A Probabilistic Intensity Similarity Measure based on Noise Distributions

Yasuyuki Matsushita Stephen Lin
Microsoft Research Asia
Beijing, P.R. China, 100080
Email: {yasumat,stevelin}@microsoft.com

Abstract

We derive a probabilistic similarity measure between two observed image intensities that is based on the noise properties of the camera. In many vision algorithms, the effect of camera noise is either neglected or reduced in a preprocessing stage. However, noise reduction cannot be performed with high accuracy due to lack of knowledge about the true intensity signal. Our similarity metric specifically represents the likelihood that two intensity observations correspond to the same unknown noise-free scene radiance. By directly accounting for noise in the evaluation of similarity, the proposed measure makes noise reduction unnecessary and enhances many vision algorithms that involve matching of image intensities. Real-world experiments demonstrate the effectiveness of the proposed similarity measure in comparison to the standard L^2 norm.

1. Introduction

Imaging noise produces random perturbations of intensity values that degrade the visual quality of images and reduce the reliability of computer vision algorithms. There has been much work on modeling and removing these variations; however, even when the noise characteristics of an image are known, it remains a difficult problem to remove noise in a manner that preserves the actual scene information in the measured intensity signal. Since there in general does not exist sufficient information in an image to extract the original scene data, noise reduction cannot be performed with high accuracy, and often leads to loss of image detail or introduction of image artifacts.

In this work, we present an alternative approach for dealing with the effects of imaging noise on vision algorithms. Rather than attempt to remove the noise from an image, we aim to directly account for its effects on intensity similarity measurement, which is one of the most fundamental and important operations in computer vision. Similarity between intensity observations is used to guide a broad range of algorithms including segmentation, object recognition, and optical flow. Conventionally, the similarity of two observed intensities I_1 and I_2 is measured as a function of their L^2 norm $||I_1 - I_2||_2$. While such a distance metric provides a valid comparison in theory, the quality of this measure lessens in practice due to the presence of noise in real images.

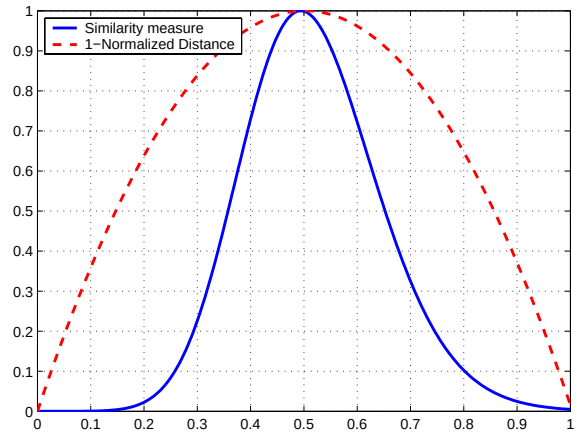


Figure 1. Plot of the proposed similarity measure and the L^2 distance, shown after normalization and with respect to intensity level 0.5. The displayed similarity measure is computed based on intensity-dependent Gaussian noise whose variance is proportional to the intensity level. The resulting similarity measure has an asymmetric profile.

Noise is introduced into intensity observations at multiple points in the imaging pipeline. Prominent components of image noise include the random noise associated with A/D conversion and uneven photon flow from the scene, fixed pattern noise due to differences in sensitivity among photon detectors in the imaging array, and dark current noise that results from measurement of thermal radiation within the camera. A detailed presentation of the various noise sources can be found in [6]. Since noise arises from a series of stochastic processes, it is natural to treat intensity observations as samples from a probability distribution defined by these imaging fluctuations. Such noise distributions may vary among pixels in the image, and can also be dependent on intensity level.

Our proposed method captures the intensity dependent

noise distribution of a camera in a calibration step, and then based on these probability distributions, we derive in this paper a probabilistic measure of the similarity between two observed intensities. This metric specifically represents the likelihood that two intensity observations correspond to the same noise-free scene radiance, with respect to the probabilistic noise distributions.

The effect of considering image noise in similarity measurement is illustrated by the example in Figure 1. While metrics based on the L^2 norm do not account for the particular imaging fluctuations within a camera, our proposed intensity similarity measure provides statistically meaningful values with respect to noise characteristics. The likelihood that two observations have the same noise-free value depends upon these noise characteristics, such that the similarity is high when the two intensities are both well within the noise distributions of certain true intensities, and becomes significantly lower otherwise. This is in contrast to conventional metrics which have a fixed structure that is intended to measure *distance* between observations rather than their *similarity*. These distance measures do not represent the likelihood that two intensity observations correspond to the same scene radiance value, since they do not consider how intensities fluctuate in the imaging system. We contend that it is the similarity of intensities that is generally important for many image understanding tasks, and that the use of a similarity-based measure is especially beneficial in real applications where intensity perturbations exist.

The presented intensity similarity measure can be directly extended to handle color values, and can also be employed in subpixel and multiresolution applications. With this proposed metric, we experimentally demonstrate improved performance on the common vision process of block matching.

1.1. Prior work

How to determine the similarity of a pair of $\mathbb{R}^1$ intensity observations is a problem that has received little attention. In fact, it has become a convention in computer vision to compute the dissimilarity between two intensity observations as a function of their Euclidean distance. While these metrics are valid measures of distance between two intensity observations, they do not represent their similarity in the sense of how likely the two observations correspond to the same scene radiance value. A likelihood measure should be a function of intensity fluctuations in the imaging system, which are not accounted for in these fixed distance metrics.

Omer and Werman [8] proposed an approach to model color distributions in an image as lines. These distributions are collected from regions that have homogeneous color. The proposed color model enables comparison of color values that is invariant to shading, highlights and color distortions caused by the imaging process. Comaniciu and Meer [3] also took an image-specific clustering approach that is based upon the mean shift algorithm. An obtained affinity measure leads to good segmentation results. A learning-based approach to modeling pixel affinity was presented by Fowlkes *et al.* [4], in which human-segmented images are used as a training set. A similarity measure called the bottleneck affinity was presented by Omer and Werman [10]. The similarity measure is defined as a likelihood that two features belong to the same cluster, based on the histogram profiles along the straight-line path between the two points. Similar affinity measures based on geodesic distances [12] or geodesic paths [9] have also been proposed. These previous affinity-based methods all seek to determine how likely two observations belong to the same *class*, based on modeling of feature distributions for each class. In contrast, our work aims to measure how likely two observations correspond to the same *scene radiance value*, based on models of intensity distributions due to noise.

Closer to our work, Alter *et al.* [2] established an intensity similarity measure for low-light conditions, which uses Maximum Likelihood (ML) estimation to define the similarity between two intensity observations. While it more accurately measures similarity than other basic metrics, it relies heavily on accurately estimating the *true* intensity signal, which is difficult to achieve from only two observations. In our work, we derive a new intensity similarity measure that avoids having to estimate the true intensity signal. Instead, our similarity measure is based on the marginal density of observations with respect to every possible true intensity signal.

2. Probabilistic framework

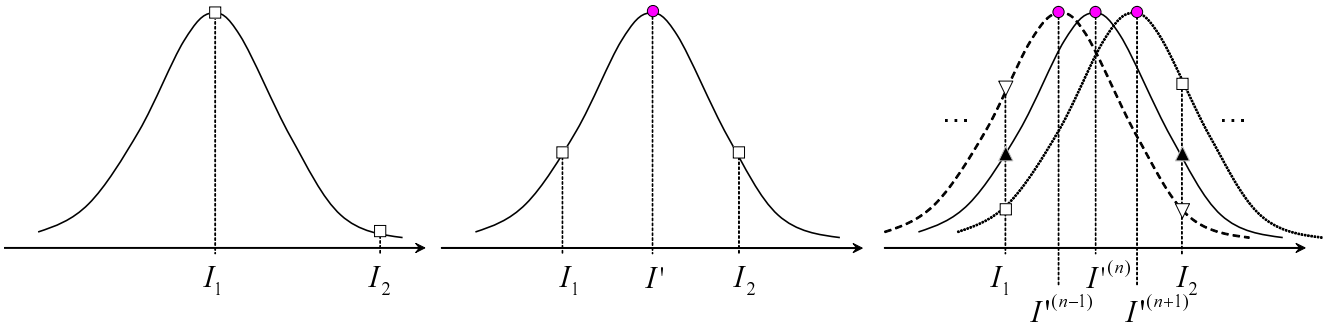
The measured intensity of a pixel can be considered as a random variable that takes a value I in the space of observations Ω. Given an observed I, the probability that it resulted from a true intensity I' of scene radiance is given by the conditional probability density $p(I|I')$.

Alter *et al.* [2] presents a similarity measure based on ML estimation. From two observations, it first estimates the true intensity I^*, then derives similarity as the product of conditional probabilities:

$$I^* = \operatorname*{argmax}_{I'} p(I_1|I')p(I_2|I'), \qquad (1)$$
$$S_{ML}(I_1, I_2) = p(I_1|I^*)p(I_2|I^*).$$

The ML similarity S_{ML} is defined as the likelihood of two intensity observations resulting from the *single* true intensity I^* that gives the greatest likelihood.

The ML similarity gives a sound measure in theory if the estimated true intensity is correct. However, the true intensity generally cannot be estimated accurately from two observations in practice, especially under noisy imaging conditions. Our approach differs from ML similarity in that it



(a) (b) (c)

Figure 2. Comparison of similarity measures in the case of Gaussian noise. (a) Gaussian similarity measure, which assumes that one of the two observations (I_1 or I_2) is free of noise. (b) ML similarity measure, which tries to estimate the true intensity signal I' by ML estimation and then derives the similarity as the product of conditional probabilities $p(I_1|I')p(I_2|I')$. (c) Our probabilistic similarity measure, which sums up the probability of observing I_1 and I_2 from every possible true intensity I'. $I'^{(n)}$ represents the n-th possible true intensity value.

does not attempt to estimate the true intensity. We define the similarity of two observed intensities I_1 and I_2 as the likelihood that they each resulted from the *same* but *unknown* true intensity I', i.e., $p(I_1|I')p(I_2|I')$. Since the true intensity is unknown, *all* possible true intensities are taken into account, rather than using a specific true intensity as done in the ML similarity measure. Our probabilistic similarity measure is therefore formulated as

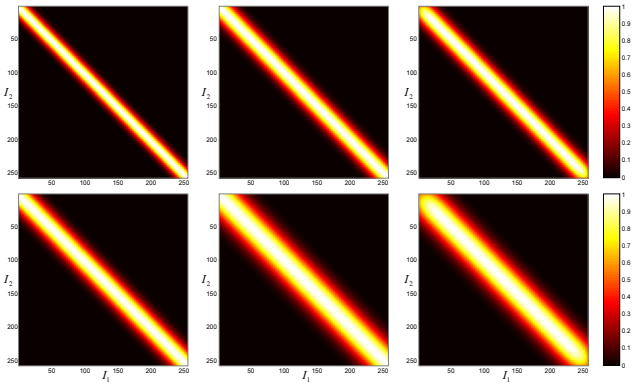
$$\begin{aligned} S(I_1, I_2) &\stackrel{\text{def}}{=} p\big(t(I_1) = t(I_2)\big) \\ &= \int_\Omega p(I_1|I')p(I_2|I')p(I')dI', \end{aligned} \tag{2}$$

where $t(I)$ is the true intensity of observation I, and $p(I')$ is the prior density of the true intensity, which is uniform. With this definition in the form of marginal densities, the proposed similarity measure is able to naturally avoid the uncertainty that exists in the estimation of a single true intensity.

The proposed intensity similarity measure is formulated as a marginal density of the likelihood function over the parameter I' as shown in Equation (2). This formulation is equivalent to the *marginal likelihood* in Bayesian estimation [13], in which there is not a single parameter value (in our case of I') that characterizes the entire model, but rather an entire set of possible likelihood functions is used. In the Bayesian framework, the marginal likelihood is computed using the likelihood functions with the associated prior distributions of the parameters. Our similarity measure also uses all possible likelihood functions $p(I_1|I')p(I_2|I')$ instead of relying upon a specific parameter I^* as done in ML similarity.

2.1. Case studies with different noise models

With different models of noise, we examine the proposed similarity measure and explain its differences from prior similarity metrics.



(a) (b) (c)

Figure 3. Visualization of different intensity similarity measures in the case of intensity-independent Gaussian noise (top row : $\sigma^2 = 100$, bottom row : $\sigma^2 = 250$). (a) Gaussian similarity, (b) ML similarity, and (c) probabilistic similarity. I_1 and I_2 are two observations, and brighter values indicate greater similarity of the two observations. Similarity values are mapped to range [0,1] for visualization purposes.

Intensity-independent Gaussian noise

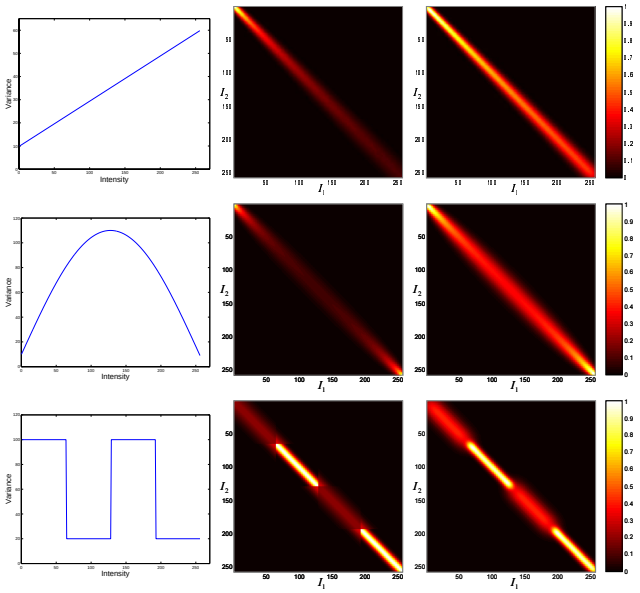
Suppose that image noise follows a Gaussian distribution, and that this Gaussian noise is the same at all intensities. The most common approach for determining the similarity of two observations (I_1 and I_2) in this case is to assume that either I_1 or I_2 is the true intensity, and then evaluate a Gaussian function as follows:

$$S_G(I_1, I_2) = p(I_2|I_1) = p(I_1|I_2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(I_1-I_2)^2}{2\sigma^2}}. \tag{3}$$

A visualization of this metric for various values of I_1 and I_2 is shown in Figure 2 (a).

With the ML similarity measure defined in Equation (1), similarity in this case can be expressed as

$$\begin{aligned} S_{ML}(I_1, I_2) &= p(I_1|I^*)p(I_2|I^*) \\ &= \frac{1}{2\pi\sigma^2} e^{-\frac{(I_1-I^*)^2+(I_2-I^*)^2}{2\sigma^2}}, \end{aligned} \tag{4}$$



(a) (b) (c)

Figure 4. Visualization of different intensity similarity measures in the case of intensity-dependent Gaussian noise. (a) Noise level function, (b) ML similarity, and (c) probabilistic similarity. From top row to bottom row, the noise variances σ^2 are respectively a linear, sinusoidal, and step function with respect to intensity. I_1 and I_2 are two observations, and brighter values indicate greater similarity of the two observations. Similarity values are mapped to range [0,1] for visualization purposes.

where $I^* = \text{argmax}_{I'}\, p(I_1|I')p(I_2|I')$. Figure 2 (b) displays the form of this function. Since the true intensity I^* is obtained through ML estimation, rather than by assuming one of the observations to be noise-free, it should provide a better similarity estimate than the basic Gaussian distance function. However, it still relies heavily on an accurate estimate of the true intensity value.

Our approach differs from these previous two in that it does not attempt to estimate the true intensity. Rather, it sums up probabilities that these two observations result from the same, but unknown, true intensity:

$$S(I_1, I_2) = \int_\Omega p(I_1|I')p(I_2|I')p(I')dI' \qquad (5)$$
$$= \frac{1}{4\sigma|\Omega|\sqrt{\pi}} e^{\frac{-(I_1-I_2)^2}{4\sigma^2}} \left\{ \text{erf}\left(\tfrac{2I_m-(I_1+I_2)}{2\sigma}\right) + \text{erf}\left(\tfrac{I_1+I_2}{2\sigma}\right) \right\},$$

where $I_m \overset{\text{def}}{=} \max I'$, $|\Omega| \overset{\text{def}}{=} p(I') = \text{const}$, and $\text{erf}(\cdot)$ is the error function [1] defined as

$$\text{erf}(x) \overset{\text{def}}{=} \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

Derivation of Equation (5) is detailed in Appendix. Figure 2 (c) illustrates the form of this measure. Since this approach takes into account all possible true intensity signals, it theoretically gives a more accurate similarity estimate than previous approaches.

Intensity-dependent Gaussian noise

As described in [7], noise level or variance is often a function of intensity in real images. In this case, it is unclear how to define the Gaussian similarity measure of Equation (3), since $p(I_1|I_2) \neq p(I_2|I_1)$. Variance σ now becomes $\sigma_{I'}$, which is a function of I'. For ML similarity and our proposed measure, the similarity metrics become

$$S_{ML}(I_1, I_2) = p(I_1|I^*)p(I_2|I^*) \qquad (6)$$
$$= \frac{1}{2\pi\sigma_{I^*}^2} e^{-\frac{(I_1-I^*)^2+(I_2-I^*)^2}{2\sigma_{I^*}^2}},$$

where $I^* = \text{argmax}_{I'}\, p(I_1|I')p(I_2|I')$, and

$$S(I_1, I_2) = \int_\Omega p(I_1|I')p(I_2|I')p(I')dI' \qquad (7)$$
$$= \frac{1}{2\pi|\Omega|} \int_\Omega \frac{1}{\sigma_{I'}^2} e^{-\frac{(I_1-I')^2+(I_2-I')^2}{2\sigma_{I'}^2}} dI'.$$

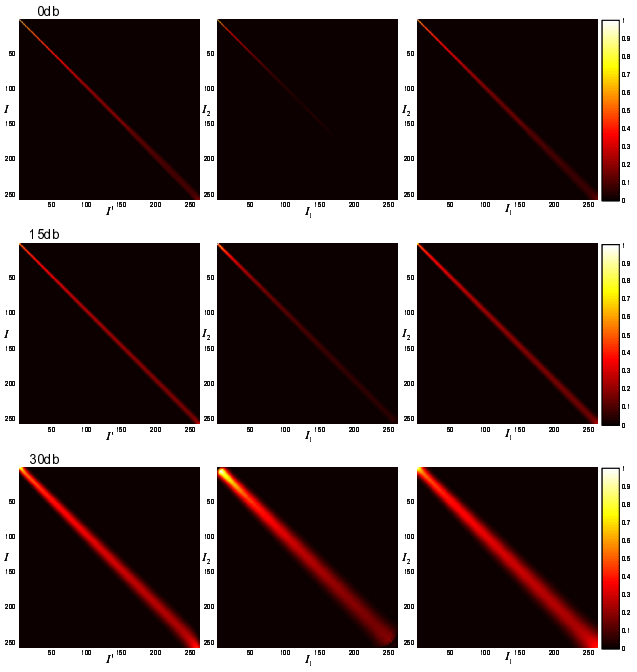
Visualizations of these similarity measures built upon intensity-dependent Gaussian noise are shown in Figure 4. We note that unlike the other similarity measures, our probabilistic metric may take an asymmetric form that arises from the intensity dependency of noise, as also illustrated in Figure 1.

Nonparametric noise model

When the noise cannot be modelled analytically, conditional pdfs can be expressed more generally in the form of discrete histograms. Histograms representing noise distributions can be obtained through real-world measurements, such as by capturing multiple shots of a static scene from a fixed viewpoint. In this setup, a histogram is constructed for each pixel from all of its recorded intensity observations. To relate a histogram to a conditional pdf $p(\cdot|I')$, it is necessary to estimate the true intensity signal I' from the histogram. Various criteria may be used for this purpose. For zero-mean noise, the mean value of the distribution can be used to determine the true intensity signal. For an assumption that the true intensity signal should correspond to most common value in the distribution, the intensity value at the mode of the histogram is taken as the true intensity signal.

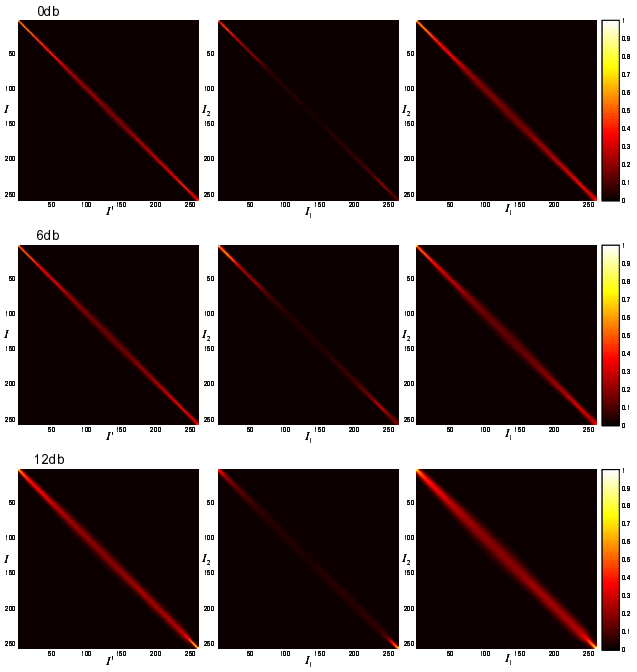
Once the true intensity signal I' of a histogram is established, the histogram is normalized such that it sums to one. In this way, the histogram provides a representation of the joint probability density $p(I, I')$. The conditional pdfs $p(I|I')$ are directly computed from the joint pdfs $p(I, I')$ as

$$p(I|I') = \frac{p(I, I')}{\int_\Omega p(I, I')dI}. \qquad (8)$$



(a) (b) (c)

Figure 5. Nonparametric case with different gain parameters. (a) conditional pdf $p(I|I')$, (b) ML similarity, and (c) probabilistic similarity. From top to bottom, the results of gain parameters 0 db, 15 db, and 30 db are shown. The data is obtained from a real-world scene using the green color channel of a PointGrey Research Dragonfly camera. For visualization purposes, similarity values are mapped to range $[0, 1]$ and are shown over the true intensity range of $[1, 254]$.



(a) (b) (c)

Figure 6. Nonparametric case with different gain parameters. (a) conditional pdf $p(I|I')$, (b) ML similarity, and (c) probabilistic similarity. From top to bottom, the results of gain parameters 0 db, 6 db, and 12 db are shown. The data is obtained from a real-world scene using the green color channel of a Sony DSR-PD190P camera. For visualization purposes, similarity values are mapped to range $[0, 1]$ and are shown over the true intensity range of $[1, 254]$.

Using the obtained conditional pdfs and Equation (2), the nonparametric probabilistic similarity measure is obtained.

Figures 5 and 6 show the intensity similarity measures derived from observed pdfs at different camera gain levels, which can affect noise characteristics. As seen in the figures, actual conditional pdfs have complex structures that are hard to model with simple parametric functions, but can be easily handled in our probabilistic similarity measure.

2.2. Color similarity measure

This similarity measure for pixel intensity can be directly extended to handle pixel features represented as vector quantities, such as color or local texture attributes. Using RGB color as an example, the similarity between vectors $\mathbf{v}_1 = [r_1, g_1, b_1]$ and $\mathbf{v}_2 = [r_2, g_2, b_2]$ can be expressed as

$$S(\mathbf{v}_1, \mathbf{v}_2) = \int_{\Omega_\mathbf{v}} p(\mathbf{v}_1|\mathbf{v}')p(\mathbf{v}_2|\mathbf{v}')d\mathbf{v}' \tag{9}$$

$$= \int_{\Omega_b}\int_{\Omega_g}\int_{\Omega_r} p(r_1|r')p(g_1|g')p(b_1|b') \; p(r_2|r')p(g_2|g')p(b_2|b')p(r')p(g')p(b')dr'dg'db'$$

$$= S(r_1, r_2) \cdot S(g_1, g_2) \cdot S(b_1, b_2).$$

2.3. Subpixel & multi-resolution similarity measure

There are many computer vision algorithms that utilize subpixel or multi-resolution matching, e.g., image alignment and optical flow. The proposed intensity similarity measure can be extended to handle subpixel/multi-resolution cases.

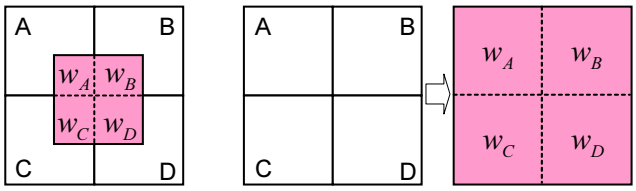
As shown in Figure 7, a subpixel value and a lower-resolution pixel value I_s can both be expressed using combination weights w_i with the original pixels:

$$I_s = \sum_i w_i I_i. \tag{10}$$

Accordingly, the conditional pdf $p_s(I_s|I')$ can be derived by linear interpolation of the original pdfs:

$$p_s(I_s, \mathbf{\Theta}|I') = \sum_i w_i p(I_i|I'), \tag{11}$$

where parameter $\mathbf{\Theta}$ indicates the combination of I_i and w_i that is used for interpolation. Inserting the conditional pdfs p_s into Equation (2), we obtain the probabilistic similarity measure for the subpixel/multiresolution case.



(a) Subpixel case (b) Multi-resolution case

Figure 7. Subpixel and multi-resolution cases. Similarity is recomputed based upon combination weights w, which are determined by the new pixel's coverage of the original pixels.

3. Application and Experiments

The proposed similarity metric provides the likelihood that two observations correspond to the same scene radiance. To evaluate the effectiveness of the proposed similarity measure, we compare it to the conventional L^2 distance in the application of basic block matching [5]. Block matching is chosen as our example since this fundamental image matching algorithm is employed in a wide range of applications.

Application formulation Block matching finds the most similar image block by minimizing the sum of squared differences to a given block:

$$\mathbf{u} = \underset{\mathbf{u}}{\operatorname{argmin}} \sum_{\mathbf{x}\in\mathbf{B}} \big(F_1(\mathbf{x}) - F_0(\mathbf{x}-\mathbf{u})\big)^2, \qquad (12)$$

where $\mathbf{B}$ is the given block, $F_i(\cdot)$ is a function at the i-th frame that maps the pixel coordinate to its intensity, $\mathbf{x}$ is a 2D vector in the image space, and $\mathbf{u}$ is a 2D motion vector to be estimated. By minimizing the distance between blocks in consecutive frames, a motion vector u is obtained. With the proposed similarity measure, Equation (12) can be rewritten as

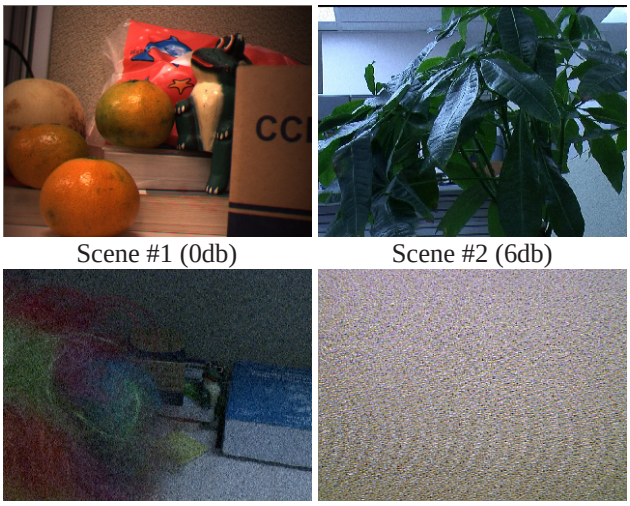
$$\begin{aligned}\mathbf{u} &= \underset{\mathbf{u}}{\operatorname{argmax}} \prod_{\mathbf{x}\in\mathbf{B}} S(F_1(\mathbf{x}), F_0(\mathbf{x}-\mathbf{u})) \qquad (13)\\ &= \underset{\mathbf{u}}{\operatorname{argmax}} \prod_{\mathbf{x}\in\mathbf{B}} \int_\Omega p(F_1(\mathbf{x})|I')p(F_0(\mathbf{x}-\mathbf{u})|I')p(I')dI'.\end{aligned}$$

To avoid computational errors caused by tiny numerical values, the actual computation is performed in the log domain:

$$\mathbf{u} = \underset{\mathbf{u}}{\operatorname{argmin}} \sum_{\mathbf{x}\in\mathbf{B}} -\log S(F_1(\mathbf{x}), F_0(\mathbf{x}-\mathbf{u})). \qquad (14)$$

Experimental configuration Experiments were performed using a PointGrey DragonFly and a Sony DSR-PD190P camera. For each camera, the similarity measure is precomputed using sequences shot of static scenes. At each gain setting, 300 frames are recorded for the 708×564-resolution DSR-PD190P and the 640×480-resolution DragonFly. Observed noise distributions that have the same true intensity value are merged together, since more observations per intensity level bring greater statistical reliability to the similarity measure. This results in roughly 450000 samples per intensity level for the DSR-PD190P and 36000 samples per intensity level for the DragonFly. Once the probabilistic intensity similarity measure is computed, the result is stored in a symmetric matrix (look-up table) of size $n \times n$, where n is the number of intensity levels.

To obtain datasets for block matching, we mounted a video camera at a fixed position and simply captured an image sequence of a static scene. Since there do not exist intensity changes except for noise, the correct matching score can be computed by counting the number of times blocks stay at their original positions, as done in [2]. The experimental datasets are of the four scenes shown in Figure 8. Scene #1 was captured using the DragonFly, and Scenes



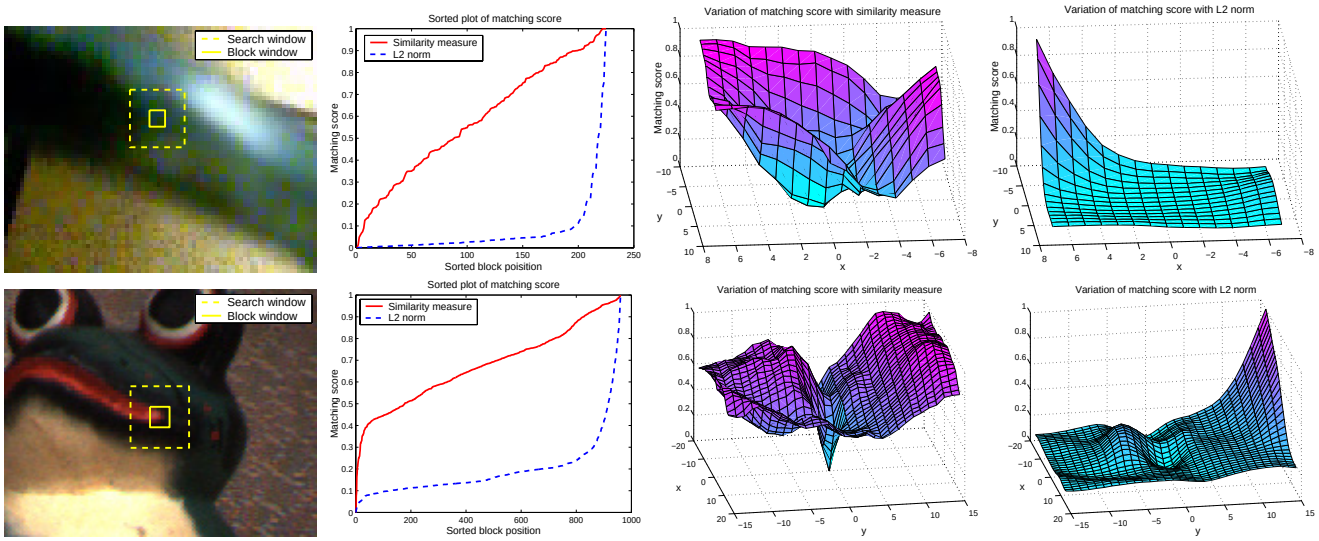
Scene #1 (0db) Scene #2 (6db)

Scene #3 (18db) Scene #4 (18db)

Figure 8. Scenes used in block matching experiments, with noise level given in parentheses.

Scene	#1	#2	#3	#4
Similarity	57.12 %	57.84 %	20.38 %	52.56 %
L^2 norm	56.78 %	57.53 %	19.35 %	48.47 %
Mean dev.	0.34 %	0.31 %	1.03 %	4.09 %
Max dev.	0.59 %	3.64 %	1.60 %	6.06 %
Min dev.	-0.09 %	-1.69 %	0.51 %	2.05 %

Table 1. Results of block matching. The top two rows show the correct matching rate using the proposed similarity measure and the standard L^2 norm. Mean dev. is the mean performance improvement of the probabilistic similarity measure in comparison to the L^2 norm. Max dev. and Min dev. represent the maximum and minimum performance improvement, respectively.



(a) Scene (b) Plot of matching scores (c) Score variation (similarity) (d) Score variation (L^2)

Figure 9. Variations in block matching scores. Top row: 5×5 block size, 15×15 search window size. Bottom row: 10×10 block size, 30×30 search window size. (a) shows the scene, one block and its search range. (b) shows a plot of sorted matching scores using the proposed similarity measure in comparison to the L^2 norm. (c) and (d) show the distribution of matching scores at different block positions in the neighbouring frames. Matching scores using the similarity measure are expressed in the negative log domain as described in Equation (14). Matching scores are mapped into the range $[0, 1]$, with lower scores indicating better matches.

#2, 3, and 4 are from the DSR-PD190P.

For performance evaluation, the captured image sequences are cropped to a 320×240 spatial resolution. We use a 3×3 block size and a 21×21 search range centered at the original block position. Block matching is evaluated using each of the non-boundary pixels, and 100 frames of each scene are used to statistically evaluate performance.

Results The results of block matching are summarized in Table 1. The top two rows list the percentage of correct matches using the proposed similarity measure and the standard L^2 norm. The third row shows the performance difference of the proposed similarity measure in comparison to the L^2 norm. Max dev. and Min dev. indicate the maximum and minimum performance difference of the proposed similarity measure over the L^2 norm. As we can see from the table, the probabilistic similarity measure leads to better results under noisy conditions (Scenes #3 and 4) in which the standard L^2 norm breaks down. At low noise levels (Scenes #1 and 2), the performance of the proposed similarity measure is similar to the L^2 norm.

While the probabilistic similarity measure does not deviate substantially from the L^2 norm in low-noise conditions, it is nevertheless useful in evaluating the reliability of estimated matches [11]. Since our proposed metric represents a measure of similarity rather than distance, it can better indicate the confidence of a match than the L^2 norm, and does so in a statistically meaningful way. Such a confidence measure can provide useful information to subsequent processes

	Mean of RMSD
Similarity	0.547
L^2	0.262

Table 2. Mean of root mean squared distance (RMSD) of the correct matching score to the set of incorrect matching scores. This value is computed after scaling the matching scores into the range [0,1]. A larger RMSD indicates greater discrimination power. The left figure shows a frame from the test image sequence.

in which the estimated motion is used, e.g., eliminating unreliable estimates. In addition, the proposed similarity measure offers greater discrimination power than the standard L^2 norm. Figure 9 shows the variation of block matching scores. In the figure, matching scores obtained within the search range are plotted in (b), (c), and (d). The matching scores using the similarity measure are computed in the $-\log$ domain as described in Equation (14), and all the obtained matching scores are scaled into the range $[0, 1]$, with smaller matching scores corresponding to better matches. In these examples, the best matches all correspond to correct matches. As we can see in the sorted matching scores in Figure 9 (b), the proposed measure provides a steeper distribution than the L^2 norm. This sharper discrimination between correct and incorrect matches brings greater confidence in matching results. Figure 9 (c) and (d) exemplify the variations in matching scores using the proposed similarity measure and the L^2 norm, respectively. It can be observed that the proposed similarity measure more clearly

differentiates correct from incorrect matches.

The same experiment is conducted using a sequence of 100 images at 320×240-resolution to measure statistical performance. To characterize the discrimination power, the root mean squared distance (RMSD) is computed, which measures the distance from the correct-matching score to each of the incorrect-matching scores, defined as:

$$RMSD = \left(\frac{1}{n_s} \sum_{s} (s^* - s)^2 \right)^{\frac{1}{2}}, \tag{15}$$

where s^* is the matching score of the correct match, s is a matching score at any position in the search range, and n_s is the number of matches tested in the search range. Table 2 shows the mean of the RMSD scores obtained from all correct matching results in the sequence. The larger RMSD of the proposed similarity measure in comparison to the L^2 norm indicates its greater power in discriminating between correct and incorrect matches.

4. Conclusion

Intensity similarity is a quantity that is difficult to define; however, the consideration of imaging noise distributions enables a statistical approach to this problem. In this paper, we propose a new intensity similarity measure in a probabilistic form that directly embeds the effect of noise. The similarity measure is formulated as the likelihood of two intensity observations having the same scene radiance. For the cases of intensity-independent and intensity-dependent Gaussian noise, we derive analytic expressions of the similarity measure. We also describe a method to compute the similarity measure for a data-driven noise model. It is also shown that the intensity similarity measure can be directly extended to handle color values, subpixel shifts, and multiresolution applications.

The proposed similarity measure is evaluated on a basic block matching technique in comparison to the standard L^2 norm. Better performance is demonstrated, particularly under noisy conditions. Moreover, we show that the proposed similarity measure naturally indicates the confidence of a match, and provides greater discrimination power between correct and incorrect matches.

References

[1] M. Abramowitz and I. A. Stegun, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, chapter 7. Error Function and Fresnel Integrals, pages 297–309. Dover, 1972. 4, 8

[2] F. Alter, Y. Matsushita, and X. Tang. An intensity similarity measure in low-light conditions. In *Proc. of European Conference on Computer Vision*, pages 267–280, 2006. 2, 6

[3] D. Comaniciu and P. Meer. Robust analysis of feature spaces: color image segmentation. In *Proc. of Computer Vision and Pattern Recognition*, pages 750–755, 1997. 2

[4] C. Fowlkes, D. R. Martin, and J. Malik. Learning affinity functions for image segmentation: Combining patch-based and gradient-based approaches. In *Proc. of Computer Vision and Pattern Recognition*, volume 2, pages 54–64, 2003. 2

[5] R. Gonzalez and R. Woods. *Digital Image Processing*. Addison-Wesley, 1992. 6

[6] G. E. Healey and R. kondepudy. Radiometric ccd camera calibration and noise estimation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(3):267–276, 1994. 1

[7] C. Liu, W. T. Freeman, R. Szeliski, and S. B. Kang. Noise estimation from a single image. In *Proc. of Computer Vision and Pattern Recognition*, volume 1, pages 901–908, 2006. 4

[8] I. Omer and M. Werman. Color lines: Image specific color representation. In *Proc. of Computer Vision and Pattern Recognition*, volume 2, pages 946–953, 2004. 2

[9] I. Omer and M. Werman. The bottleneck geodesic: Computing pixel affinity. In *Proc. of Computer Vision and Pattern Recognition*, volume 2, pages 1901–1907, 2006. 2

[10] I. Omer and M. Werman. Image specific feature similarities. In *Proc. of European Conference on Computer Vision*, volume 2, pages 321–333, 2006. 2

[11] I. Patras, E. A. Hendriks, and R. L. Lagendijk. Confidence measures for block matching motion estimation. In *Proc. of International Conference on Image Processing*, volume 2, pages 277–280, 2002. 7

[12] M. Rattray. A model-based distance for clustering. In *Proc. of the IEEE-INNS-ENNS International Joint Conference on Neural Networks*, volume 4, pages 13–16, 2000. 2

[13] T. A. Severini. *Likelihood Methods in Statistics*. Oxford University Press, 2001. 3

Appendix

When the noise variance σ^2 is constant for all intensity values, the similarity of two observations I_1 and I_2 is derived as follows:

$$\begin{aligned}
S(I_1, I_2) &= \int_\Omega p(I_1|I')p(I_2|I')p(I')dI' \\
&= \int_\Omega \frac{1}{2\pi\sigma^2|\Omega|} e^{-(I_1-I')^2/2\sigma^2} e^{-(I_2-I')^2/2\sigma^2} dI' \\
&= \underbrace{\frac{1}{2\pi\sigma^2|\Omega|} e^{-(I_1^2+I_2^2)/2\sigma^2}}_{\text{Independent of I}'} \int_\Omega e^{(I_1+I_2-I')I'/\sigma^2} dI' \\
&= \frac{1}{4\sigma|\Omega|\sqrt{\pi}} e^{-\frac{(I_1-I_2)^2}{4\sigma^2}} \left\{ \operatorname{erf}\left(\tfrac{2I'-(I_1+I_2)}{2\sigma}\right)\Big|_\Omega \right\} \\
&= \frac{1}{4\sigma|\Omega|\sqrt{\pi}} e^{-\frac{(I_1-I_2)^2}{4\sigma^2}} \left\{ \operatorname{erf}\left(\tfrac{2I_{max}-(I_1+I_2)}{2\sigma}\right) + \operatorname{erf}\left(\tfrac{I_1+I_2}{2\sigma}\right) \right\},
\end{aligned}$$

where $|\Omega| \stackrel{\text{def}}{=} p(I') = \text{const}$, $I_{max} \stackrel{\text{def}}{=} \max I'$, and $\operatorname{erf}(\cdot)$ is an error function [1] defined as

$$\operatorname{erf}(x) \stackrel{\text{def}}{=} \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$