A Locally Adaptive Window for Signal Matching *

Masatoshi Okutomi†
Information Systems Research Center
Canon Inc.
890-12 Kashimada, Saiwai-ku, Kawasaki, Japan

Takeo Kanade
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213, USA

Abstract

This paper presents a signal matching algorithm that can select an appropriate window size adaptively so as to obtain both precise and stable estimation of correspondences.

Matching two signals by calculating the sum of squared-differences (SSD) over a certain window is a basic technique in computer vision. Given the signals and a window, there are two factors that determine the difficulty for obtaining precise matching. The first is the variation of the signal within the window, which must be large enough, relative to noise, that the SSD values exhibit a clear and sharp minimum at the correct disparity. The second factor is the variation of disparity within the window, which must be small enough that signals of corresponding positions are duly compared. These two factors present conflicting requirements to the size of the matching window, since a larger window tends to increase the signal variation, but at the same time tends to include points of different disparity. A window size must be adaptively selected depending on local variations of signal and disparity in order to compute a most certain estimate of disparity at each point.

There has been little work on a systematic method for automatic window size selection. The major difficulty is that, while the signal variation is measurable from the input, the disparity variation is not, since disparities are what we wish to calculate. In this paper, we introduce a statistical model of disparity variation within a window, and employ it to establish a link between the window size and the uncertainty of the computed disparity. This allows us to choose the window size that minimizes uncertainty in the disparity computed at each point. This paper presents a theory for the model and the resultant algorithm, together with analytical and experimental results that demonstrate their effectiveness.

1 Introduction

One of the most basic methods for signal matching is calculating the sum of squared differences (SSD) between two signals over a certain window and locating the position of their minimum. [MSK88][FP86][Woo83][MKA73][LOY73]. However, precise localization of such minima, i.e., precise determination of disparity, is difficult and unreliable for two cases. The first is that when there is not enough signal variation, relative to noise, the SSD values are noisy and do not exhibit a clear and sharp minimum. The second case is that when there is too much disparity variation within the window, corresponding positions within the window are not equally shifted, and the differences are not calculated for the corresponding positions. As a result, the minimum SSD value may not occur at the correct match position. These properties pose conflicting requirements on the size of the window. Increasing the window size alleviates the first difficulty, since it increases the signal to noise ratio, while decreasing the window size alleviates the second difficulty since it limits the SSD computation to only a local, likely relevant portion of the signals. Figure 1 illustrates this problem; it shows results of matching two intensity signals by calculating SSD's with different window sizes. The intensity signals are shown in figure 1 (a) and the true disparity is shown in figure 1 (b). We can observe that for the smaller window the computed disparity is noisy, but the disparity edges are sharp; while for the larger window the computed disparity is smoother, but the disparity edges are more blurred or misplaced. Clearly, we want a larger window for the flat regions (small disparity variation) and a smaller window near the disparity edges (large disparity variation).

There has been little work on a systematic method for automatically selecting a window size locally and adaptively for signal matching. In appearance, the problem seems to be very similar to that of smoothing a signal while detecting and preserving discontinuities, for which various powerful techniques have been developed including use of Markov Random Fields [Mar84], continuation methods [Ter86], weak continuity [BZ86], and optimal amount of smoothing [Bou86]. These techniques, however, are for smoothing a signal itself by observing its properties, such as signal variation. The fundamental difference of our problem from smoothing, which makes it difficult, is that we have to deal with a disparity function which is embedded in the input signals. While the signal variation is measurable from the input, the disparity variation is not, since disparities are what we want to calculate.

As a solution to the problem, we propose to introduce a statistical model of the disparity pattern, which assumes that the disparity values within a window are generated by a random walk process starting from the value of the center point. Thus at each point within a window, its disparity is expected to be the same as that of the center point, but its variation is higher as the point is farther from the center. We employ this model to establish a link between the window size and the uncertainty of the computed disparity. This allows us to choose the window size that minimizes uncertainty in the disparity computed at each point.

In section 2 we present a mathematical framework necessary to discuss the statistical properties of disparity calculation based on SSD values in a window. Then in section 3, we analyze how the disparity variation within a win-

*This research was supported by the Defense Advanced Research Projects Agency (DOD) and monitored by the Avionics Laboratory, Air Force Wright Aeronautical Laboratories, Aeronautical Systems Division (AFSC), Wright-Patterson AFB, Ohio 45433-6543 under Contract F33615-87-C-1499, ARPA Order No. 4976, Amendment 20. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of DARPA or the U.S. government.

†This research was performed while the first author was with Carnegie Mellon University.

CH2934-8/90/0000/0190$01.00 © 1990 IEEE

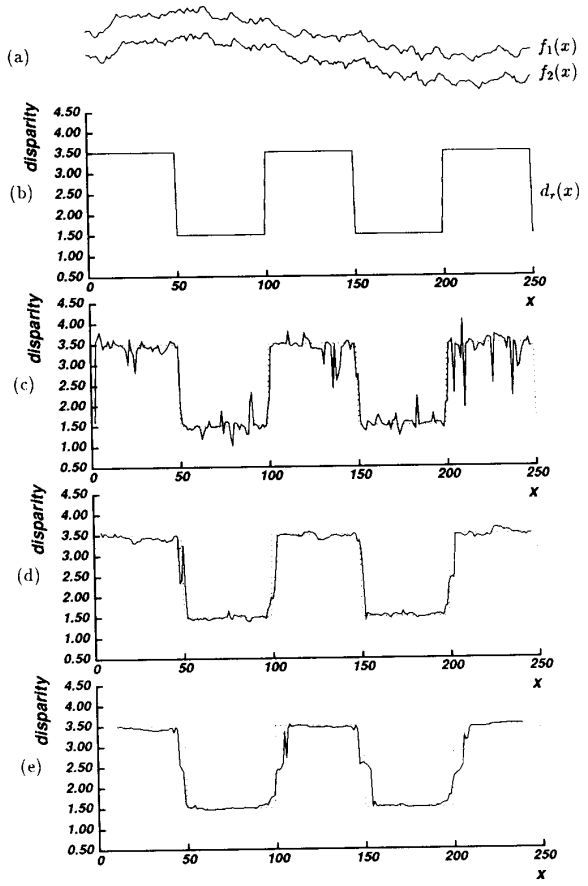


Figure 1: Matching by SSD with different window sizes. (a) Signals (b) True disparity pattern (c) Computed disparity, $w = 3$ (d) $w = 7$ (e) $w = 21$

dow affects the estimation of disparity at the center point. This analysis leads to the proposed algorithm for signal matching with a locally adaptive window. In section 4, we analyze the uncertainty of the disparity for a few typical disparity patterns. Finally, we show experimental results in section 5 which demonstrate the advantages of the proposed method.

2 Window Size in Signal Matching by SSD

Signal matching algorithms which compute the sum of squared differences (SSD) of intensity patterns within a window implicitly assume that all points in the window have equal disparity; that is, all points within the window have shifted in parallel. In this section, we will review the behavior of an iterative matching algorithm, similar to ones in [FP86] and [MO89], that makes this assumption. We will use this analysis to develop the new formulation in the next section.

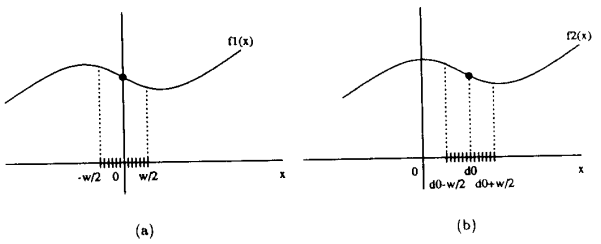


Figure 2: Signals and window settings

2.1 Matching by SSD

Let $f_1(x)$ and $f_2(x)$ be signals that have a constant disparity d_r. We can write;

$$f_1(x) = f(x) + n_1(x) \quad (1)$$
$$f_2(x) = f(x - d_r) + n_2(x), \quad (2)$$

where $n_1(x)$ and $n_2(x)$ are zero mean Gaussian white noise such that

$$n_1(x),\ n_2(x) \sim N(0, \sigma_n^2)^*$$

Then

$$f_1(x) - f_2(x + d_r) = n_1(x) - n_2(x + d_r) \equiv n(x), \quad (3)$$

where $n(x)$ is also Gaussian white such that

$$n(x) \sim N(0, 2\sigma_n^2). \quad (4)$$

If d_0 is an initial estimate of the disparity, by using the Taylor expansion,

$$f_2(x + d_r) \approx f_2(x + d_0) + \Delta d f_2'(x + d_0), \quad (5)$$

where $\Delta d = d_r - d_0$ is the correction. From (3) and (5),

$$f_1(x) - f_2(x + d_0) - \Delta d f_2'(x + d_0) = n(x). \quad (6)$$

Let

$$\psi_1(x) = f_1(x) - f_2(x + d_0)$$
$$\psi_2(x) = f_2'(x + d_0), \quad (7)$$

then

$$\psi_1(x) - \Delta d \psi_2(x) = n(x). \quad (8)$$

Next we set a window around the point whose disparity we would like to find, as shown in figure 2. For simplicity, suppose that the point whose disparity we would like to compute is at $x = 0$. Within the window, suppose that we select N sample points, $x_0, x_1, \ldots, x_{N-1}$, with equal intervals, and calculate the values of $\psi_1(x_i)$ and $\psi_2(x_i)$ from the sampled values of the intensity patterns. Let us define η_i such that

$$\eta_i = \psi_1(x_i) - \Delta d \psi_2(x_i) \qquad i = 0, \ldots, N-1. \quad (9)$$

From equations (8) and (4), the conditional distribution function of η_i, given Δd, is

$$p(\eta_i | \Delta d) = \frac{1}{2\sqrt{\pi}\sigma_n} \exp\left(-\frac{(\psi_1(x_i) - \Delta d \psi_2(x_i))^2}{4\sigma_n^2}\right). \quad (10)$$

*Here, $N(a, b)$ denotes a Gaussian distribution with mean a and variance b.

Since $n(x)$ is Gaussian white noise, the η_i's are independent of each other. So we get

$$p(\eta_0, \eta_1, \ldots, \eta_{N-1}|\Delta d) = \prod_{i=0}^{N-1} p(\eta_i|\Delta d). \quad (11)$$

From the continuous version of Bayes' theorem,

$$\begin{aligned} & p(\Delta d|\eta_0, \eta_1, \ldots, \eta_{N-1}) \\ = & \frac{p(\eta_0, \eta_1, \ldots, \eta_{N-1}|\Delta d)p(\Delta d)}{\int_{-\infty}^{\infty} p(\eta_0, \eta_1, \ldots, \eta_{N-1}|\Delta d)p(\Delta d)d(\Delta d)} \\ = & \frac{\prod_{i=0}^{N-1} p(\eta_i|\Delta d)p(\Delta d)}{\int_{-\infty}^{\infty} \prod_{i=0}^{N-1} p(\eta_i|\Delta d)p(\Delta d)d(\Delta d)}. \end{aligned} \quad (12)$$

Assuming no prior information, let $p(\Delta d) = constant$. Then we obtain

$$p(\Delta d|\eta_0, \eta_1, \ldots, \eta_{N-1}) = \frac{\prod_{i=0}^{N-1} p(\eta_i|\Delta d)}{\int_{-\infty}^{\infty} \prod_{i=0}^{N-1} p(\eta_i|\Delta d)d(\Delta d)}. \quad (13)$$

Substituting (10) into (13), we get

$$p(\Delta d|\eta_0, \eta_1, \ldots, \eta_{N-1}) = \frac{1}{\sqrt{2\pi}\sigma_{\Delta d}} \exp\left(-\frac{(\Delta d - \hat{\Delta d})^2}{2\sigma_{\Delta d}^2}\right), \quad (14)$$

where

$$\hat{\Delta d} = \frac{\sum_{i=0}^{N-1}(\psi_1(x_i)\psi_2(x_i))}{\sum_{i=0}^{N-1}(\psi_2(x_i))^2} \quad (15)$$

$$\sigma_{\Delta d}^2 = \frac{2\sigma_n^2}{\sum_{i=0}^{N-1}(\psi_2(x_i))^2}. \quad (16)$$

That is, the conditional probability density function of Δd, given the observed intensity pair, becomes a Gaussian distribution with mean $\hat{\Delta d}$ and variance $\sigma_{\Delta d}^2$. The $\hat{\Delta d}$ and $\sigma_{\Delta d}^2$ derived here are the same as those obtained from maximum likelihood estimation and standard error propagation techniques, which have been presented by multiple researchers including [MO89][FP86][RGH80].

2.2 Window Size and Uncertainty

Now, let

$$\Delta x = \frac{w}{N} \quad (17)$$

be the sampling interval, where w is the size of the window. Multiplying the numerators and the denominators of equations (15) and (16) by Δx, we obtain

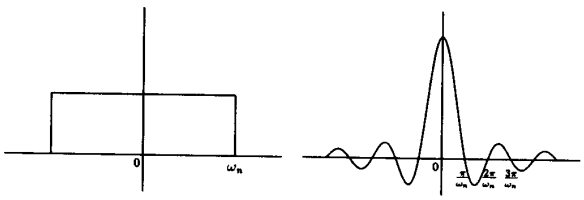
$$\hat{\Delta d} = \frac{\sum_{i=0}^{N-1}(\psi_1(x_i)\psi_2(x_i))\Delta x}{\sum_{i=0}^{N-1}(\psi_2(x_i))^2\Delta x} \quad (18)$$

$$\sigma_{\Delta d}^2 = \frac{2\sigma_n^2\Delta x}{\sum_{i=0}^{N-1}(\psi_2(x_i))^2\Delta x}. \quad (19)$$

As $N \to \infty$,

$$\hat{\Delta d} = \frac{\int_{-\frac{w}{2}}^{\frac{w}{2}}(\psi_1(x)\psi_2(x))dx}{\int_{-\frac{w}{2}}^{\frac{w}{2}}(\psi_2(x))^2dx} \quad (20)$$

$$\sigma_{\Delta d}^2 = \frac{2\sigma_n^2\Delta x}{\int_{-\frac{w}{2}}^{\frac{w}{2}}(\psi_2(x))^2dx} \to 0. \quad (21)$$



(a)Power spectral density function (b)Autocorrelation function

Figure 3: Band-limited white noise

This result is somewhat counterintuitive. Here, it appears as if the variance of the estimated Δd could be made arbitrarily small by sampling as many points as we need. This would only be possible if $n(x)$ were actually *white*. However, in practice it is not, and we must model that $n(x)$ is *band-limited* Gaussian white noise. Figures 3 (a) and (b) show the power spectral density function and the autocorrelation function of band-limited white noise. We can see that points evenly spaced by $\frac{\pi}{\omega_n}$ are uncorrelated, where ω_n is the maximum frequency of $n(x)$. They are also independent of each other, since uncorrelated Gaussian variables are statistically independent [dC86]. Therefore, the sampling interval Δx that keeps the samplings independent of each other and makes $\sigma_{\Delta d}^2$ the smallest is

$$\Delta x = \frac{\pi}{\omega_n}. \quad (22)$$

With this sampling period we can obtain

$$\sigma_{\Delta d}^2 = \frac{2\pi\sigma_n^2}{\omega_n \int_{-\frac{w}{2}}^{\frac{w}{2}}(\psi_2(x))^2dx}. \quad (23)$$

Substituting equation (7) into equations (20) and (23), we get

$$\hat{\Delta d} = \frac{\int_{-\frac{w}{2}}^{\frac{w}{2}}(f_1(x) - f_2(x + d_0))f_2'(x + d_0)dx}{\int_{-\frac{w}{2}}^{\frac{w}{2}}(f_2'(x + d_0))^2dx} \quad (24)$$

$$\sigma_{\Delta d}^2 = \frac{2\pi\sigma_n^2}{\omega_n \int_{-\frac{w}{2}}^{\frac{w}{2}}(f_2'(x + d_0))^2dx}. \quad (25)$$

These equations show that, given the signals $f_1(x)$ and $f_2(x)$ and an initial estimate d_0, we can obtain a disparity correction $\hat{\Delta d}$ and its uncertainty of $\sigma_{\Delta d}$. We can revise the disparity estimate by replacing d_0 with $d_0 + \hat{\Delta d}$, and iterate the process. The results shown in figure 1 were obtained with this method.

Equation (25) indicates two characteristics. First, the larger the absolute value of the derivative of the signal is, the smaller the uncertainty. Second, the larger the window size is, the smaller the uncertainty. The first characteristic is intuitive. That is, the more fluctuation in the intensity pattern, the more reliable the matching. The second characteristic is also understandable because a large window can average out the effect of noise. However, this is true *only because* we have assumed that the two signals have equal disparity everywhere. If the disparity varies from position to position, as happens in practice, a large window is likely to compare intensities which are not actually at corresponding positions, even when the center point

of the window is at the corresponding position. As a result, a larger window has the effect that disparity edges are blurred and misplaced, while a smaller window yields sharp but noisy results. This is the phenomena that we have observed in figure 1. Clearly, there should be an appropriate window size for each position.

3 Matching with a Locally Adaptive Window

In the previous section, we analyzed an iterative matching method with the assumption that the disparity within the window is constant. If this assumption were true, a larger window would be better. However, in practice the actual disparity varies from point to point, and too large a window is actually harmful. An appropriate window size must exist depending on local intensity *and* disparity patterns.

3.1 Assumption of Non-constant Disparity

In this section, we let the disparity be a function of position, i.e. $d_r(x)$. Therefore instead of equation (3), we have

$$f_1(x) - f_2(x + d_r(x)) = n(x). \tag{26}$$

Using the Taylor expansion,

$$f_2(x+d_r(x)) \approx f_2(x+d_r(0)) + (d_r(x)-d_r(0))f_2'(x+d_r(0)). \tag{27}$$

Substituting equation (27) into equation (26), we get

$$f_1(x) - f_2(x+d_r(0)) = (d_r(x)-d_r(0))f_2'(x+d_r(0)) + n(x). \tag{28}$$

Comparing this with equation (3), we observe an additional term $e(x)$, i.e.

$$e(x) = (d_r(x) - d_r(0))f_2'(x + d_r(0)). \tag{29}$$

The interpretation of this term is the following: suppose that we place the center point of a window at the corresponding points of both signals and compare the intensity values at each position x within the window. Then $e(x)$ represents the *apparent* intensity difference between the two signals caused not by mismatch, but by the difference between the disparity $d_r(x)$ at that position and the disparity $d_r(0)$ at the center point of the window. Figure 4 illustrates this situation. We can see that although the centers of the windows are placed at the exactly corresponding points (p1 and p'1 in the figure), other points in the window do not correspond properly.

3.2 A Statistical Model of Disparity

In order to advance our analysis of matching, we must consider the effect of $e(x)$ on disparity estimate. The difficulty is that $e(x)$ includes $d_r(x)$, which is what we wish to calculate. As a solution, we introduce the following statistical model for the disparity $d_r(x)$ within a window:

$$d_r(x) - d_r(0) \sim N(0, \alpha_d|x|), \tag{30}$$

where α_d is a constant that represents fluctuation of the disparity. This model assumes that the difference in disparity between a point x in the window and the center point of the window has a zero-mean Gaussian distribution with variance proportional to the distance between these points. In other words, the farther the point is from the center, the more likely it is that it will have a different disparity (though the expected value of the disparity is

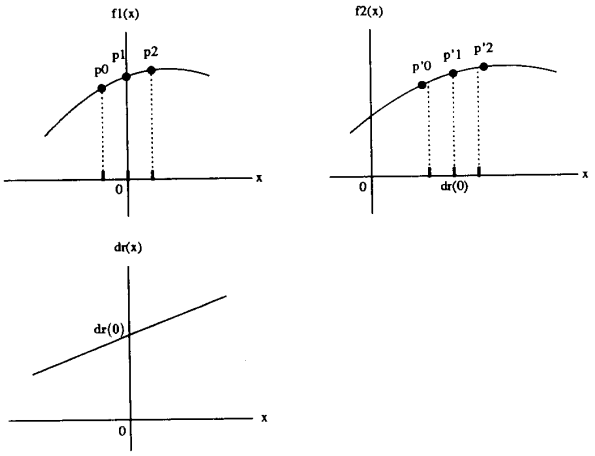


Figure 4: Matching in the presence of linearly changing disparity: Although the window centers may be placed correctly for the corresponding points, p1 and p'1, other points in the window do not correspond properly.

the same).[†] We also assume that the intensity derivatives $f_2'(x)$ within a window follow a zero-mean Gaussian white distribution[‡] and are independent of $d_r(x)$.

From these assumptions and equation (29), we can show that $e(x)$ can be approximated by Gaussian white noise such that

$$e(x) \sim N(\mu_e, \sigma_e^2), \tag{31}$$

where

$$\begin{aligned} \mu_e &= E[e(x)] \\ &= E[d_r(x) - d_r(0)]E[f_2'(x + d_r(0)] \\ &= 0 \end{aligned} \tag{32}$$

$$\begin{aligned} \sigma_e^2 &= E[(e(x))^2] \\ &= E[(d_r(x) - d_r(0))^2]E[(f_2'(x + d_r(0)))^2] \\ &= \alpha_f \alpha_d |x|, \end{aligned} \tag{33}$$

and

$$\alpha_f = E[(f_2'(x + d_r(0)))^2]. \tag{34}$$

Appendix A of [KO90] shows this approximation. From equations (28), (29), (32), and (33), we obtain

$$\begin{aligned} f_1(x) - f_2(x + d_r(0)) &= e(x) + n(x) \\ &\equiv n_s(x), \end{aligned} \tag{35}$$

where $n_s(x)$ is Gaussian white noise such that

$$n_s(x) \sim N[0, \sigma_s^2(x)], \tag{36}$$

[†]The statistical model of (30) can be shown equivalent to assuming that $d_r(x)$ is generated by Brownian motion (refer to [BN68][Vos87]). More generally, we can assume $d_r(x)$ to be a fractal. This corresponds to choosing a different degree of $|x|$ in the variance in (30). The Brownian motion is the simplest case in which the degree is 1 (see Appendix A). However, our preliminary experiments have shown no noticeable advantage in using a general fractal assumption.

[‡]This is also equivalent to assuming the pattern $f_2(x)$ to be the result of Brownian motion.

and

$$2\sigma_n^2 + \alpha_f \alpha_d |x| = \sigma_s^2(x). \quad (37)$$

Intuitively, $n_s(x)$ can be considered as noise added to signals, whose variance is the sum of a constant term and a term proportional to $|x|$. The first term comes from the noise added to the signals. The second term comes from an *uncertain local support*. That is, while the surrounding points of the center point in the window are used to support the matching for the center point, the information from these points adds some uncertainty, too, because of the disparity difference between the center point and the supporting points. This uncertainty is represented as if additional noise were added whose power is proportional to the distance from the center point in the window. If the disparity is constant over the window, i.e. $\alpha_d = 0$, this additional noise is zero. As the disparity varies more in the window (i.e., larger α_d), this additional noise becomes larger and the information from the supporting points becomes more uncertain. Also, note that the noise effect of the disparity variation is amplified by a factor of α_f, that is, by the amount of the intensity variation. This is because wrong correspondences due to disparity variation affect more severely when the intensity variation is higher.

3.3 Uncertainty of Estimation as a Function of Signal and Disparity Variations

Now, as we obtained equation (6) from equation (3) for iterative estimation, we can obtain from equation (35) the following:

$$f_1(x) - f_2(x + d_0) - \Delta d f_2'(x + d_0) = n_s(x). \quad (38)$$

Dividing both sides of this equation by $\sigma_s(x)$,

$$\frac{f_1(x) - f_2(x + d_0) - \Delta d f_2'(x + d_0)}{\sigma_s(x)} = \frac{n_s(x)}{\sigma_s(x)} \sim N(0,1). \quad (39)$$

By letting

$$\phi_1(x) = \frac{f_1(x) - f_2(x + d_0)}{\sigma_s(x)} \quad (40)$$

$$\phi_2(x) = \frac{f_2'(x + d_0)}{\sigma_s(x)}, \quad (41)$$

we have

$$\phi_1(x) - \Delta d \phi_2(x) \sim N(0,1), \quad (42)$$

which corresponds to equation (8) for the case of constant disparity within a window. So, instead of equation (20) and (23), we obtain

$$\hat{\Delta d} = \frac{\int_{-\frac{w}{2}}^{\frac{w}{2}} (\phi_1(x)\phi_2(x))dx}{\int_{-\frac{w}{2}}^{\frac{w}{2}} (\phi_2(x))^2 dx} \quad (43)$$

$$\sigma_{\Delta d}^2 = \frac{\pi}{\omega_n \int_{-\frac{w}{2}}^{\frac{w}{2}} (\phi_2(x))^2 dx}. \quad (44)$$

Substituting equation (40), (41), and (37), we finally obtain

$$\hat{\Delta d} = \frac{\int_{-\frac{w}{2}}^{\frac{w}{2}} \frac{(f_1(x) - f_2(x+d_0)) f_2'(x+d_0)}{2\sigma_n^2 + \alpha_f \alpha_d |x|} dx}{\int_{-\frac{w}{2}}^{\frac{w}{2}} \frac{(f_2'(x+d_0))^2}{2\sigma_n^2 + \alpha_f \alpha_d |x|} dx} \quad (45)$$

$$\sigma_{\Delta d}^2 = \frac{\pi}{\omega_n \int_{-\frac{w}{2}}^{\frac{w}{2}} \frac{(f_2'(x+d_0))^2}{2\sigma_n^2 + \alpha_f \alpha_d |x|} dx}. \quad (46)$$

These are the equations for correction of the disparity estimate and uncertainty of the correction under the assumption of non-constant disparity within a window. If $\alpha_d = 0$, that is, if all points have the same disparity, then these equations reduce to equations (24) and (25) in the previous section. The values of α_f and α_d depend on the intensity and disparity patterns respectively, and they change locally. Since they include $d_r(x)$, we cannot obtain their true values. Instead we calculate their estimates by using the current estimates of $\hat{d}_r(x)$. From equations (34) and (30).

$$\hat{\alpha_f} = \frac{1}{w} \int_{-\frac{w}{2}}^{\frac{w}{2}} (f_2'(x + \hat{d}_r(0)))^2 dx \quad (47)$$

$$\hat{\alpha_d} = \frac{1}{w} \int_{-\frac{w}{2}}^{\frac{w}{2}} \frac{(\hat{d}_r(x) - \hat{d}_r(0))^2}{|x|} dx. \quad (48)$$

These are the key values that furnish a link between the uncertainty of disparity estimate and the size of a window for matching.

3.4 Iterative Algorithm

We propose the following algorithm that uses a locally adaptive window based on the preceding analysis:

1. Set $\hat{d_r}(x)$ to an initial estimate $d_0(x)$.
2. For a window size w, compute $\hat{\alpha_f}$ and $\hat{\alpha_d}$ from the signals $f_1(x)$, $f_2(x)$, and $\hat{d_r}(x)$ by using equations (47) and (48).
3. Compute a correction $\hat{\Delta d}$ and an uncertainty of the correction $\sigma_{\Delta d}$ by using equations (45) and (46).
4. Repeat steps 2 and 3 for various window sizes and use the one that provides the estimate with the smallest uncertainty.
5. Update $\hat{d}_r(x)$ by the amount $\hat{\Delta d}$.
6. Repeat steps 2 through 5 until convergence or up to a certain number of iterations.

4 Uncertainty Analysis for Typical Disparity Patterns

In the previous section, we introduced a statistical model for the disparity $d_r(x)$. We then derived equations that relate the disparity estimate and its uncertainty to intensity and disparity variations within the window. The value $\sigma_{\Delta d}^2$ in equation (46) is the variation of the estimated disparity or the uncertainty of the estimation. The adaptive window selection method proposed in the previous section chooses the window size that minimizes this value. In this section, we will analyze the behavior of $\sigma_{\Delta d}^2$ for a few typical disparity patterns: constant, linear, and step. By examining how $\sigma_{\Delta d}^2$ changes with the window size, we can tell how the proposed method will work for those typical cases.

Even when we fix the disparity pattern in the window, the value of $\sigma_{\Delta d}^2$ still depends on a particular intensity pattern $f_2(x)$. Therefore we will instead use the expected values of $\sigma_{\Delta d}^2$ over an ensemble of intensity patterns whose derivatives $f_2'(x)$ follow a zero-mean Gaussian white distribution such that

$$f_2'(x) \sim N(0, \alpha_f). \quad (49)$$

In the following analysis, we assume that α_f is constant, i.e. the intensity fluctuates equally over the whole signal. Then, we have

$$E\left[\int_{-\frac{w}{2}}^{\frac{w}{2}} \frac{(f_2'(x+d_0))^2}{2\sigma_n^2+\alpha_f\alpha_d|x|}dx\right] = \alpha_f\int_{-\frac{w}{2}}^{\frac{w}{2}} \frac{1}{2\sigma_n^2+\alpha_f\alpha_d|x|}dx$$
$$= \frac{2}{\alpha_d}\log\left(1+\frac{\alpha_d\alpha_f w}{4\sigma_n^2}\right) \quad (50)$$

Substituting this into equation (46),

$$\sigma_{\hat{\Delta}d}{}^2 = E[\sigma_{\Delta d}^2] = \frac{\pi\alpha_d}{2\omega_n\log\left(1+\frac{\alpha_d r^2 w}{4}\right)}, \quad (51)$$

where

$$r = \frac{\sqrt{\alpha_f}}{\sigma_n}. \quad (52)$$

Since α_f is the variation in the first derivative of the intensity signal and σ_n^2 is the power of additive noise, r represents *the ratio of the intensity fluctuation (ie. signal) to the noise*. This is an important parameter in matching two intensity patterns.

4.1 Constant

Suppose all points have an equal disparity within a window. That is,

$$\hat{\alpha_d} = 0. \quad (53)$$

Then,

$$\sigma_{\hat{\Delta}d} = \sqrt{\lim_{\alpha_d\to 0}\frac{\pi\alpha_d}{2\omega_n\log\left(1+\frac{\alpha_d r^2 w}{4}\right)}}$$
$$= \frac{1}{r}\sqrt{\frac{2\pi}{\omega_n w}}. \quad (54)$$

Here, we can see that the uncertainty $\sigma_{\hat{\Delta}d}$ is inversely proportional to the ratio r and to the square root of window size $\sqrt{w}$. Therefore our adaptive window method will choose the largest window size allowed. This is consistent with the observation made in section 2.

4.2 Linear

Suppose that the disparity changes linearly within a window as shown in figure 5; that is, the window includes a slanted surface:

$$d_r(x) = ax + b, \quad (55)$$

where a and b are constants. Then,

$$\hat{\alpha_d} = \frac{a^2 w}{4}. \quad (56)$$

Substituting this into equation (51),

$$\sigma_{\hat{\Delta}d} = \frac{|a|}{2}\sqrt{\frac{\pi w}{2\omega_n\log\left(1+\frac{a^2r^2w^2}{16}\right)}}. \quad (57)$$

Figure 6 shows how $\sigma_{\hat{\Delta}d}$ changes with the window size w for a few values of the slope a. We observe that for a fixed window size the steeper the slope is, the larger the uncertainty. There is a minimum of $\sigma_{\hat{\Delta}d}$ for each slope a. The dotted line connects those minima. If we denote by w_{opt} the window size that gives the minimum of $\sigma_{\hat{\Delta}d}$, then

$$w_{opt} = \frac{4k_1}{|a|r}, \quad (58)$$

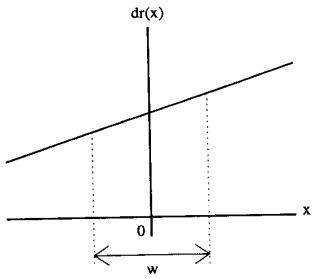


Figure 5: Linear disparity pattern

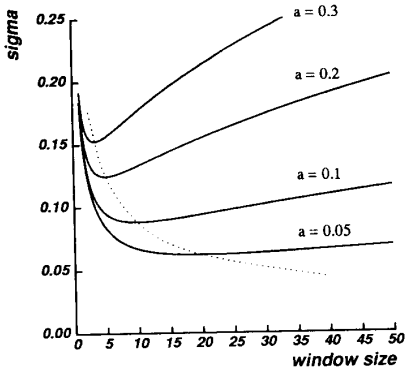


Figure 6: $\sigma_{\hat{\Delta}d}$ vs. window size w (linear disparity pattern, r=8)

where k_1 is a constant that satisfies

$$\log(1+k_1^2) - \frac{2k_1^2}{1+k_1^2} = 0. \quad (59)$$

We can see that given the ratio r, there is an optimal window size which is inversely proportional to the absolute value of the slope a.

4.3 Step

Suppose that there is a disparity jump within the window, that is, the window includes an occluding edge. Figure 7 shows a step pattern of disparity, where the step is positioned $x1$ away from the window center:

$$d_r(x) = \begin{cases} d_1 & x < x_1 \\ d_1 + d_s & x \geq x_1. \end{cases} \quad (60)$$

Then,

$$\hat{\alpha_d} = \begin{cases} 0 & \frac{w}{2} < x_1 \\ \frac{d_s^2}{w}\log\frac{w}{2x_1} & \frac{w}{2} \geq x_1 \end{cases}. \quad (61)$$

Substituting this into equation (51),

$$\sigma_{\hat{\Delta}d} = \begin{cases} \frac{1}{r}\sqrt{\frac{2\pi}{\omega_n w}} & \frac{w}{2} < x_1 \\ d_s\sqrt{\frac{\pi\log\frac{w}{2x_1}}{2\omega_n w\log(1+\frac{d_s^2r^2}{4}\log\frac{w}{2x_1})}} & \frac{w}{2} \geq x_1 \end{cases}. \quad (62)$$

Figure 8 shows plots of $\sigma_{\hat{\Delta}d}$ versus the window size w for different step sizes d_s. In general, the uncertainty has

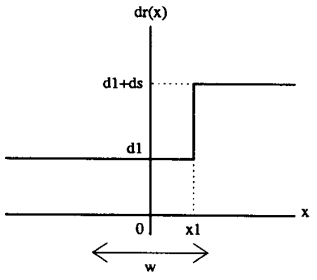


Figure 7: Step disparity pattern

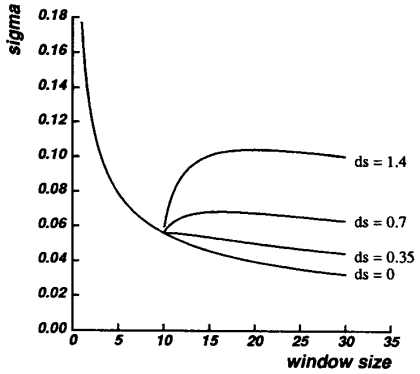


Figure 8: $\sigma_{\hat{\Delta}d}$ vs. window size w (step disparity pattern, $x_1 = 5$, $r = 8$)

a local minimum at $w = 2x_1$, if $d_s^2 > 8/r^2$ (in the case of the figure, $d_s > 0.35$). That is, the uncertainty will reach a minimum when the window is just about to cover the step, unless the disparity step is very small relative to the intensity noise level. Therefore, the proposed adaptive window method will choose the largest window that doesn't cross the disparity edge.

5 Experimental Results

In this section, we show some experimental results of matching using synthesized signals and scanlines of real stereo images. The underlying intensity pattern (ie., $f(x)$) of the synthesized signals is created by Brownian motion so that its derivatives are Gaussian white noise. The pattern is then transformed into the two intensity signals according to equations (1) and (2); by the addition of Gaussian white noise and transformation by an embedded disparity function $d_r(x)$. The signals which have been used in Figure 1(a) are examples where the disparity function is a square wave.

First we examine the case where the disparity pattern is constant ($d_r(x) = 2.5$). Figure 9 is a plot of the RMS error of the computed disparity vs. the ratio r. Figure 10 shows the RMS error as a function of the window size w. The dotted lines show the theoretical values from equation (54) in the previous section.

Next we look at the case where the disparity changes linearly. In figure 11, the solid curve shows for various

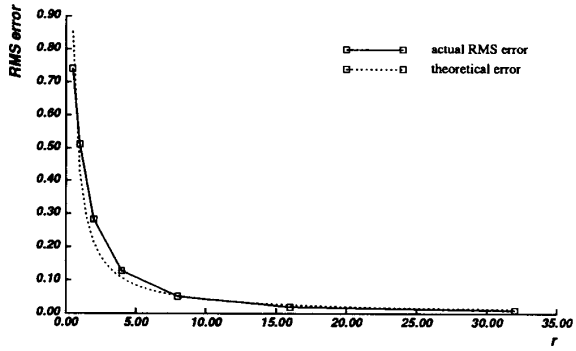


Figure 9: RMS error vs. r (constant disparity pattern)

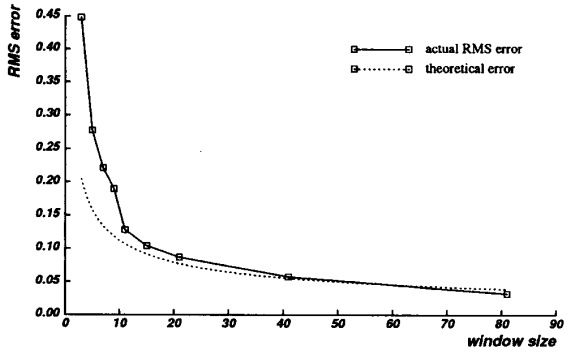


Figure 10: RMS error vs. window size (constant disparity pattern)

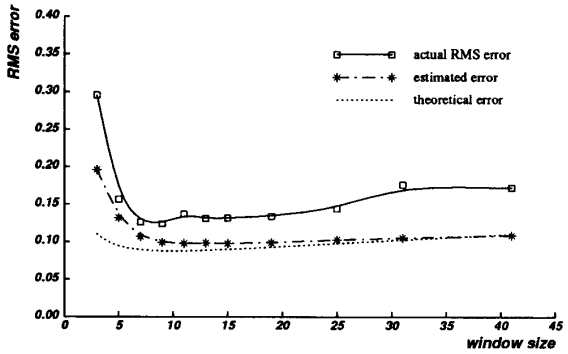


Figure 11: RMS error vs. window size (linear disparity pattern)

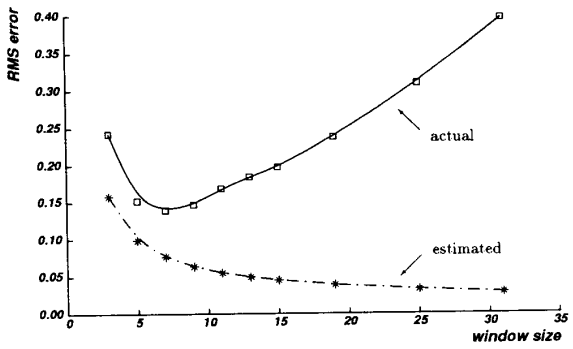


Figure 12: RMS error vs. window size (linear disparity pattern) by a conventional method

window sizes the actual RMS errors of the disparity values obtained by using equation (45) in the new formulation in section 3, and the dashed line shows the estimated error $\sigma_{\Delta d}$ calculated by using equation (46). The dotted line shows the theoretically expected error from equation (57). For comparison, figure 12 shows the results by the conventional formulation with the (implicit) constant disparity assumption, ie., equation (24) in section 2: the actual RMS errors are shown by the solid curve, together with the estimated error from equation (25). First of all, by comparing the solid curves in figures 11 and 12 we can see that the computed disparity from the new formulation has less RMS error than the conventional method over all of the window sizes. More importantly, the new formulation can give consistently better estimate for the error in computed disparity: that is, the estimated error is closer to the actual error. This provides a solid foundation for automatic selection of an appropriate window size. In the conventional formulation, the estimated error is far from the actual error as shown in figure 12. This is due to the inappropriate assumption of constant disparity as we mentioned before. Figure 13 shows the selected window sizes that give the minimum error estimation for each slope a. The dotted line shows the theoretically optimal window size from equation (58).

For the third experiment, figure 14 shows the computed disparities for a square-wave disparity pattern. In figure 1, we showed how the window size affects matching results. Comparing figure 14 with figure 1, clear improvement can be seen. Figure 14 (b) shows the window size that was actually selected by the algorithm (minimum and maximum sizes were limited to 3 and 21, respectively). We can observe that in general the method selects a small window near disparity edges according to the distance to the edges and a larger window in the flat regions.

We also tested a more complicated disparity pattern, shown in figure 15 (b), which includes slopes, steps, and a curve. Ten experiments were performed using different but statistically the same intensity patterns and noise for the same disparity pattern to evaluate the robustness of the method. One of the ten signal pairs used are shown in figure 15 (a). Figure 15 (c) shows the computed disparities by the matching algorithm with a locally adaptive window: results for the ten experiments are overlaid. Again for comparison, the test signals were also subjected to the

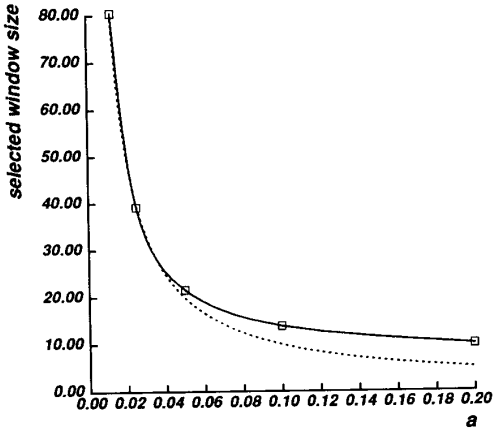


Figure 13: Selected window size vs. slope a

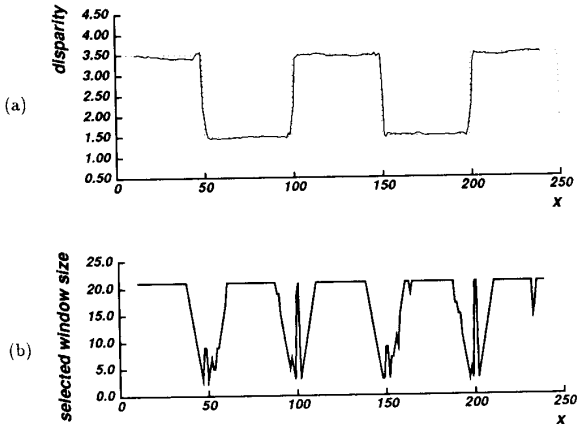


Figure 14: Step disparity (a) Computed disparity (b) Selected window size

conventional matching method with a fixed size window. Figure 16 is a plot of the RMS error over the whole pattern for various fixed window sizes. All of these RMS errors are larger than the RMS error (0.10 shown by the dotted line in the figure) achieved by the locally adaptive window method.

Figure 15 (d) shows the selected window sizes (averaged over ten experiments). Note that the window size is adaptively chosen near the step of disparity (e.g. $x = 50, 150, 190, 230, 260$), around the round shape of the disparity (i.e. $50 < x < 150$), and over the different disparity slope (i.e. $300 < x < 450$).

Figure 17 plots the actual RMS error vs. the estimated error (uncertainty $\sigma_{\Delta d}$) for all the points in figure 15. Despite the fact that they are the mixture of various disparity patterns and various window sizes, the estimated error and the actual RMS error are linearly related.

Lastly, we will show disparity estimation by using signals from real stereo images. Figure 18 (a) shows the stereo images (top down views of a scale model of buildings).

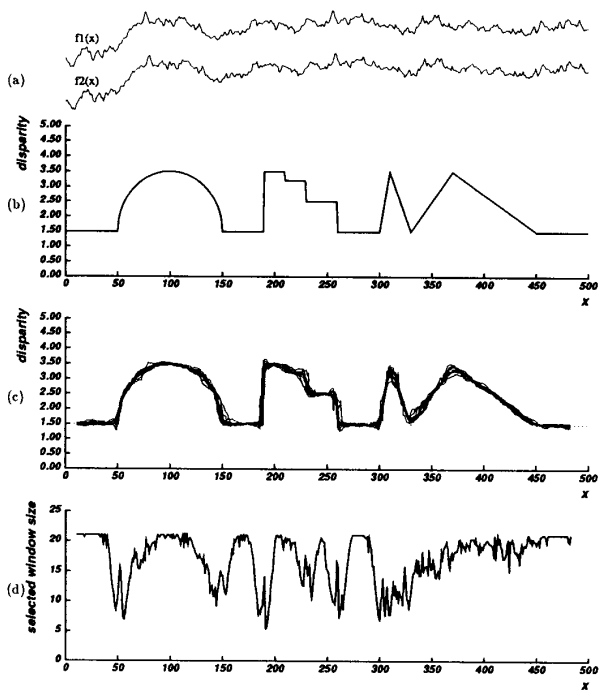


Figure 15: (a) Signals (b) True disparity (c) Computed disparity (d) Selected window size

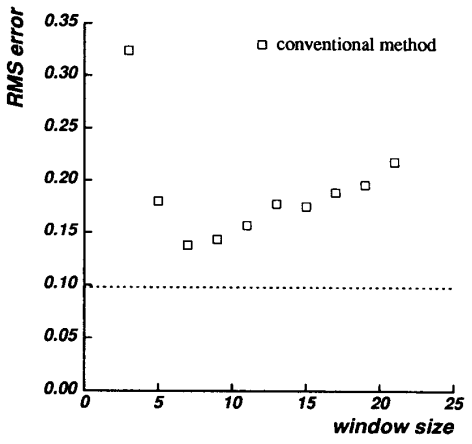


Figure 16: RMS error

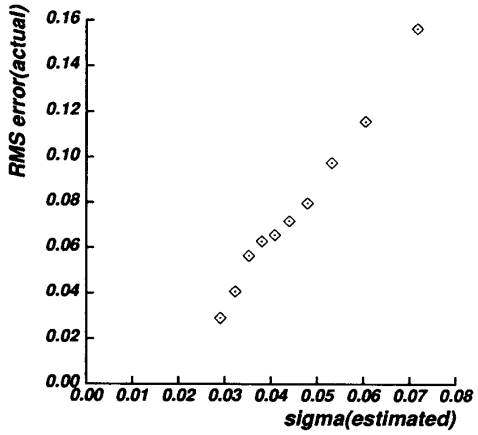


Figure 17: Actual RMS error vs. estimated $\sigma_{\Delta d}$

Scanlines marked by black lines are shown in figure 18 (b). Figure 18 (c) is the computed disparity. For comparison, the oblique view of the model is shown to the right. The selected window size is plotted in figure 18 (d).

6 Conclusions

We have presented a new signal matching method which can select appropriate window sizes adaptively. This method is based on a statistical model of disparity distribution within the window. We assume that disparities have the same expected value, but their variation from that expected value increases with the distance from the center point of the window. This model has enabled us to correctly evaluate the influence of the disparity fluctuation within the window on the computation of disparity, so that the estimated uncertainty of the computed disparity is close to the real error of the computed disparity. As a result we can choose the window size that provides the disparity estimate with minimum uncertainty.

The analytical and experimental results have shown that the method is effective for various disparity and intensity patterns. The estimated disparities obtained here are better than those from a conventional method with any fixed window size, and the estimated uncertainties show good correlation with the actual RMS error of the estimated disparities. We have extended this idea to the two-dimensional case where an appropriate size and shape of a window must be selected for stereo matching [KO90].

Acknowledgment

The authors would like to thank John Krumm, Carlo Tomasi, Atsushi Yokoyama, and Larry Matthies for useful discussions and comments on this paper. Thanks to Keith Gremban, Carol Novak, and Jim Rehg who read the manuscript carefully and greatly improved its readability.

Appendix A: Assuming Fractal for $d(x)$

Here we assume $d(x)$ to be fractal, then instead of equation (30), we have

$$d_r(x) - d_r(0) \approx N(0, \alpha_d |x|^{2H}), \tag{63}$$

where the parameter H has a value $0 < H < 1$. When $H = \frac{1}{2}$, this equation represents the case that $d(x)$ is Brownian

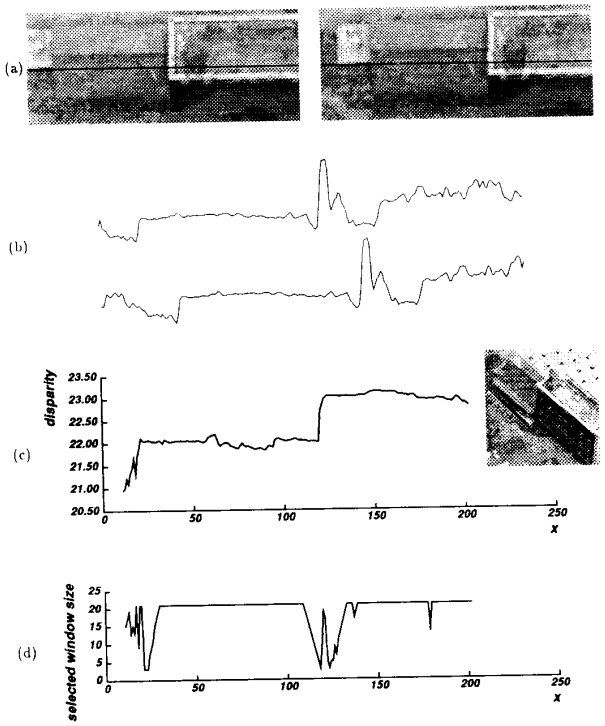


Figure 18: Signals from real images (a) Stereo images (b) Selected scan lines (c) Computed disparity (d) Selected window size

motion.
Then, instead of the final equations (45) and (46), we get

$$\hat{\Delta d} = \frac{\int_{-\frac{w}{2}}^{\frac{w}{2}} \frac{(f_1(x)-f_2(x+d_0))f_2'(x+d_0)}{2\sigma_n^2+\alpha_f\alpha_d|x|^{2H}}dx}{\int_{-\frac{w}{2}}^{\frac{w}{2}} \frac{(f_2'(x+d_0))^2}{2\sigma_n^2+\alpha_f\alpha_d|x|^{2H}}dx} \quad (64)$$

$$\sigma^2_{\Delta d} = \frac{\pi}{\omega_n \int_{-\frac{w}{2}}^{\frac{w}{2}} \frac{(f_2'(x+d_0))^2}{2\sigma_n^2+\alpha_f\alpha_d|x|^{2H}}dx}. \quad (65)$$

Furthermore, instead of equation (48), we obtain:

$$\hat{\alpha_d} = \frac{1}{w}\int_{-\frac{w}{2}}^{\frac{w}{2}} \frac{(d_r(x)-d_r(0))^2}{|x|^{2H}}dx. \quad (66)$$

References

[BN68] B.B.Mandelbrot and B.J.Van Ness. Fractional brownian motion, fractional noises and applications. *SIAM*, 10(4):422–438, 1968.

[Bou86] T. E. Boult. *Information Based Complexity in Non-Linear Equations and Compiuter Vision.* PhD thesis, Dept. of Computer Science, Columbia University, 1986.

[BZ86] A. Blake and A. Zisserman. Invariant surface reconstruction using weak continuity constraint. In *International Conference on Computer Vision and Pattern Recognition*, pages 62–68. IEEE, 1986.

[dC86] Frederic de Coulon. *Signal Theory and Processing.* Artech House, Inc., 1986.

[FP86] Wolfgang Forstner and Alfred Pertl. *Photogrammetric Standard Methods and Digital Image Matching Techniques for High Precision Surface Measurements*, pages 57–72. Elsevier Science Publishers B.V., 1986.

[KO90] Takeo Kanade and Masatoshi Okutomi. A stereo matching algorithm with an adaptive window: Theory and experiment. Technical Report CMU-CS-90-120, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, 1990.

[LOY73] Martin D. Levine, Douglas A. O'Handley, and Gary M. Yagi. Computer determination of depth maps. *Computer Graphics and Image Processing*, 2(4):131–150, 1973.

[Mar84] J. L. Marroquin. Surface reconstruction preserving discontinuities. Technical Report A. I. Memo 792, MIT, 1984.

[MKA73] Kenichi Mori, Masatsugu Kidode, and Haruo Asada. An iterative prediction and correction method for automatic sterocomparison. *Computer Graphics and Image Processing*, 2:393–401, 1973.

[MO89] Larry Matthies and Masatoshi Okutomi. A bayesian foundation for active stereo vision. In *SPIE, Sensor Fusion II: Human and Machine Strategies*, November 1989.

[MSK88] Larry Matthies, Richard Szeliski, and Takeo Kanade. Incremental estimation of dense depth maps from image sequenses. In *Proc. CVPR88*, pages 366–374, June 1988.

[RGH80] T. W. Ryan, R. T. Gray, and B. R. Hunt. Prediction of correlation errors in stereo-pair images. *Optical Engineering*, 19(3):312–322, May 1980.

[Ter86] Demetri Terzopoulos. Regularization of inverse visual problems involving discontinuities. *IEEE transaction on pattern analysis and machine inteligence*, 8(4):413–424, July 1986.

[Vos87] Richard F. Voss. Fractals in nature. In *Course note on FRACTALS: Introduction, Basics, and Perspectives*, 1987.

[Woo83] G.A. Wood. Realities of automatic correlation problem. *Photogrammetric Engineering and Remote Sensing*, 49:537–538, April 1983.