DAYTIME FOG DETECTION AND DENSITY ESTIMATION WITH ENTROPY MINIMIZATION

Laurent Caraffa, Jean-Philippe Tarel
IFSTTAR, LEPSiS, 14-20 Boulevard Newton, Cité Descartes, Champs-sur-Marne, F-77420. Paris-Est University, France
laurent.caraffa@ifsttar.fr, jean-philippe.tarel@ifsttar.fr

ABSTRACT:

Fog disturbs the proper image processing in many outdoor observation tools. For instance, fog reduces the visibility of obstacles in vehicle driving applications. Usually, the estimation of the amount of fog in the scene image allows to greatly improve the image processing, and thus to better perform the observation task. One possibility is to restore the visibility of the contrasts in the image from the foggy scene image before applying the usual image processing. Several algorithms were proposed in the recent years for defogging. Before to apply the defogging, it is necessary to detect the presence of fog, not to emphasis the contrasts due to noise. Surprisingly, few a reduced number of image processing algorithms were proposed for fog detection and characterization. Most are dedicated to static cameras and can not be used when the camera is moving. Daytime fog is characterized by its extinction coefficient, which is equivalent to the visibility distance. A visibility-meter can be used for fog detection and characterization, but this kind of sensor performs an estimation in a relatively small volume of air, and is thus sensitive to heterogeneous fog, and air turbulence with moving cameras.

In this paper, we propose an original algorithm, based on entropy minimization, to detect fog and estimate its extinction coefficient by the processing of stereo pairs. This algorithm is fast, provides accurate results using low cost stereo camera sensor and, the most important, can work when the cameras are moving. The proposed algorithm is evaluated on synthetic and camera images with ground truth. Results show that the proposed method is accurate, and, combined with a fast stereo reconstruction algorithm, should provide a solution, close to real time, for fog detection and visibility estimation for moving sensors.

1. INTRODUCTION

Even if fog is rare, it impacts the quality of the results of many images processing algorithms when applied to outdoor scenes, see Fig. 1. For example, fog reduces vehicle and obstacle visibility when driving, which leads to death injuries. Solutions have been investigated to tackle the problem of reduced visibility when fog is around the vehicle. One possibility is to detect fog, turn on fog lights and depending of the importance of the fog which is measured by the extinction coefficient $\beta$, adapt the strength of the fog lights, as demonstrated in ICADAC project. An alternative is to display to the driver an image of the scene ahead of the vehicle after visibility restoration by image processing, see for instance (Tarel et al., 2012).

Fog detection and characterization is an important question to deal with these problems. Several algorithms were proposed for defogging, in particular in the recent years. A first class is based on color and contrast enhancement, for instance (Jobson et al., 1997, Zuiderveld, 1994). A second class is based on Koschmieder’s law to model the effect of fog, for instance (Tan, 2008, He et al., 2010, Tarel and Hautière, 2009, Tarel et al., 2010). With a linear response camera, Koschmieder’s law gives the apparent intensity $I$ of an object as a function of its intrinsic intensity $I_0$ and of the fog extinction coefficient $\beta$:

$$ I = I_0 e^{-\beta D} + I_s (1 - e^{-\beta D}) $$

where $D$ is the object depth, and $I_s$ is the intensity of the sky. The "meteorological visibility distance" is defined conventionally, as a function of $\beta$, by:

$$ d_{visibility} = -\frac{\ln(0.05)}{\beta} $$

In (Tarel et al., 2012), experiments shows that algorithms based on Koschmieder’s law generally do better than the ones not based on this law. If most of the time, the value of $\beta$ is not necessary to perform the defogging, one may notice that, when applied to road images, the value of $\beta$ is necessary to use the planar road assumption. The detection of the fog is also necessary, not to apply defogging methods when it is useless.

Surprisingly only few methods were proposed for detecting fog with a camera and for estimating the extinction coefficient $\beta$. Most of these methods apply only when the camera is static. The first approach is to estimate the decrease of contrast of distant objects, see for instance (Bush and Debes, 1998, Hautière et al., 2011). In (Kopf et al., 2008), the attenuation model due to fog is estimated assuming that a 3D model of the scene is known. This method is interesting since it shows that Koschmieder’s law is not always valid on long range distances due to heterogeneous fog. The difficulty with this method is that an accurate 3D model is required. When the camera is moving, only three approaches were proposed, the first one is based on the use of lane-markings, the second on the planar road assumption and the third on the use of the stereo. In (Pomerleau, 1997), $\beta$ is obtained by looking at the contrast attenuation of road markings at various distances. The disadvantage of this method is that it requires the presence of road markings in order to proceed. In (Lavenant et al., 2002, Hautière et al., 2006b), a method is proposed wich is dedicated to in-vehicle applications. This method assumes that the road is planar with a homogeneous gray level, and that the camera calibration is known with respect to the road plane. In case of fog, the
Figure 1: Contrast fading on the same scene due to various values of the extinction coefficient $\beta$.

gray level profile on the road, in the vertical direction, is derived from (1) and it shows an inflection point. The position of this inflexion point can be used to estimate the value of $\beta$. The weakness of this method is that several assumptions are required for the method to work. Another problem is that the inflexion point may be hidden by a close object in case of a light fog. In (Hautière et al., 2006a), a stereovision approach is proposed. It consists in detecting the most distant object with a contrast above 5%. This method is simple but it requires the presence of objects with high contrasts at every distances, and this is not always the case in practice.

From the drawbacks of the previous fog detection and fog characterization methods when the camera is moving, we see that a generic method is needed. We thus proposed a generic method which is based on stereovision. In (Caraffa and Tarel, 2013), a MRF model for stereo reconstruction and defogging is proposed. By minimizing this energy, $\beta$ can be estimated. However, this approach will lead to a slow iterative algorithm to produce good results. Inspired by (Caraffa and Tarel, 2013), we here investigate the estimation of the extinction coefficient $\beta$ when an approximate depth map is provided up to a critical distance. We thus assume that a fast stereo reconstruction is first applied and that this algorithm provides enough good matches up to the critical distance. When the depth is known, the main idea is to estimate $\beta$ from the attenuation of the contrast w.r.t the distance on homogeneous areas in the scene. Similarly to the case of shadows removal (Finlayson et al., 2009, Kwatra et al., 2012), the Shannon entropy is used as a criterion and the proposed method for estimating $\beta$ consists in the entropy minimization of the restored image. The advantage of the proposed approach over previous approaches is that it is generic and we do not need the whole range of distances in the depth map. The proposed algorithm, combined with a real time stereo reconstruction algorithm, should provide a fast algorithm for estimating the extinction coefficient (which includes the detection of foggy weather using a threshold on the visibility distance) when the camera is moving. As shown in the experiments, it also leads to accurate results when the camera is static.

First, in section 2, the problem of fog detection and extinction coefficient estimation is explained and our approach is presented. Then, the algorithm is detailed in section 3. In section 4, the proposed algorithm is evaluated on a synthetic database for a moving camera and evaluated with a ground truth database for static camera images.

2. PROBLEM STATEMENT

When a region with a uniform albedo is observed in presence of fog, its image is not of constant intensity. Indeed, its image becomes whiter with an increasing depth. This is due to the last additive term in (1), the so-called atmospheric veil which is an increasing function of the depth $D$. The Shannon entropy is a measure of the redundancy of a probability distribution. As a consequence, the entropy of the intensity distribution of a region of constant albedo, i.e without fog, is very small. On the contrary, the entropy of the same region in presence of fog is larger if the region spans a large range of depth. Our main idea is thus to estimate the extinction coefficient $\beta$ as the value which gives the restored image with the smallest entropy on regions of constant albedo.

The Fig. 2 illustrates the principle of the proposed method. A foggy image is shown in the Fig. 2(a). We search for regions of constant albedo. We thus propose to segment the foggy image in to regions where the intensity varies slowly, see Fig. 2(b). Some of the obtained regions, such as the fronto-parallel regions, or the regions with small depth variations are not adequate for the estimation of $\beta$ and thus must be rejected. Relevant segmented regions are thus selected using a criterion on the region area and on the region depth variations. In Fig. 2(c), this criterion leads to the selection of only three regions. In Fig. 2(d), 2(e) and 2(f), the image after restoration assuming different values of $\beta$ are shown, respectively with $\beta = 0.0024$, $\beta = 0.0086$ and $\beta = 0.0167$. These restorations are simply obtained by inverting (1), i.e by computing $I_0$, for each pixel, from the foggy image $I$, from the value of $I$, and from the depth map $D$. The correct value for $\beta$ is the second one, and was obtained using a visibility-meter. We can see that, when $\beta$ is too small, the restored image is still foggy and the gradient due to fog on homogeneous regions is perceptible. Conversely, when $\beta$ is too large, far away parts of the image become too dark. For the ground-truth value $\beta = 0.0086$, the restored image is more homogeneous and thus looks better than the others. In Fig. 2(g), 2(h) and 2(i), the gray level distributions on the 3 selected regions are displayed. We can see, for the correct value $\beta = 0.0086$, that the gray level histogram of each region is tighter than for the other values of $\beta$. Therefore, the entropy is smaller for this value compared to the two others. Fig. 2(j) shows the entropy of each selected regions versus $\beta$. It appears that a minimum of the entropy is reached around the ground-truth value of $\beta$ for the tree selected regions. Finally, in Fig. 2(k), the total entropy over the three selected regions is displayed versus $\beta$. The total entropy on the image is also shown. The previously used values of $\beta$ are displayed as vertical lines. It is clear from that figure, that the minimization of the sum of the entropy over the selected regions allows to better estimate the value of $\beta$ by doing an averaging on the three selected regions. The minimum of the sum of the entropy over the selected regions is very close to the ground truth value. On the contrary, the entropy over the image does not lead to a good estimate. Indeed, the entropy over the image intensities is mixing several random variables and can not be a coherent choice.

3. ALGORITHM

The used sensor is stereo cameras. For every stereo pair, the stereo reconstruction is performed by one or another method. The choice of the stereo reconstruction algorithm is studied in the next section experimentally, where it is shown that the stereo reconstruction only needs to be correct at a relatively short range of distances close to the cameras. The input of the proposed fog
Figure 2: Variation of the entropy with and without segmentation for different values of $\beta$. On the first line, 2(a) is the foggy image where the ground-truth is known: $\beta = 0.0086$, 2(b) is the used segmentation and 2(c) are the three selected regions. On the second line are the restored images for different values of $\beta$, when the depth-map $D$ is known. Next, the gray level histograms of the intensity distribution on the selected regions are shown for three different values of $\beta$. Then, 2(j) shows the entropy of the three selected regions versus $\beta$ and 2(k) shows the sum over the selected regions of the entropy compared to the image entropy. In the last figure, vertical lines displays the value of the three $\beta$ values used previously.
detection and characterization algorithm are thus an input camera image \(I(x)\) observed with or without atmospheric scattering and the corresponding depth map \(D(x)\) obtained from the stereo reconstruction.

When \(I(x)\) and \(D(x)\) are known, in the equation (1), the extinction coefficient \(\beta\), the restored image \(I_0\) and the intensity of the sky \(I_S\) remains unknown. The first step of the algorithm is thus to estimate \(I_S\) in an original way. Then, given a value of \(\beta\), the restored image \(I_0\) can be computed by solving (1). The second step consists in searching for the value of \(\beta\) which minimizes the total entropy of the gray level distributions over selected regions of the restored image.

### 3.1 Estimation of \(I_S\)

In defogging methods, the intensity of the sky is usually obtained using simple methods such as the selection of the maximum intensity in the image, or as the value corresponding to a given quantile on the intensity histogram, see for instance (Tan, 2008, Tarel and Hautière, 2009). If this kind of approach gives approximate results in practice, we observed that an accurate \(I_S\) is important to estimate correctly the value of \(\beta\). In particular, the fog being not necessarily homogeneous on long range distances, it is necessary to estimate a correct value of \(I_S\) in the close surrounding of the camera.

We thus introduce a first novel and simple method, based on contrast detection using the Sobel operator to estimate \(I_S\). First, we compute the Sobel operator on the input image. Second, pixels with a gradient magnitude larger than an arbitrary threshold are selected. Third, we choose \(I_S\) as the highest intensity value in the vicinity of the pixels selected by the gradient thresholding.

This previous can be applied with static or moving cameras. When the depth-map of the observed scene is known with enough accuracy, in particular when the scene is static. In such a case, the value of \(I_S\) can be estimated more accurately by taking the maximum intensity at the farthest distance detected by the pixels selected by the gradient thresholding. This refined method gives better results as shown in the experiments.

### 3.2 Region selection

The entropy minimization is meaningful only on a set of regions with constant albedo and with varying enough depth. The input image is thus segmented into regions where the intensity varies slowly with depth. The regions with constant intensity or fronto-parallel to the image are rejected as well as the tiny regions which are useless. The used criterion for this selection is to select regions with a depth range higher than 10% of the whole image depth range and with an area higher than 2% of the number of image pixels.

This segmentation can also be performed with advantages on a clear day image when a static sensor is used. The segmentation can be also performed on both the image and the depth map, when the depth map is accurate enough.

### 3.3 Estimation of \(\beta\) and \(I_0\)

Once \(I_S\) is estimated, \(I_0\) can be obtained for a given \(\beta\) by solving equation (1). To measure the quality of the restored image \(I_0\), we use the Shannon formulation of the entropy defined for a random variable \(y \in Y\):

\[
H(Y) = -E[\log p(y)] = -\sum_{y \in Y} p(y) \log p(y)
\]

where \(E\) is the expectation function and \(p(y)\) is the probability density function of variable \(y\). The entropy is defined for a single random variable, but for several independent variables, the entropy of the set of these variables is the sum of the individual entropy.

Actual scene contains of many objects, and thus its image is composed of several regions, and each region has its own random variable. Let \(S_c\) be the set of selected regions in the previously described steps. The total entropy is thus computed on this set of selected region as:

\[
TH = \sum_{Y_i \in S_c} H(Y_i) = -\sum_{Y_i \in S_c} \sum_{y \in Y} p(y) \log p(y) \quad (4)
\]

The total entropy is computed for a set of possible \(\beta\) values, and the estimate of \(\beta\) is obtained as the value minimizing the total entropy (4). This computation is quite fast, since only pixels in the selected regions with valid depths need to be considered and restored.

Due to wrong estimates in the depth map, due to heterogeneous fog, the intensity after restoration can be saturated and becomes black. This case appears for the ground-truth value of \(\beta = 0.0001\) in the Fig. 3(e) where remote objects are black on the right side of the image. This is due to the fact that the fog is lighter in this area. Not to be perturbed by this saturation, a maximal percentage of black pixels is allowed and \(\beta\) values with restorations with a larger percentage are discarded.

### 3.4 Summary

In short, the proposed algorithm is the following:

- Detection of the pixels with a gradient magnitude higher than a threshold on the input intensity image \(I(x)\). \(I_S\) is set as the highest intensity in the vicinity of these detected pixels.
- Segmentation of \(I(x)\) to regions and selection of the reliable regions \(S_c\) with thresholds on the depth range and on the area.
- For each \(\beta\), in a set of possible increasing values:
  - Compute the restored image for the pixels in the selected regions \(S_c\), when the depth map \(D(x)\) is known, using:
    \[
    I_0(x) = I(x)e^{\beta D(x)} - I_S(e^{\beta D(x)} - 1)
    \]
  - Compute the total entropy \(TH\) over the segmented regions of the restored intensity distributions \(I_0\) using (4).
  - Count the number of black pixels in the restored image \(I_0\), and if this number is higher than a threshold, the search on higher \(\beta\) values is stopped.
- The solution is given as the value of \(\beta\) which minimizes the total entropy (4).

In Fig. 3, the total entropy on the three selected regions is shown versus the value of \(\beta\). The ground-truth \(\beta\) values are shown as vertical lines. This figure shows how the ground-truth is close of the minimum of the total entropy, in various weather conditions.
3.5 Complexity

Computing the entropy of each intensity is linear according to the number of pixel. If $\beta$ has $K$ possible values and the image $N$ pixels, the complexity of the proposed algorithm is $O(KN)$. The proposed algorithm can be easily parallelized. Indeed, the restored image for each $\beta$ can be computed independently and the entropy can be computed with a MapReduce approach.

4. EVALUATIONS

To evaluate the proposed algorithm, we need foggy images with a visibility distance ground-truth. A visibility-meter can be used for the ground-truth. However and until now, we were not able to have moving stereo images of foggy scenes with a ground-truth. Indeed, the measures obtained by a visibility-meter are very perturbed by air turbulence around the sensor when it is on a moving vehicle. This is probably due to the fact that a visibility-meter uses a relatively small volume of air for its measure. As a consequence, we rely on synthetic images to experiment with moving scenes. For static scene, we rely on camera images.

4.1 Static sensor

The input of the proposed algorithm being an intensity image and a depth-map, we evaluated it using the Matilda database introduced in (Hautière et al., 2011). The Matilda database provides a sequence of more than 300 images from a static low cost camera during three days (night and day) with the visibility distances from a visibility-meter. An approximate depth map of the scene is also provided. In this database, the visibility varies from 100m to 15km, and the accuracy of the used visibility-meter is around 10%. The maximum distance in the depth map is 1300m.

To challenge the proposed algorithm, we also performed tests where the image is processed in a reduced range of depths. Thus, the evaluation results are presented in two cases: in the first case, the whole depth map is used (denoted $D_{\text{max}}(I_{S})$), and in the second case, only the pixels with a depth lower than 50m are used (denoted $D_{\text{max}} = 50m$). The camera being static, the segmentation into regions is performed a single time on an image with clear weather conditions using the algorithm proposed in (Felzenszwalb and Huttenlocher, 2004). We thus use the same selected regions for all images in the sequence. This allows to evaluate the visibility distance estimation without the extra source of errors due to the segmentation process. The effect of applying the segmentation on every image is also tested.

Tab. 1 shows the evaluation results in term of mean relative error on the visibility distance and the ground-truth on the Matilda database for different distance ranges, with a single image segmentation. The first line is obtained from the maximum distance $D_{\text{max}}$ where $I_{S}$ is observed and the second line is for a shorter and fixed distance $D_{\text{max}} = 50$m. On the last line, the segmentation is computed on every image, to simulate the situation of embedded sensors in a moving vehicle.

![Figure 3: Normalized total entropy versus $\beta$ on the same scene with different weather conditions. The ground-truth $\beta$ measured with a visibility-meter are shown (0.0001, 0.0046, 0.0086 and 0.0244) as vertical lines.](image)

![Graph](image)

<table>
<thead>
<tr>
<th>Visibility (m)</th>
<th>0-250</th>
<th>0-500</th>
<th>0-1000</th>
<th>all</th>
</tr>
</thead>
<tbody>
<tr>
<td>number of images</td>
<td>9</td>
<td>19</td>
<td>24</td>
<td>198</td>
</tr>
<tr>
<td>$D_{\text{max}}(I_{S})$</td>
<td>10.0%</td>
<td>9.1%</td>
<td>11.7%</td>
<td>97.3%</td>
</tr>
<tr>
<td>$D_{\text{max}} = 50$m</td>
<td>7.1%</td>
<td>7.1%</td>
<td>10.0%</td>
<td>91.5%</td>
</tr>
<tr>
<td>Images segmentation</td>
<td>$D_{\text{max}}(I_{S})$</td>
<td>25.4%</td>
<td>17.2%</td>
<td>20.7%</td>
</tr>
</tbody>
</table>

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
doi:10.5194/isprsannals-II-3-25-2014
Figure 4: Comparison between estimated visibility distance and ground-truth during 1600 minutes (one image every 10 minutes, images during the night are skipped) on the Matilda database. The proposed method is tested with two depth ranges: maximum distance where $I$ is obtained (green crosses) and a fixed range $[0,50]$m (red crosses), using a single image segmentation and is also shown the case with a different segmentation at every image (blue crosses). The ground-truth value is shown as a black line. The horizontal blue line shows the 500m visibility distance which is considered as the distance were fog is no more considered in driving situations.

Table 2: Mean relative error between the estimated visibility distance and the ground-truth value on 66 synthetic stereo images from Frida3 database. In the top part, the ground-truth depth-map is used and in the bottom part, the depth-map comes from the stereo reconstruction using Libelas algorithm.

<table>
<thead>
<tr>
<th></th>
<th>Mean relative error</th>
</tr>
</thead>
<tbody>
<tr>
<td>With ground truth distance</td>
<td>4.07%</td>
</tr>
<tr>
<td>With stereo reconstruction depth map</td>
<td>16.03%</td>
</tr>
</tbody>
</table>

4.2 Moving sensor

In the case of a moving sensor, we evaluate the proposed method on 66 synthetic stereoscopic images with uniform fog from the database named Frida3. This database was introduced in (Tarel et al., 2012) for stereo reconstruction and restoration evaluation and contains a ground-truth: the images without fog and the associated depth-maps.

To study the effect, on the estimated $\beta$, of bad matches in the stereo reconstruction, we compared the case where the ground-truth depth-map is used with the case where the depth-map comes from a fast stereo reconstruction algorithm. In practice, we use the Libelas algorithm which is very fast and is dedicated to road scene images (Geiger et al., 2010). Tab. 2 shows the obtained results in terms of mean relative error on the estimated visibility distance in the two cases. Results obtained using the ground-truth depth-map are very accurate, then the stereo reconstruction is used, accuracy is lower but still of good quality.

5. CONCLUSION

In this article, we propose an original method for estimating the extinction coefficient of the fog or equivalently the meteorological visibility distance. The proposed method is based on the entropy minimization of the restored image on pixels where the depth is known. The depth-map can come from a 3D model of the scene, or from a stereo reconstruction as well. The results obtained on a static scene are as accurate that what is usually obtained with a visibility-meter, when the depth-map is of good quality. The proposed algorithm can be also used with a moving sensor and is tested with success on synthetic stereo images.

The proposed algorithm has several advantages. First, a small depth range can be used to characterize a visibility distance 10 times larger than the maximum depth in the depth-map. This is important for vehicular applications where obstacles or other vehicles can hide a large range of distances. Another important consequence is that the implemented stereo reconstruction algorithm needs to provide good matches only up to a critical distance which is much lower than the visibility distance. Second, the proposed method is fast enough to consider real-time applications for vehicular applications. In particular, it can be easily speeded-up using parallel computations.

REFERENCES


Lavenant, J., Tarel, J.-P. and Aubert, D., 2002. Procédé de détermination de la distance de visibilité et procédé de détermination de la présence d’un brouillard. French patent number 0201822, INRETS/LCPC.

