



HAL
open science

Stereo Reconstruction and Contrast Restoration in Daytime Fog

Laurent Caraffa, Jean Philippe Tarel

► **To cite this version:**

Laurent Caraffa, Jean Philippe Tarel. Stereo Reconstruction and Contrast Restoration in Daytime Fog. Asian Conference on Computer Vision, Nov 2012, France. pp 13-25. hal-00852835

HAL Id: hal-00852835

<https://hal.science/hal-00852835>

Submitted on 21 Aug 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Stereo Reconstruction and Contrast Restoration in Daytime Fog

Laurent Caraffa and Jean-Philippe Tarel*

Université Paris-Est, LEPSiS, IFSTTAR,
58 Boulevard Lefèbvre, F-75015 Paris, France

Abstract. Stereo reconstruction serves many outdoor applications, and thus sometimes faces foggy weather. The quality of the reconstruction by state of the art algorithms is then degraded as contrast is reduced with the distance because of scattering. However, as shown by defogging algorithms from a single image, fog provides an extra depth cue in the gray level of far away objects. Our idea is thus to take advantage of both stereo and atmospheric veil depth cues to achieve better stereo reconstructions in foggy weather. To our knowledge, this subject has never been investigated earlier by the computer vision community. We thus propose a Markov Random Field model of the stereo reconstruction and defogging problem which can be optimized iteratively using the α -expansion algorithm. Outputs are a dense disparity map and an image where contrast is restored. The proposed model is evaluated on synthetic images. This evaluation shows that the proposed method achieves very good results on both stereo reconstruction and defogging compared to standard stereo reconstruction and single image defogging.

1 Introduction

The first dense stereo reconstruction algorithms were proposed forty years ago. There is now more than one hundred algorithms listed on the Middlebury evaluation site. Nevertheless, several new algorithms or improvements are proposed each year. The reason for this constant interest is the high usefulness of the 3D reconstruction which serves in many applications such as: driver assistance, automatic driving, environment simulators, augmented reality, data compression, 3D TV. While the Middlebury database contains only indoor scenes of good quality, outdoor applications are confronted with more difficult weather conditions such as fog, rain and snow. These weather conditions reduce the quality of the stereo pairs and introduce artifacts. Reconstruction results are thus usually degraded.

The principle of stereo reconstruction is to find, for every pixel in the left image, the pixel in the right image which minimizes a matching cost along the epipolar line. Depending on the scene, the matching cost can be ambiguous or

* Thanks to the ANR (French National Research Agency) for funding, within the ICADAC project (6866C0210).

wrongly minimal. A prior on the disparity map is thus added, for instance to enforce that close pixels have similar disparity. As a consequence, the stereo reconstruction is set as the minimization of an energy which derives from a Markov Random Field (MRF) model, see for instance [1, 2]. Thanks to recent advances in numerical analysis, the optimization of this energy can be performed quickly without being trapped by most of the local minima.

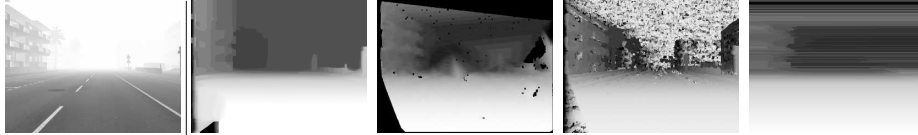


Fig. 1. From left to right: original left image of the stereo pair, disparity maps obtained using α -expansion on MRF [1], Libelas [3], correlation windows and dynamic programming on each line.

We observed that stereo reconstructions are degraded in the presence of fog. As an illustration, in Fig. 1, we show disparity maps obtained on a foggy stereo image by four stereo reconstruction algorithms: α -expansion on MRF [1], Libelas [3], correlation windows and dynamic programming on each line. Results are not satisfactory; in the best case, they are correct only up to a critical distance. Indeed, in a foggy scene, the more distant an object, the whiter its color. As a consequence, contrast is a decreasing function of distance, which makes matching all the more difficult to perform. If stereo disparity is important for 3D reconstruction, in foggy scenes, the gray-level of distant objects is also a depth cue. This depth cue is used in contrast restoration algorithms but had not been used in 3D reconstruction yet. The defogging problem can also be set as a MRF problem, see [4]. The atmospheric veil depth cue is particularly interesting since it is complementary to the stereo depth cue: the former is reliable only for remote objects, while the latter is reliable only for near by objects. Our idea is thus to combine a MRF model of both stereo reconstruction and defogging problems into a unified MRF model to take advantage of both depth cues. As far as we know, there is no algorithm dedicated to dense stereo reconstruction in foggy weather conditions.

The article is structured as follows. In Sec. 2, we state the problem, and explain how fog affects the scene image. The classic dense stereo reconstruction and image defogging problems are derived from a general formulation. In Sec. 3, our model of the stereo reconstruction and defogging problem is proposed. At last, Sec. 4 is dedicated to an evaluation on synthetic images and tests on camera images.

2 Problem Statement

The inputs are the left and right images of a stereo pair $\{I_L, I_R\}$. These images are observed after perturbation by atmospheric scattering and camera optics. The images without all these perturbations are denoted I_{0L} and I_{0R} , respectively, and are of course unknown. Also unknown is the depth map represented by its disparity map D . Our goal being to fuse depth cues from the stereo and from the atmospheric veil to achieve better reconstruction, it seems natural to search for a Bayesian formulation of the problem so that prior knowledge can be included to remove possible ambiguities. The two unknowns that we want to estimate are the disparity map D and the clean left image I_{0L} . The right one I_{0R} is not an unknown since, not considering occluded objects, it is a function of D and I_{0L} .

The maximum a posteriori principle tells us to maximize the following posterior probability, which can be rewritten using Bayes' rule as:

$$p(D, I_{0L}|I_L, I_R) \propto p(I_L, I_R|D, I_{0L}) P(D, I_{0L}) \quad (1)$$

where $p(I_L, I_R|D, I_{0L})$ is the data likelihood and $P(D, I_{0L})$ is the prior on the unknowns (D, I_{0L}) . Instead of posterior probability maximization, in practice, it is its log which is minimized, leading to the following formulation in terms of energy, or log-likelihood:

$$E(D, I_{0L}|I_L, I_R) = \underbrace{E(I_L, I_R|D, I_{0L})}_{E_{data}} + \underbrace{E(D, I_{0L})}_{E_{prior}} \quad (2)$$

The term E_{data} is also known as the data cost or fidelity term, and E_{prior} as the prior or regularization term.

2.1 Dense Stereo Reconstruction Without Fog

Without fog, I_L and I_R are only affected by the noise of the sensor which is generally low. Following [2], the Bayesian formulation of the dense stereo reconstruction is approximated by assuming that I_L is without noise. In (2), the unknown variable I_{0L} can be thus substituted by I_L leading to the approximate but simpler energy minimization:

$$E(D|I_L, I_R) = \underbrace{E(I_R|D, I_L)}_{E_{data_stereo}} + \underbrace{E(D|I_L)}_{E_{prior_stereo}} \quad (3)$$

Data term: E_{data_stereo} is the error in intensity between a pixel in the left image and a pixel in the right image given a disparity. It is usually chosen as:

$$E_{data_stereo} = \sum_{(i,j) \in X} \rho_S\left(\frac{|I_L(i,j) - I_R(i - D(i,j), j)|}{\sigma_S}\right) \quad (4)$$

where X is the set of image pixels, ρ_S is a function related to the distribution of the intensity noise with scale σ_S . This intensity noise takes into account the camera noise, but also the occlusion, and it can be one of the functions used in robust estimation to remove outliers.

Prior term: This term enforces the smoothness of the disparity map. Because of constant intensity objects, the data term can be rather ambiguous. It is thus necessary to introduce a prior on the disparity map to interpolate the ambiguous areas correctly. The smoothness prior tells that two close pixels have a greater chance to be the projection of a same object with the same depth than remote pixels. This assumption is not always true due to gaps in depth for example. As a consequence, a robust function ρ_D should be used in this term. The classical prior term is:

$$E_{prior_stereo} = \lambda_D \sum_{(i,j) \in X} \sum_{(k,l) \in N} W_D(\nabla I_L(i,j)) \rho_D(|D(i,j) - D(i+k, j+l)|) \quad (5)$$

where λ_D is a factor weighting the strength of the prior on D , N is the set of relative positions of pixel neighbors (4, 8 connectivity or other), and W_D is a monotonically decreasing function of image intensity gradients. The weight W_D is introduced to smooth low-gradient ambiguous areas more than gradient edges. Usually W_D is chosen as a decreasing exponential function of the image gradient: $W_D(\nabla I) = e^{-\frac{|\nabla I|}{\sigma_g}}$, where σ_g is a scale parameter. It is even better to use a function of the image Laplacian in order to avoid sensitivity to linear intensity variations: $W_D(\nabla I) = e^{-\frac{|\Delta I|}{\sigma_g}}$.

2.2 Effects of Fog

With a linear response camera, assuming an object of intrinsic intensity I_0 , the apparent intensity I in presence of a fog with extinction coefficient β is modeled by Koschmieder law:

$$I = I_0 e^{-\beta p} + \underbrace{I_s(1 - e^{-\beta p})}_V \quad (6)$$

where p is the object depth, and I_s is the intensity of the sky. From (6), it can be seen that fog has two effects: first an exponential decay $e^{-\beta p}$ of the intrinsic luminance I_0 , and second the addition of the atmospheric veil V which is an increasing function of the object distance p . The depth p can be rewritten as a function of the disparity $p = \frac{\delta}{D}$ where δ is related to the stereo calibration parameters. It is important, for the following, to notice that there is one situation where D can be obtained from a single image using V : when I_0 is close to zero, i.e when the object is dark. It is also important to notice that when the disparity D is zero, the intensity I_0 cannot be obtained. Moreover, I_0 being positive, the photometric constraint $V < I$ is deduced from (6).

For road images, several algorithms exist for detecting the fog and estimating the extinction coefficient β , see for instance [5]. The parameter β is thus assumed known in the following, as well as I_s .

2.3 Single Image Defogging Knowing the Depth

Before we describe our model for fused stereo reconstruction and defogging, we focus on the simpler problem of defogging from a single image I given the

disparity map D . Using the previous notations, only the left image is used in this section. We thus drop L in the indexes. The unknown I_0 is the image without fog and noise. Both I and D are assumed known. Even though D or the depth p is not accurately known, an approximate map is enough. The defogging problem knowing the disparity D can be set as a particular case of (1), i.e the maximization of the posterior probability:

$$p(I_0|D, I) \propto p(I|D, I_0)P(I_0|D) P(D) \quad (7)$$

or equivalently as the minimization of the energy:

$$E(I_0|D, I) = \underbrace{E(I|D, I_0)}_{E_{data-fog}} + \underbrace{E(I_0|D)}_{E_{prior-fog}} \quad (8)$$

Data term: The data term is the log-likelihood of the noise probability on the intensity, taking into account that I_0 is observed through the fog, see (6):

$$E_{data-fog} = \sum_{(i,j) \in X} \rho_P \left(\frac{|I_0(i,j)e^{-\frac{\beta\delta}{D(i,j)}} + I_s(1 - e^{-\frac{\beta\delta}{D(i,j)}}) - I(i,j)|}{\sigma_P} \right) \quad (9)$$

where ρ_P is a function related to the intensity noise due to the camera and σ_P is the scale of this noise. ρ_P and σ_P are thus directly related to the probability density function (pdf) of the camera noise and can be estimated off-line when calibrating the camera. It can be noticed for D close to zero that the data term does not constrain the distribution of I_0 which tends to the uniform pdf.

Prior term: We found that the following prior term produces nice restoration results:

$$E_{prior-fog} = \lambda_{I_0} \sum_{(i,j) \in X} \sum_{(k,l) \in N} e^{-\frac{\beta\delta}{D(i,j)}} W_{I_0}(\nabla D(i,j)) \rho_{I_0}(|I_0(i,j) - I_0(i+k, j+l)|) \quad (10)$$

where λ_{I_0} is a factor weighting the strength of the prior on I_0 . Function W_{I_0} is the equivalent to W_D in the stereo, only now it is applied on the disparity map gradient rather than on image gradient. We use $W_{I_0}(\nabla D) = e^{-\frac{|\Delta D|}{\sigma'_g}}$, where σ'_g is a scale parameter. Function ρ_{I_0} is a robust function used for similar reasons as ρ_D . An extra weight $e^{-\frac{\beta\delta}{D(i,j)}}$ is introduced, and it is a key point, to take into account that in presence of fog, there is an exponential decay of contrast with respect to (w.r.t.) depth. This has the effect of giving less and less importance to the prior as depth increases. This is necessary to be consistent with the fact that the distribution of I_0 is less and less constrained by the data term for large distances. Without this extra factor, the intensity of close objects may wrongly diffuse on remote objects.

2.4 Optimization

While MRF formulations are successful to model image processing and computer vision problems, it is also necessary to have reliable optimization algorithms to minimize the derived energies. A large class of useful MRF energies is of the form:

$$f(Y) = \sum_{x \in X} \Phi_x(Y_x) + \sum_{x \in X, x' \in X} \Phi_{x,x'}(Y_x, Y_{x'}) \quad (11)$$

When the variable Y is binary, it has been shown long ago that for sub-modular functions Φ , the global minimum of the previous problem can be obtained in polynomial time. For non-binary variables, one of the most efficient technique to optimize (11) approximately is the α -expansion algorithm, which is based on the decomposition of the problem in successive binary problems. The global optimum of each binary sub-problem is obtained in polynomial time, when the prior term is sub-modular.

When the function Φ is not sub-modular, other heuristics such as $\alpha - \beta$ swap, Belief propagation, TRW, roof duality were proposed which produce interesting results.

3 Stereo Reconstruction and Defogging

The model we now propose for fused stereo reconstruction and defogging shares similarities with the single image defogging model presented in [4]. Indeed in [4], the model is set as a MRF model and both depth p and restored image I_{0L} are estimated successively. The main difference is that stereo is used in our approach, while the approach in [4] is monocular. In particular, this last approach cannot work with gray-level images, contrary to our stereo approach. Another difference is that, in [4], Koschmieder's law (6) is rewritten, after algebraic manipulations and use of the log function, in such a way that the depth and intensity appear as a linear combination of independent functions of each of these two variables. This rewriting allows a simpler optimization. However, the noise is non linearly transformed and this is not taken into account. The stereo approach we now present contains non-linear equations where the image noise is better handled.

3.1 MRF Model

Data term: In stereo with fog, the data term (9) applies on the left image. On the right, a similar term taking into account the disparity D is also introduced. This leads to the following log-likelihood of the stereo data in fog:

$$\begin{aligned} E_{data_fog_stereo} = & \sum_{(i,j) \in X} \rho_P \left(\frac{|I_{0L}(i,j)e^{\frac{-\beta\delta}{D(i,j)}} + I_s(1 - e^{\frac{-\beta\delta}{D(i,j)}}) - I_L(i,j)|}{\sigma_P} \right) \\ & + \rho_P \left(\frac{|I_{0L}(i,j)e^{\frac{-\beta\delta}{D(i,j)}} + I_s(1 - e^{\frac{-\beta\delta}{D(i,j)}}) - I_R(i - D(i,j), j)|}{\sigma_P} \right) \end{aligned} \quad (12)$$

Notice that when $\beta = 0$, i.e without fog, the first term in (12) enforces $I_{0L} = I_L$, and the second term is the stereo log-likelihood E_{data_stereo} . This shows that D can be estimated from both log-likelihoods. We thus propose to linearly combine the two log-likelihoods in the data term:

$$E_{data^*} = \alpha E_{data_stereo} + (1 - \alpha) E_{data_fog_stereo} \quad (13)$$

with $0 \leq \alpha \leq 1$. During the estimation of both I_{0L} and D , the value of I_{0L} can be temporarily far from the true value. The advantage of introducing E_{data_stereo} in the data term is that the minimization of E_{data_stereo} provides correct estimates of D at short distances even if I_{0L} is badly estimated.

Photometric constraint and assumption on white pixels: As introduced in Sec. 2.2, the photometric constraint on the atmospheric veil V must be verified both on the left and right images of the stereo pair. Due to noise, the photometric constraint is not very strict but it helps to reduce the search space of I_{0L} .

Due to fog, the contrast of remote objects is very low and stereo does not work. As remote objects are nearly white, we add a zero disparity assumption on those pixels with an intensity equal to I_s . This assumption is of course wrong for white objects. Taking into account the photometric constraint and the assumption on white pixels, the data term is:

$$E_{data} = \begin{cases} E_{data^*} & \text{if } V(i, j) \leq I_L(i, j) + 3\sigma_P \\ & \text{and } V(i, j) \leq I_R(i - D(i, j), j) + 3\sigma_P \\ & \text{and } I_L(i, j) \neq I_s \\ 0 & \text{if } I_L(i, j) = I_s \text{ and } D(i, j) = 0 \\ +\infty & \text{else.} \end{cases} \quad (14)$$

Prior term: In (1), the prior probability $P(D, I_{0L})$ is related to two variables: the disparity D and the intensity I_{0L} . Unfortunately, this kind of mixed prior term is actually difficult to optimize. To be consistent with previous stereo and defogging prior terms, (5) and (10) respectively, the two variables D and I_{0L} cannot be assumed independent of one another. We thus propose to write the prior probability as $P(D, I_{0L}) = P(D|\ddot{I}_{0L})P(I_{0L}|\ddot{D})$, where \ddot{D} and \ddot{I}_{0L} are fixed approximations of D and I_{0L} , given as priors. We thus propose the following prior term for the stereo reconstruction and defogging problem:

$$E_{prior} = \sum_{(i,j) \in X} \sum_{(k,l) \in N} \lambda_{I_0} e^{-\frac{\beta \delta}{\ddot{D}(i,j)}} W_{I_0}(\nabla \ddot{D}(i, j)) \rho_{I_0}(|I_{0L}(i, j) - I_{0L}(i + k, j + l)|) \\ + \lambda_D W_D(\nabla \ddot{I}_{0L}(i, j)) \rho_D(|D(i, j) - D(i + k, j + l)|) \quad (15)$$

The fact that \ddot{D} and \ddot{I}_{0L} are approximated is not a problem since they appear only in the weights, such as W_{I_0} and W_D , which are very smooth functions. Indeed, the weight W_D is set, like in the stereo reconstruction case, to $W_D(\nabla I) = e^{-\frac{|\Delta I|}{\sigma_g}}$. The weight W_{I_0} is set, like in the defogging case, to $W_{I_0}(\nabla D) = e^{-\frac{|\Delta D|}{\sigma_g}}$.

Initial \ddot{D} and \ddot{I}_{0L} : Variables \ddot{D} and \ddot{I}_{0L} are the approximate disparity and approximate intensity in E_{prior} . The atmospheric veil can be approximately estimated on the left image using a single image defogging algorithm, see for instance [6, 7]. Here, it is approximated by minimizing the following w.r.t. \ddot{V} :

$$\sum_{(i,j) \in X} |I_L(i,j) - \ddot{V}(i,j)| + \lambda \sum_{(k,l) \in N} |\ddot{V}(i,j) - \ddot{V}(i+k, j+l)| \quad (16)$$

using α -expansion. The small features in the image I_L are lost in \ddot{V} , but thanks to the L_1 robust terms, large objects with low contrast are kept. This atmospheric veil \ddot{V} has the important property: it contains object edges. The weights W_D and $W_{I_{0L}}$ in E_{prior} are introduced to attenuate the regularization through these edges. By definition from (6), $V = 1 - e^{-\frac{\beta\delta}{D}}$ (assuming $I_s = 1$ without loss of generality). As a consequence, \ddot{D} can be obtained from \ddot{V} . This implies that the factor $e^{-\frac{\beta\delta}{D}}$ in E_{prior} can be substituted by $1 - \ddot{V}$. Another consequence is that $\Delta\ddot{D}$ in W_{I_0} can be approximated by $\Delta\ddot{V}$. Rather than search for a close approximation of \ddot{I}_{0L} , we use $\frac{\Delta\ddot{V}}{1-\ddot{V}}$ as a good approximation of $\Delta\ddot{I}_{0L}$.

Complete model: In summary, the stereo reconstruction and defogging problem is set as the following minimization:

$$\min_{D, I_{0L}} E_{data} + E_{prior} \quad (17)$$

In practice, the functions ρ_D and ρ_{I_0} are chosen as the identity. The noise on the image being assumed Gaussian, ρ_P is the square function. For those pixels which verify the photometric constraint and which are not white, the energy which is minimized is, after introduction of \ddot{V} :

$$\begin{aligned} E(D, I_{0L}) = & \sum_{(i,j) \in X} \left\{ \frac{1-\alpha}{\sigma_P^2} \left(|I_{0L}(i,j)e^{\frac{-\beta\delta}{D(i,j)}} + I_s(1 - e^{\frac{-\beta\delta}{D(i,j)}}) - I_L(i,j)|^2 \right. \right. \\ & + |I_{0L}(i,j)e^{\frac{-\beta\delta}{D(i,j)}} + I_s(1 - e^{\frac{-\beta\delta}{D(i,j)}}) - I_R(i - D(i,j), j)|^2 \Big) \\ & + \alpha \rho_S \left(\frac{|I_L(i,j) - I_R(i - D(i,j), j)|}{\sigma_S} \right) \\ & + \sum_{(k,l) \in N} \left\{ (1-\alpha)\lambda_{I_0}(1 - \ddot{V}(i,j))e^{-\frac{|\Delta\ddot{V}(i,j)|}{\sigma_g}} |I_{0L}(i,j) - I_{0L}(i+k, j+l)| \right. \\ & \left. \left. + \lambda_D e^{-\frac{|\Delta\ddot{V}(i,j)|}{\sigma_g(1-\ddot{V}(i,j))}} |D(i,j) - D(i+k, j+l)| \right\} \right\} \quad (18) \end{aligned}$$

As this energy is known up to a scale factor, (18) can be arbitrarily divided by $(1-\alpha)\lambda_{I_0}$. This is used in the next section to estimate σ_P from image residuals.

3.2 Optimization

In (18), D and I_{0L} appear in non-linear unary functions and independently in binary functions. It is thus possible to optimize (18) by means of a two-step alternate minimization: one step consists in minimizing w.r.t. I_{0L} and the other in minimizing w.r.t. D . The first step is defogging and the second step is stereo reconstruction. The energies in both steps being sub-modular, α -expansion is used for the minimization. With the alternate minimization, convergence towards a local minima is guaranteed. Before the first step, the disparity is initialized by stereo reconstruction assuming no fog, i.e by minimizing (18) with $\alpha = 1$.

As pointed in [4], the gradient distribution of a hazy image can be very different from that of a foggy image. This implies that after division of (18) by $(1 - \alpha)\lambda_{I_0}$, the factor $\sigma_P\sqrt{\lambda_{I_0}}$ must be set differently from one image to another. When this factor is not correctly set, the chance to converge towards an interesting local minimum decreases. Hopefully, the first term of (18) being quadratic, the factor $\sigma_P\sqrt{\lambda_{I_0}}$ can be easily estimated by estimating the standard deviation of the left intensity residuals $I_{0L}(i, j)e^{\frac{-\beta\delta}{D(i, j)}} + I_s(1 - e^{\frac{-\beta\delta}{D(i, j)}}) - I_L(i, j)$.

In summary, the optimization scheme is:

- Compute \ddot{V} by minimization of (16), using α -expansion.
- Initialize D by minimizing (18) w.r.t D , with $\alpha = 1$, using α -expansion.
- Until convergence, iterate:
 1. Until convergence, iterate:
 - (a) Minimization of (18) w.r.t I_{0L} , using α -expansion.
 - (b) Minimization of (18) w.r.t D , using α -expansion.
 2. Update σ_P by computing the standard deviation of the left intensity residuals.
- $\sigma_P\sqrt{\lambda_{I_0}}$ is enforced to value 1 and a last optimization w.r.t. I_{0L} is performed to better emphasize the detailed texture.

4 Evaluation

4.1 Parameters Setting

The proposed MRF model is mainly parametrized by α which is the weight between the photometric log-likelihood $E_{photo_fog_stereo}$ of left and right images and the log-likelihood E_{photo_stereo} of the stereo. When α is close to zero, the obtained disparity map is smooth in homogeneous areas, but the disparity of close objects may be biased as well as the intensity I_{0L} . When α is close to one, the disparity obtained from the stereo log-likelihood is usually correct for close objects but the quality of the reconstruction decreases with the contrast and thus with the depth. Therefore, we recommend to set α close to 0.5 or a little higher.

Another important parameter is the initial value of $\sigma_P\sqrt{\lambda_{I_0}}$. The choice of this value can have an effect on the local minima selected at convergence.

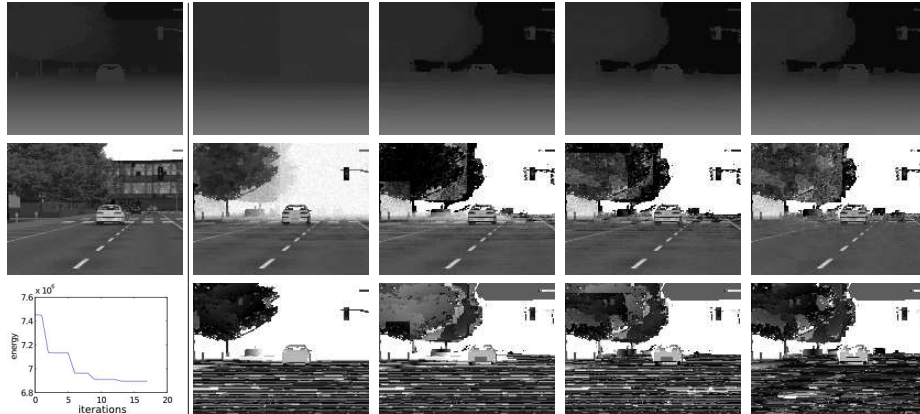


Fig. 2. First column, from top to down: the ground truth disparity map, the image without fog and the energy decrease with iterations. Second column: the disparity map, the restored image with $\sigma_P \sqrt{\lambda_{I_0}} = 1$ and the disparity error map obtained with stereo reconstruction without considering fog. Third column: results of the proposed method when $\sigma_P \sqrt{\lambda_{I_0}} = 20$ at the initialization. Fourth column: the first iteration. Last column: after convergence. For comparison purpose, restorations are processed using the last optimization step to emphasize details.

The bigger $\sigma_P \sqrt{\lambda_{I_0}}$ at the initialization, the smoother is the depth map after convergence. Fig. 2 shows several iterations of the algorithms with $\sigma_P \sqrt{\lambda_{I_0}} = 20$. We can notice that, after one iteration, the large scale of $\sigma_P \sqrt{\lambda_{I_0}}$ allows a better reconstruction and restoration around the closest vehicle. When the number of iteration increases, the scale $\sigma_P \sqrt{\lambda_{I_0}}$ becomes smaller, and the restoration is improved step by step for remote objects. Thus, the two far away vehicles appear. A too large scale can cause wrong stereo matching. However, when α is larger than 0.5, these wrong matches are unusual.

4.2 Synthetic Images

To evaluate the stereo reconstruction in foggy weather, we rely on synthetic images due to the difficulty to have the 3D model of a scene and images of this scene with and without fog. We generate synthetic stereo images using SiVICTM software which allows to build physically-based road environments. Uniform fog is added knowing the depth map, see Fig. 3. To make the image more realistic and evaluate the ability of the algorithm to manage the noise, we also added a Gaussian noise on every pixels of left and right images, with standard deviation 1. This database is named FRIDA3 and is available online for comparative studies¹.

We compared the results of three methods: first, the stereo reconstruction based on the classic MRF model without fog; second, the first iteration of the

¹ <http://perso.lpc.fr/tarel.jean-philippe/visibility/fogstereo.zip>

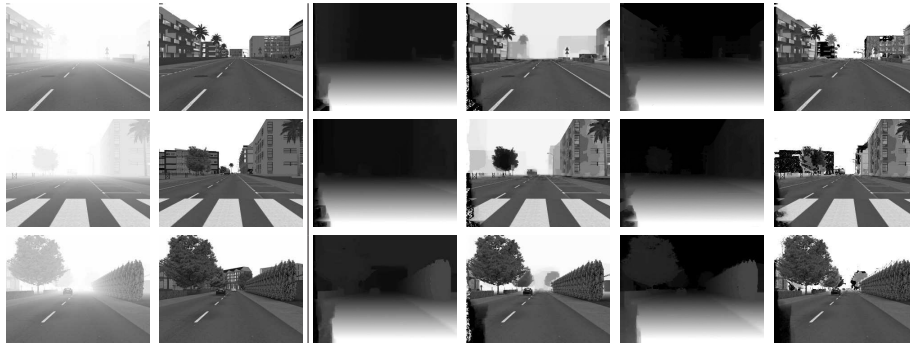


Fig. 3. Results on three images of the synthetic FRIDA3 stereo image database. First column: foggy left images. Second column: same scene without fog. Third and fourth columns: disparity maps obtained using stereo reconstruction without fog and restored images using these disparity maps. Fifth and sixth columns: disparity maps with the proposed method and associated restored images.

Table 1. Comparison of the percentage of correct disparities in average on 66 synthetic stereo pairs using the classic MRF approach without fog (STEREO, see Sec. 2.1), with the photometric constraint and assumption on white pixels added to stereo (STEREO+PC), at the first iteration (FIRST) and after converging (FINAL). Percentages are given for different values of the maximum error err on the disparity (in pixel).

Algorithm	$err < 1$	$err < 0.66$	$err < 0.33$
STEREO	0.776	0.722	0.514
STEREO+PC	0.811	0.764	0.548
FIRST	0.822	0.771	0.552
FINAL	0.828	0.780	0.573

proposed method; third, the proposed method after convergence (with initial $\sigma_P \sqrt{\lambda I_0} = 20$ and $\alpha = 0.5$). Results are shown in Tab. 4.2, in average on 66 stereo pairs. This percentage takes into account only the pixels seen in both images with disparity larger than one, i.e not considering the sky. The stereo without fog (STEREO) achieves 72.2% of correct disparities in the whole image, for a maximum error of 0.66 pixels. When the photometric constraint due to fog veil is added (STEREO+PC), the percentage of correct disparities is improved to 76.4%. This step STEREO+PC corresponds to the initialization of the proposed method. The first iteration of the proposed method (FIRST) achieves 77.1%. After convergence (FINAL), this percentage is increased to 78.0%. From Tab. 4.2, it is clear that the proposed method outperforms the classic stereo reconstruction which does not take the presence of fog into account. In percentage, the improvement due to iterations may seem reduced on the whole image, but these iterations are important to improve correct disparities at long distances.

This fact is illustrated in Fig. 3 which displays obtained disparity maps and restored left images on three stereo pairs of the FRIDA3 database.

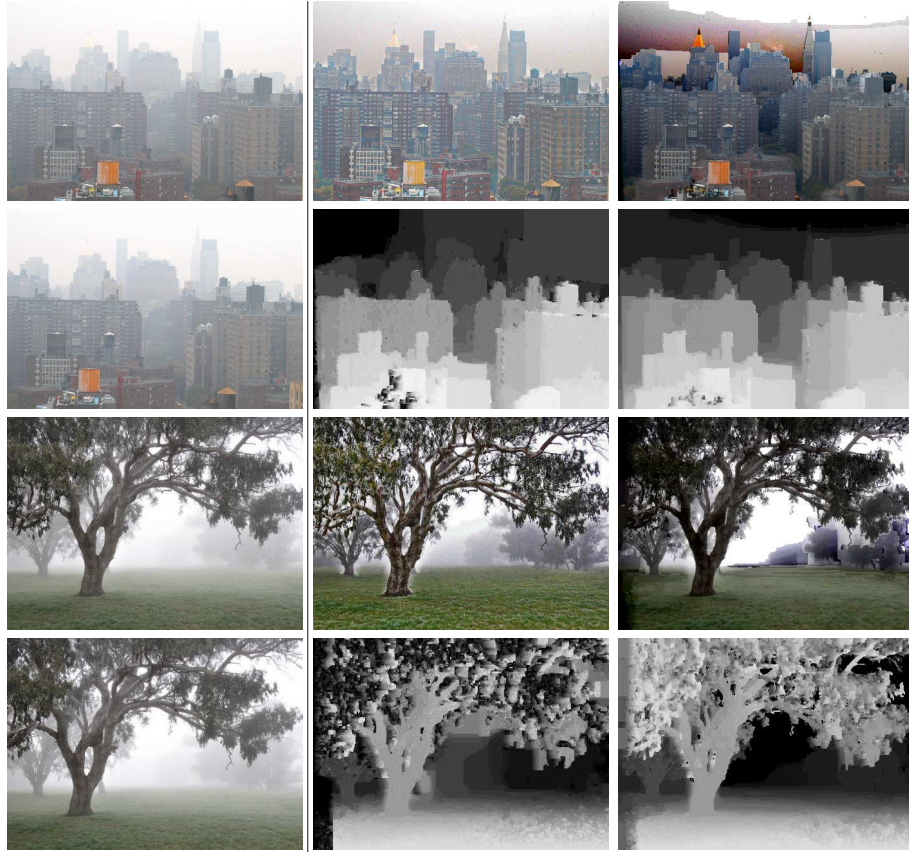


Fig. 4. First column: foggy stereo pair. Second column: Single image defogging with [6] and disparity map obtained by stereo reconstruction without fog. Third column: restored left image and disparity map obtained using the proposed method.

4.3 Camera Images

We compared the proposed method to the stereo reconstruction without fog described in Sec. 2.1 and image defogging described in [6]. β is manually selected. Results show that both the reconstruction and restoration are of better quality. In Fig. 4, results are compared on urban and country side stereo pairs. One may note that the obtained stereo reconstruction are dense at both short and long distances, contrary to stereo reconstruction without taking into account

the fog. The stereo restoration obtained by the proposed method is of good quality compared to single image defogging results. At close distances, outliers are avoided thanks to the photometric constraint and the true intensity of objects is kept. At a far distances, the contrast is greatly enhanced without amplifying the noise to much.

5 Conclusion

We proposed a MRF model to solve the stereo reconstruction and image defogging in daytime fog. It is an extension of two sub-models: the classical stereo reconstruction without fog and newly introduced image restoration when the depth is known. The proposed model includes the photometric constraint and priors on white pixels. It leads to the optimization of an energy which can be solved by an alternate scheme based on the application of successive α -expansion optimizations. The convergence towards a local minimum is thus guaranteed. Tests on both synthetic stereo pairs and camera stereo pairs show the relevance of the model. Thanks to the stereo depth clue, the disparity is correct at short distances, and thanks to the atmospheric veil depth cue, the disparity is drastically improved at long distances. The obtained restored results are better than the ones obtained without stereo thanks to the simultaneous estimation with the disparity map. Perspectives for future research are to take into account non constant sky, non Gaussian noise to improve scale estimation, to explicitly take into account occlusions in the formulation, to speed up the algorithm for real time applications and to extend the previous model to heterogeneous fog.

References

1. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23** (2001) 1222–1239
2. Woodford, O., Torr, P., Reid, I., Fitzgibbon, A.: Global stereo reconstruction under second-order smoothness priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31** (2009) 2115–2128
3. Geiger, A., Roser, M., Urtasun, R.: Efficient large-scale stereo matching. In: *Proceedings of the 10th Asian conference on Computer vision. Volume Part I of ACCV'10.* (2011) 25–38
4. Nishino, K., Kratz, L., Lombardi, S.: Bayesian defogging. *International Journal of Computer Vision* **98** (2012) 263–278
5. Hautière, N., Tarel, J.P., Lavenant, J., Aubert, D.: Automatic fog detection and estimation of visibility distance through use of an onboard camera. *Machine Vision and Applications* **17** (2006) 8–20
6. Tarel, J.P., Hautière, N.: Fast visibility restoration from a single color or gray level image. In: *Proceedings of IEEE International Conference on Computer Vision (ICCV'09), Kyoto, Japan* (2009) 2201–2208
7. Tarel, J.P., Hautière, N., Caraffa, L., Cord, A., Halmaoui, H., Gruyer, D.: Vision enhancement in homogeneous and heterogeneous fog. *IEEE Intelligent Transportation Systems Magazine* **4** (2012) 6–20