}$ is the input image filtered by a $kx$. $k$ spatial Gaussian filter with variance $\sigma^2$. The pixel blurriness map $P_{init}$ is computed as:

$$P_{init}(x) = \frac{1}{n} \sum_{i=1}^{n} \left( | I_g(x) - G^r \Omega(x) | \right)$$

where $r_i = 2^{l_n} + 1$ and $I_g$ is the grayscale version of the input image $I_c$. As an example, we can assign $n = 4$. This assignment can be altered based upon factors that include image resolutions and image characteristic (quality, sharpness, etc.). The assignment can also be dependent upon the precision of blurriness estimation. There is a tradeoff between precision and computational cost. Generally, a larger $n$ results in a more precise estimation and higher computation cost than a lower $n$. A preferred range of $n$ for a 1280 x 720 image is 1-6.

The max filter can be applied to $B$ to calculate a rough depth map $P_r$ as:

$$P_r(x) = \max_{y \in \Omega(x)} P_{init}(y),$$

where $\Omega(\chi)$ is a $z x z$ local patch centered at $x$. As an example, we set $z = 7$. More generally, a patch size of $z = 7$ to $z - 31$ works well for image
sizes ranging from 800 x 600 to 1280 x 720. The value of z is can be
determined by image resolutions and image contents, such as the size of
objects in the image. For example, a larger z does not work well for small
objects, which could cause halo effects around such objects in the restored
image. According to experiments that were conducted, we found a patch size
of z=7 to z=31 works well for image sizes ranging from 800 x 600 to 1280 x
720. Generally, an image with a higher resolution requires a larger z, i.e., z
preferably increases with image resolution. The rough depth map \( P_r \) can then
be refined. One option for refinement is hole filling, which can fill holes
caused by the flat regions in objects during the closing by morphological
reconstruction (CMR).

Specifically, refine \( P_r \) by filling holes caused by flat regions in objects using
morphological reconstruction followed by soft matting or guided filtering
applied to generate a refined blurriness map \( P_{btr} \)

\[
P_{btr}(x) = F_g[C_r[P_r(x)]], \tag{8}
\]

where \( C_r \) is a hole-filling morphological reconstruction and \( F_g \) represents
soft matting or guided filtering.

Preferred embodiments modify the depth estimation above. This
modification improves depth estimation, especially when red light content is
significant. In this modification, a candidate selection process chooses from
a plurality of candidate blurriness regions. There can be many regions that
consist of a top predetermined percentage of blurry pixels, which are global
to the image. A preferred candidate selection process selects from a top
predetermined percentage of blurry pixels in the input image, a variance
region within a predetermined lowest variance range and a blurriness region
within a predetermined largest blurriness range. A specific preferred
candidate selection process selects from among a predetermined number of
background light (BL) candidates that are within a predetermined percentage
of the top most blurry pixels, the lowest variance region and the largest blurriness region. The BL candidates are average values of the candidate regions. If three BL candidates are chosen, these correspond to three average pixel values of three BL candidate regions. For example, a preferred embodiment selects three BL candidates from the top 0.1% blurry pixels in the input image, the lowest variance region and the largest blurriness region. These candidate regions are preferably of equal size. For example, a preferred technique divides the input image $I$ into four equally sized regions $I_q^1, I_q^2, I_q^3, and I_q^4$ until \( \frac{\text{Size}(I_q)}{\text{Size}(I)} \leq \epsilon_s \), where $\epsilon_s = 2^{-10}$. The lowest variance region and the largest blurriness region (which may or may not be the same) can be determined, for example, by quadtree decomposition, which iteratively divides the input into a predetermined number of equal size blocks according to variance or blurriness, for example four equal sized blocks. The blurriness of a region can be obtained by averaging $P_{blr}(x)$ in the corresponding region in the blurriness map.

After selection of the predetermined BL candidates, in this example, separate BL is then selected for each color channel according to the input image. FIG. 2 shows pseudo code for the selection. The code of FIG. 2 describes a preferred background light estimation based on image pixel variance and blurriness. The sigmoid function $S$ in FIG. 2 is given by

$$S(a, \nu) = \left[1 + e^{-s(a-\nu)}\right]^{-1}$$ (9)

where $s$ is an empirical constant. In an example, $s$ was set to 32. If $s$ is small, the BL estimate is calculated using a smoother weighted combination. FIG. 3 shows a set of curves illustrating (from left to right) curves plotted using $s = 8$, $s = 16$, and $s = 32$. Fixed thresholds are set with reference to the input image. For example, $\nu$ is a threshold (we assign $\nu = e_n$) that determines whether the input image was taken under sufficient lighting. According to
the percentage of bright pixels \((I^k > 0.5)\), \(\left( \frac{|I^k > 0.5|}{\text{Size}(I^k)} \gg \epsilon_n \right)\) represents the image has sufficient lighting. The value of \(\epsilon_n\) can be determined empirically. Example threshold used in experiments were \(\epsilon_S = 2^{-10}\) and \(\epsilon_n = 0.2\). The quadrant selection lowest variance (QUAD-SELECT-LV) is similar to the quadrant selection with the largest blurriness (QUAD-SELECT-LB), except that largest blurriness is replaced with the lowest variance and there is no need to consider \(P_{blir}\).

The function BL-ESTIMATE in FIG. 2 determines BL for each color channel between the darkest and brightest BL candidates according to the percentage of bright pixels, which for the example is the top 0.5 \% \((I^k > 0.5)\). When the percentage of bright pixels in the input image is high enough \(\left( \frac{|I^k > 0.5|}{\text{Size}(I^k)} \gg \epsilon_n \right)\), then the image is deemed to have sufficient lighting, and BL estimated as being brighter is more suitable. Generally, BL should be estimated as being brighter (with regard to the channel \(k\)) than most (at least a majority and preferably 50-90\% or more) of pixels in the input image. The image is determined as having insufficient lighting with \(\left( \frac{|I^k > 0.5|}{\text{Size}(I^k)} \ll \epsilon_n \right)\), and BL is estimated as darker. BL should be estimated as being darker (with regard to the channel \(k\)) than most (at least a majority and preferably 50-90\% or more) of pixels in the input image. Between these extremes, the BL estimate can be calculated by a weighted combination of the darkest and brightest BL candidates. This selection process for the BL estimate followed by restoration has demonstrated visually pleasing results. Objective measures also reveal good performance.

In another modification, both image blurriness and light absorption are leveraged for the depth estimation. This method can be considered to leverage three different depth estimations that are then sigmoidally combined in view of lighting and image conditions.
A first estimate of depth is obtained directly from the red channel. The red channel map \(R\) is defined as:

\[
R(x) = \max_{y \in \Omega(x)} I^r(y),
\]

(10)

From this, an estimate of depth is obtained:

\[
\widetilde{d}_R = 1 - F_S(R)
\]

(11)

where \(F_S\) is a stretching function. The stretching function \(F_S\) can be defined as:

\[
F_S(V) = \frac{V - \min(V)}{\max(V) - \min(V)},
\]

(12)

where \(V\) is a vector.

A second estimate of depth is:

\[
\widetilde{d}_D = 1 - F_S(D_{mip})
\]

(13)

This estimation is based upon maximum intensity prior (MIP). The MIP can be determined through the different between the maximum intensity of the red channel compared to the green and blue channels, as follows:

\[
D_{mip}(x) = \max_{y \in \Omega(x)} I^r(y) - \max_{y \in \Omega(x)} \{I^g(y), I^b(y)\}
\]

(14)

Large values of \(D_{mip}(x)\) represent closer scene points having red light that attenuates less than that of farther scene points. This concept was applied by others to estimate the transmission map, while it is adapted here to estimate the depth directly.

A third estimation uses the image blurriness \(P_r\) of Equation (7) to estimate depth. The third depth estimation based upon blurriness is defined as:

\[
\widetilde{d}_B = 1 - F_S(C_r(P_r))
\]

(15)

Combining the three estimates of questions (11), (13) and (15) provides for the estimation of depth in a distorted medium based upon light absorption and image blurriness according to the estimated BL \(\widetilde{B}\) and the average input red value according to the following:

\[
\widetilde{d}(x) = \Theta_b \left[ \theta_a \widetilde{d}_D(x) + (1 - \theta_a) \widetilde{d}_R(x) \right] + (1 - \Theta_b) \widetilde{d}_B(x)
\]

(16)
where $\Theta_a = S(\text{avg}_c(B^c), 0.5)$ and $\Theta_b = S(\text{avg}_c(I^r), 0.1)$ are determined by the sigmoid function of Equation (9). After that the depth map is refined and smoothed, such as by soft matting or guided filtering, as discussed above. The resultant estimated depth map $d_n E [0,1]$ can be considered as a map of normalized relative distance for the scene points of the image for any other image enhancement or adjustment processes to be conducted.

With this third variation, when the image has some reasonable level of red content overall, e.g., $\text{avg}(I^r) \gg 0.1$, and the background light is relatively dim, $\text{avg}(B^c) \ll 0.5$, then $d_R$ alone provides a good depth representation. In such case, $\Theta_a \approx 1$ and $\Theta_b \approx 1$, and $d_n(x) \approx d_R(x)$.

As BL gets brighter, the possibility that $d_R(x)$ fails to represent scene depth gets higher. Because the BL accounts for more of the observed intensity for a scene point farther from the camera, far scene points may still have large values in the red channel and be wrongly judged as being close according to Equation (11). When an underwater image has a brighter BL, then $d_D$ is more reliable for scene depth. The red light of a farther scene point is absorbed more compared to the green and blue light. For this reason, when the image has some reasonable level of red content overall $\text{avg}(I^r) \gg 0.1$ and the background is relatively bright $(B^c \gg 0.5)$ then $d_D$ is itself a good measure of depth. In such case, $\Theta_a \approx 1$ and $\Theta_b \approx 1$, and $d_n(x) \approx d_D(x)$.

When there is very little red light $\text{avg}(I^r) \ll 0.1$ then both Equations (13) and (15), which directly use red channel values are likely to fail to estimate scene depth properly. In this case, and $\Theta_a \approx 0$, and $d_n(x) \approx d_D(x)$, mean that the depth estimation reverts to using the blurriness map alone, between these various extremes, the depth map comes from a weighted combination of the three approaches.

With the depth map determined, TM can then be calculated and scene radiance can be recovered. Preferred embodiments determine TM by
measuring the distance from the camera to each scene point. To measure the
distance from the camera to each scene point, the distance $d_0$ between the
closest scene point and the camera must be estimated as well. Via the
maximum difference between the estimated $\tilde{B}^c$ and the observed intensities
$I^c$ in the input image, the estimated $\tilde{d}_0 \in [0,1]$ can be calculated by:

$$\tilde{d}_0 = 1 - \max_{x, c \in \{r, g, b\}} \frac{\max \{|\tilde{B}^c - I^c|\}}{\max(\tilde{B}^k, 1 - \tilde{B}^k)}$$  \hspace{1cm} (17)

where $k = \text{argmax}_{c \in \{r, g, b\}}(\max |\tilde{B}^c - I^c|)$. If the BL accounts for a large
portion of the observed intensities for the closest scene point, the maximum
difference would be small, and $\tilde{d}_0$ would be large indicating that the distance
from the camera to the closest object in the scene is long.

The final scene depth $\tilde{d}_f$ can be determined by combining Equations (16) and
(17) as follows:

$$\tilde{d}_f(x) = D_\infty \times (\tilde{d}_n(x) + \tilde{d}_0)$$  \hspace{1cm} (18)

where $D_\infty$ is a scaling constant for transforming the relative distance to the
actual distance.

With $\tilde{d}_f$, the TM for the red channel can be calculated as:

$$\tilde{t}^r(x) = e^{-\beta^r d_f(x)}$$  \hspace{1cm} (19)

where $\beta^r \in \left[\frac{1}{8}, \frac{1}{5}\right]$ for Ocean Type-I water as a distorting medium.

Approximately 98% of the world’s open oceans and coastal waters fall into
this category so values of $\beta^r$ close to this value represent a preferred
embodiment for underwater image restoration. With the red channel TM
calculated, the values for the green and blue channels can then be calculated
according to relationships between the channels, such as through the residual
energy ratios of the color channels, as described by Chiang et al in the context
of a wavelength compensation and image dehazing method. J. Y. Chiang
and Y.-C. Chen, "Underwater image enhancement by wavelength
Chiang et al. chose ratios manually, and predetermined ratios can also be selected in preferred embodiments. Another option is automatic calculation. A 2015 publication provides relations among the attenuation coefficients of different color channels based on inherent optical properties of water that are derived from the BL. X. Zhao, J. Tao, and Q. Song. "Deriving inherent optical properties from background color and underwater image enhancement," Ocean Eng., vol. 94, pp. 163-172, (Jan. 2015). The relations are determined as:

\[
\frac{B^k}{B^r} = \frac{B^r(m\lambda^k+i)}{B^k(m\lambda^r+i)}, \quad k \in \{g, b\} \quad (20)
\]

where \(\lambda^*, c \in \{r, g, b\}\) represent the wavelengths of the red, green, and blue channels, and \(m = -0.00113, \text{ and } i = 1.62517\). The TMs for the green and blue lights are then calculated by

\[
t^k = t^r(x)\frac{B^k}{B^r}, \quad k \in \{g, b\} \quad (21)
\]

where \(t^r\) is taken from Equation (19) in this preferred embodiment. Testing has also shown the preferred embodiments perform well compared to prior systems and techniques, as indicated by qualitative and quantitative measurement results. Quantitative and qualitative measurement results of the preferred embodiments are reported in Peng and Cosman, "Underwater Image Restoration based on Image Blurriness and Light Absorption," IEEE TRANSACTIONS XXX, VOL. X, NO. X, XXX (2016) and in Peng, Zhao and Cosman, "Single Underwater Image Enhancement using Depth Estimation based on Blurriness," in Proc. IEEE Int. Conf. on Imag. Process. (ICIP), pp. 4952-4956, (Sep. 2015). Experiments discussed above restored images according to Equation (8) in that paper. Specifically, \(P_{bltr}(x) = F_g\{C_r[P_r(x)]\}\). Then, stretch \(P_{bltr}(x)\) to a proper range \([r0, r1]\), where \(r0\) is set to \([0.3, 0.5]\) and \(r1\) is 0.9, empirically, denoted as \(t(x)\). Then, use \(f(x) = \frac{t(x) - \overline{B^c}}{t(x)} + \overline{B^c}\) to calculate the enhanced image \(J(x)\).
While specific embodiments of the present invention have been shown and described, it should be understood that other modifications, substitutions and alternatives are apparent to one of ordinary skill in the art. Such modifications, substitutions and alternatives can be made without departing from the spirit and scope of the invention, which should be determined from the appended claims.

Various features of the invention are set forth in the appended claims.
CLAIMS

1. A method for depth estimation in image or video obtained from a distorting medium comprising:
   calculating a pixel blurriness map;
   generating a rough depth map from the pixel blurriness map while assuming depth in a small local patch is uniform; and
   refining the rough depth map.

2. The method of claim 1, wherein said refining comprises closing by morphological reconstruction and guided filtering.

3. The method of claim 2, wherein said calculating calculates a difference between an original and multi-scale Gaussian-filtered images to estimate the pixel blurriness map.

4. The method of claims 1-3, wherein said calculating calculates the pixel blurriness map, denoted $P_{\text{init}}$, according to
   \[ P_{\text{init}}(x) = \frac{1}{n} \sum_{i=1}^{n} (|l_g(x) - G^{T_i}r_i(x)|) \]
   where $r_i = 2^i n + 1$ and $l_g$ is the grayscale version of the input image $I^c$, and $n$ is a constant.

5. The method of claim 4, wherein said generating the rough depth map, denoted $P_r$, is determined according to
   \[ P_r(x) = \max_{y \in \Omega(x)} P_{\text{init}}(y), \]
   where $\Omega(x)$ is a $z \times z$ local patch centered at $x$. 
6. The method of any of claims 1-5, wherein said refining comprises hole filling and image smoothing.

7. The method of claim 6, wherein said holes filling fills holes caused by flat regions in objects.

8. The method of claim 6, wherein said image smoothing comprises soft matting.

9. The method of claim 6, wherein said image smoothing comprises guided filtering.

10. The method of claim 6, wherein said hole filling and image smoothing generate a refined blurriness map, denoted $P_{blr}$, according to

$$P_{blr}(x) = F_g[C_r[P_r(x)]],$$

where $C_r$ is a hole-filling morphological reconstruction and $F_g$ comprises soft matting or guided filtering.

11. The method of any of claims 1-5, wherein said calculating and generating comprise a candidate selection process to select from a plurality of background light (BL) candidates.

12. The method of claim 11, wherein the candidate selection process includes background light (BL) candidates of a top predetermined percentage of blurry pixels in the input image, a variance region within a predetermined lowest variance range and a variance region within a predetermined largest blurriness range.
13. The method of claim 11, wherein the candidate selection process includes background light (BL) candidates from the top 0.1% blurry pixels in the input image, the lowest variance region and the largest blurriness region.

14. The method of claim 13, wherein lowest variance region and the largest blurriness region (which may or may not be the same) can be determined, for example, by quadtree decomposition.

15. The method of claim 11, wherein separate BL is determined for separate color channels after the candidate selection.

16. The method of claim 15, wherein the separate BL in accordance with:

---

Algorithm 1 BL-Estimate

1: Input parameter: input image $I^e$, blurriness map $P_{blr}$.
2: Output parameter: estimated BL $B_e$.
3:
4: function BL-ESTIMATE($I^e$, $P_{blr}$) returns $B_e$
5: $B_{cand}^c ← \text{avg}_{\text{BL}}(\text{QUAD-SELECT-LV}(I^e))$;
6: $B_{cand}^b ← \text{avg}_{\text{BL}}(\text{QUAD-SELECT-LB}(I^e, P_{blr}))$;
7: $B_{cand}^r ← \frac{1}{|P_{blr}|} \sum_{x \in P_{blr}} I^e(x)$;
8: $B_{max}^c ← \max_{x \in (1,2,3)} B_{cand}^c$;
9: $B_{min}^c ← \min_{x \in (1,2,3)} B_{cand}^c$;
10: for $k \in \{r, g, b\}$ do
11: \hspace{1cm} $\alpha ← \frac{S}{\text{SSE}(P^r, \epsilon_n)}$;
12: \hspace{1cm} $B_k ← \alpha B_{max}^k + (1 - \alpha)B_{min}^k$;
13: end for
14: return $B_e$;
15:
16:
17: function QUAD-SELECT-LB($I^e$, $P_{blr}$) returns $B_{cand}^c$
18: $I_{gray} ← \text{rgb2gray}(I^e)$;
19: $I_g ← I_{gray}$;
20: while $\text{SSE}(I_{gray}) > \epsilon$ do
21: Partition $I_{gray}$ into four quadrants, $I_1^g$, $I_2^g$, $I_3^g$, and $I_4^g$;
22: Pick $I_{max}^g$ with largest blurriness computed using $P_{blr}$;
23: $I_g ← I_{max}^g$;
24: end while
25: return $I^e(\text{Position}(I_g))$;
26: end function
---
17. The method of and of claims 1-5, wherein said generating comprise by applying a maximum filter to the pixel blurriness map.

18. The method of claim 1, wherein said generating is based upon both the pixel blurriness map and light absorption.

19. The method of claim 18, wherein said generating comprises a plurality of unique depth estimations and further comprises sigmoidally combining the different depth estimations in view of lighting and image conditions.

20. The method of claim 19, wherein a first depth estimate is obtained from a red color channel.

21. The method of claim 20, wherein a second depth estimate is obtained from a maximum intensity prior that compares the red channel to green and blue channels.

22. The method of claim 21, wherein a third depth estimate is obtained from the image blurriness.

23. An imaging system according to any of the preceding claims.

24. The imaging system of claim 23, comprising a camera in a sealed housing, memory for storing images and processor that controls camera functions and performs depth estimation on acquired images.

25. The method of any of claims 1-22, further comprising estimating a transmission map from the refined rough depth map.
Algorithm 1 BL-Estimate

1: Input parameter: input image $I^c$, bluriness map $P_{blr}$.
2: Output parameter: estimated BL $\overline{B}^c$.
3:
4: function BL-ESTIMATE($I^c$, $P_{blr}$)
5:   $B_{cand_1}^c \leftarrow \text{avg}_x \left[ \text{QUAD-SELECT-LV}(I^c) \right]$;
6:   $B_{cand_2}^c \leftarrow \text{avg}_x \left[ \text{QUAD-SELECT-LB}(I^c, P_{blr}) \right]$;
7:   $B_{cand_3}^c \leftarrow \frac{1}{|P_{blr}|} \sum_{x \in P_{blr}} I^c(x)$;
8:   $B_{max}^c \leftarrow \max_{i \in \{1,2,3\}} B_{cand_i}^c$;
9:   $B_{min}^c \leftarrow \min_{i \in \{1,2,3\}} B_{cand_i}^c$;
10:   for $k \in \{r,g,b\}$ do
11:       $\alpha \leftarrow S\left(\frac{|I^c|}{\text{size}(I^c)}, \epsilon_n\right)$;
12:       $\overline{B}^k \leftarrow \alpha B_{max}^k + (1 - \alpha) B_{min}^k$;
13:   end for
14: return $\overline{B}^c$;
15: end function
16:
17: function QUAD-SELECT-LB($I^c$, $P_{blr}$)
18:   $I_{gray} \leftarrow \text{rgb2gray}(I^c)$;
19:   $I_q \leftarrow I_{gray}$;
20:   while $\frac{\text{size}(I_q)}{\text{size}(I)} > \epsilon_s$ do
21:      Partition $I_q$ into four quadrants, $I_1^q$, $I_2^q$, $I_3^q$, and $I_4^q$;
22:      Pick $I_q^n$ with largest bluriness computed using $P_{blr}$;
23:      $I_q \leftarrow I_q^n$;
24:   end while
25: return $I^c(\text{Position}(I_q))$;
26: end function

FIG. 2
International application No.  
PCT/US 16/51897

A. CLASSIFICATION OF SUBJECT MATTER
IPC(8): G06K 9/40 (2016.01)
CPC: G06T 5/001
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC(8): G06K 9/40 (2016.01)
CPC: G06T 5/001

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic database consulted during the international search (name of database and, where practicable, search terms used) PatBase, ProQuest Dialog, Google Web, Google Patents (Search terms: depth estimate, distortion, medium, blur, map, depth, uniform patch, refine, difference, gaussian, mean absolute difference, maximum, light absorption, sigmoid, combine, red channel, compare, green, blue, intensity prior, image, blurriness, etc.)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

<table>
<thead>
<tr>
<th>Category</th>
<th>Citation of document, with indication, where appropriate, of the relevant passages</th>
<th>Relevant to claim No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>US 2007/0036427 A1 (Nakamura et al.) 15 February 2007 (15.02.2007), para. [0028], [0030], [0032]-[0033], [0035], [0037]-[0038], [0041]-[0043], and [0046], and Figs. 2 and 5.</td>
<td>1-3, 18-22</td>
</tr>
<tr>
<td>Y</td>
<td>US 2009/0010546 A1 (Rossato et al.) 08 January 2009 (08.01.2009), para. [0043], [0128], [0139]-[0134], [0138], and [0148]-[0151].</td>
<td>1-3, 18-22</td>
</tr>
</tbody>
</table>

Further categories of cited documents:
- Special categories of cited documents:
  - "A" document defining the general state of the art which is not considered to be of particular relevance
  - "E" earlier application or patent but published on or after the international filing date
  - "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
  - "O" document referring to an oral disclosure, use, exhibition or other means
  - "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search: 04 November 2016 (04. 11.2016)

Date of mailing of the international search report: 07 Dec 2016

Name and mailing address of the ISA/US: 
Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450
Facsimile No. 571-273-8300

Authorized officer: Lee W. Young
PCT Helpdesk: 571-272-4300
PCT OSP: 571-272-7774

Form PCT/ISA/210 (second sheet) (January 2015)
**INTERNATIONAL SEARCH REPORT**

**International application No.**

PCT/US 16/51897

<table>
<thead>
<tr>
<th>Box No. II</th>
<th>Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)</th>
</tr>
</thead>
<tbody>
<tr>
<td>This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:</td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td>[Z] Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:</td>
</tr>
<tr>
<td>2.</td>
<td>Claims Nos.: because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:</td>
</tr>
<tr>
<td>3.</td>
<td>☒ Claims Nos.: 4-17 and 23-25 because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Box No. III</th>
<th>Observations where unity of invention is lacking (Continuation of item 3 of first sheet)</th>
</tr>
</thead>
<tbody>
<tr>
<td>This International Searching Authority found multiple inventions in this international application, as follows:</td>
<td></td>
</tr>
</tbody>
</table>

| 1. | ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims. |
| 2. | ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees. |
| 3. | ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.: |
| 4. | ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.: |

**Remark on Protest**

- ☒ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet (2)) (January 2015)