A NEW LIKELIHOOD FUNCTION FOR STEREO MATCHING
How to Achieve Invariance to Unknown Texture, Gains and Offsets?

Sanja Damjanović, Ferdinand van der Heijden and Luuk J. Spreeuwers
Chair of Systems and Signals, Faculty of EEMCS, University of Twente, Drienerlolaan 5, Enschede, The Netherlands

Keywords: Likelihood, NCC, Probabilistic framework, HMM, Stereo reconstruction.

Abstract: We introduce a new likelihood function for window-based stereo matching. This likelihood can cope with unknown textures, uncertain gain factors, uncertain offsets, and correlated noise. The method can be fine-tuned to the uncertainty ranges of the gains and offsets, rather than a full, blunt normalization as in NCC (normalized cross correlation). The likelihood is based on a sound probabilistic model. As such it can be directly used within a probabilistic framework. We demonstrate this by embedding the likelihood in a HMM (hidden Markov model) formulation of the 3D reconstruction problem, and applying this to a test scene. We compare the reconstruction results with the results when the similarity measure is the NCC, and we show that our likelihood fits better within the probabilistic frame for stereo matching than NCC.

1 INTRODUCTION

Stereo correspondence is the process of finding pairs of matching points in two images that are generated by the same physical 3D surface in space, (Faugeras, 1993). The classical approach is to consider image windows around two candidate points, and to evaluate a similarity measure (or dissimilarity measure) between the pixels inside these windows. Such an approach is based on the constant brightness assumption (CBA) stating that, apart from noise, the image data in two matching windows are equal. If the noise is white and additive, then the SSD measure (sum of squared differences), or the SAD (sum of absolute differences) is appropriate. Often, the gains and offsets with which the two images are acquired are not equal, and are not precisely known. Therefore, another popular similarity measure is the NCC (normalized cross correlation) which neutralizes these offsets and gains. An alternative is the mutual information, (Egnal, 2000), which is even invariant to a bijective mapping between the grey levels of the left and right images.

In a probabilistic approach to stereo correspondence, the similarity measures become likelihood functions being the probability density of the observed data given the ground truth. For the application of stereo correspondence (and related to that motion estimation) several models have been proposed for the development of the likelihood function, but none of them consider the situation of uncertain gains and offsets. In this paper, we introduce a new likelihood function in which the unknown texture, and the uncertainties of gains and offsets are explicitly modelled.

The solution of stereo correspondence is often represented by a disparity map. The disparity is the difference in position between two corresponding points. In the classical approach, the disparity map is estimated point by point on an individual base. Better results are obtained by raising additional constraints in the solution space. For instance, neighbouring disparities should be smooth (except on the edge of an occlusion), unique, and properly ordered. Context-dependent approaches, such as dynamic programming (Cox et al., 1996) and graph-cut algorithms (Roy and Cox, 1998), embed these contextual constraints by raising an optimization criterion that concerns a group of disparities at once, rather than individual disparities. For that purpose, an optimization criterion is defined that expresses both the compliance of a solution with the constraints, and the degree of agreement with the observed image data.

The Bayesian approach has proved to be a sound base to formulate the optimization problem on (Cox et al., 1996; Belhumeur, 1996). Here, the optimization criterion is expressed in terms of probability densities. A crucial role is the likelihood function, i.e. the conditional probability density of the data given the disparities. Suppose that a given point has a disparity x, and that for that particular point and disparity the

van der Heijden F., J. Spreeuwers L. and Damjanovic S. (2009).
A NEW LIKELIHOOD FUNCTION FOR STEREO MATCHING - How to Achieve Invariance to Unknown Texture, Gains and Offsets?.
In *Proceedings of the Fourth International Conference on Computer Vision Theory and Applications*, pages 603-608
DOI: 10.5220/0001793606030608

pixels in the corresponding image windows are given by $\mathbf{z}_1$ and $\mathbf{z}_2$. Then the likelihood function of that point is by definition the pdf $p(\mathbf{z}_1, \mathbf{z}_2|x)$.

The usual expression for this likelihood is again based on the CBA, and assumes Gaussian, additive white noise. Application of this model leads to the following likelihood:

$$p(\mathbf{z}_1, \mathbf{z}_2|x) \propto \exp\left(-\frac{1}{4\sigma_n^2}\|\mathbf{z}_1 - \mathbf{z}_2\|^2\right) \qquad (1)$$

Here, $\|\mathbf{z}_1 - \mathbf{z}_2\|^2$ is the SSD. The likelihood function in eq. (1) is a monotonically decreasing function of the SSD. It is used by (Cox et al., 1996) and (Belhumeur, 1996) albeit that both have an additional provision for occluded pixels. However, the function is inappropriate if the gains and offsets are uncertain. Yet, the differences between the grey levels in two corresponding windows is often more affected by differences in gains and offsets than by noise. This paper introduces new expressions which do include these effects. The NCC and the mutual information similarity measures are also invariant to these nuisance factors. However, these measures are parameters derived from the pdfs. But in a true probabilistic approach we really need the pdfs themselves, and not just parameters.

The paper is organized as follows. Section 2 introduces the new likelihood function. Here, a probabilistic model is formulated that explicitly describes the existence of an unknown texture, and uncertain gains and offsets. The final likelihood is obtained by marginalization of these factors. Section 3 analyses the expression that is found for the likelihood. In Section 4, we present some experimental results where the likelihood function is used within a HMM framework. A comparison is made between the newly derived likelihood and the NCC when used in a forward/backward algorithm. Section 5 gives concluding remarks and further directions.

2 THE LIKELIHOOD OF TWO CORRESPONDING POINTS

We consider two corresponding points with disparity x. The image data within two windows that surround the two points are represented by $\mathbf{z}_1$ and $\mathbf{z}_2$. The grey levels (or colours) within the windows depend on the texture and radiometric properties of the observed surface patch, but also on the illumination of the surface, and on the properties of the imaging device. We model this by:

$$\mathbf{z}_k = \alpha_k \mathbf{s} + \mathbf{n}_k + \beta_k \mathbf{e} \qquad k = 1,2 \qquad (2)$$

Here, $\mathbf{s}$ is the result of mapping the texture on the surface to the two image planes. According to the CBA, this mapping yields identical results in the two images. α_k are the gain factors of the two imaging devices. β_k are the offsets. $\mathbf{e}$ is the all 1 vector. $\mathbf{n}_k$ are noise vectors. We assume Gaussian noise with covariance matrix $\mathbf{C_n}$. Furthermore, we assume that $\mathbf{n}_1$ is not correlated with $\mathbf{n}_2$.

Strictly speaking, the CBA can only hold for fronto-parallel planar surface patches. In all other cases the local geometry of the surface around a point of interest is mapped differently to the two image planes. Thus, the texture on the surface will be observed differently in the images. This problem becomes more distinct as the size of the window increases. The problem can be solved by backmapping the image data within the two windows to the 3D surface before applying the similarity measure, (Spreeuwers, 2008). In the sequel, we will assume that either such a geometric correction has taken place, or that the windows are so small that the aperture problem can be neglected.

In order to get the expression for the likelihood function we marginalize the pdf of $\mathbf{z}_1$ and $\mathbf{z}_2$ with respect to the unknown texture $\mathbf{s}$. Next, we marginalize the resulting expression with respect to the gains α_k. The offsets can be dealt with by regarding $\mathbf{n}_k + \beta_k \mathbf{e}$ as one additive noise term. Thus, a redefinition of $\mathbf{C_n}$ suffices. This will be looked upon in more detail in Section 2.3, but for the moment we can ignore the existence of offsets.

2.1 Texture Marginalization

The likelihood function can be obtained by marginalization of the texture:

$$p(\mathbf{z}_1, \mathbf{z}_2|x, \alpha_1, \alpha_2) = \int_{\mathbf{s}} p(\mathbf{z}_1, \mathbf{z}_2|x, \mathbf{s}, \alpha_1, \alpha_2) p(\mathbf{s}|x) d\mathbf{s} \qquad (3)$$

The pdf $p(\mathbf{s}|x)$ represents the prior pdf of the texture $\mathbf{s}$. For simplicity, we assume a full lack of prior knowledge, thus leading to a prior pdf which is constant within the allowable range of $\mathbf{z}_1$ and $\mathbf{z}_2$. This justifies the following simplification:

$$p(\mathbf{z}_1, \mathbf{z}_2|x) = K \int_{\mathbf{s}} p(\mathbf{z}_1, \mathbf{z}_2|x, \mathbf{s}, \alpha_1, \alpha_2) d\mathbf{s} \qquad (4)$$

K is a normalization constant that depends on the width of $p(\mathbf{s})$. Any width will do as long as $p(\mathbf{s})$ covers the range of interest of $\mathbf{z}_1$ and $\mathbf{z}_2$. Therefore, K is undetermined. This is not really a limitation since K does not depend on x, $\mathbf{z}_1$ or $\mathbf{z}_2$.

With $\mathbf{s}$ fixed, $\mathbf{z_1}$ and $\mathbf{z_2}$ are two uncorrelated, normal distributed random vectors with mean $\mathbf{s}$, and co-

variance matrix $\mathbf{C_n}$. Therefore $p(\mathbf{z}_1,\mathbf{z}_2|x,\alpha_1,\alpha_2) = G(\mathbf{z}_1-\alpha_1\mathbf{s})G(\mathbf{z}_2-\alpha_2\mathbf{s})$, where $G(\cdot)$ is a Gaussian distribution with zero mean and covariance matrix $\mathbf{C_n}$. This expression can be further simplified by the introduction of two auxiliary variables: $\mathbf{h} \equiv \frac{\mathbf{z}_1}{\alpha_1} - \mathbf{s}$ and $\mathbf{y} \equiv \frac{\mathbf{z}_1}{\alpha_1} - \frac{\mathbf{z}_2}{\alpha_2}$ so that $\mathbf{h}-\mathbf{y} = \frac{\mathbf{z}_2}{\alpha_2} - \mathbf{s}$. The likelihood function can be obtained by substitution:

$$p(\mathbf{z_1},\mathbf{z_2}|x,\alpha_1,\alpha_2) = K\int_{\mathbf{h}} G(\alpha_1\mathbf{h})G(\alpha_2(\mathbf{h}-\mathbf{y}))d\mathbf{h}$$

and by rewriting this in the Gaussian form:

$$p(\mathbf{z}_1,\mathbf{z}_2|x,\alpha_1,\alpha_2) \propto \frac{1}{\sqrt{\alpha_1^2+\alpha_2^2}}\exp\left(-\frac{(\alpha_2\mathbf{z}_1-\alpha_1\mathbf{z}_2)^{\mathrm{T}}\mathbf{C_n}^{-1}(\alpha_2\mathbf{z}_1-\alpha_1\mathbf{z}_2)}{2(\alpha_1^2+\alpha_2^2)}\right) \quad (5)$$

Note that for $\alpha_1=\alpha_2=1$ and $\mathbf{C_n}=\sigma_n^2\mathbf{I}$ the likelihood simplifies to eq. (1). The resulting likelihood function is the same as in (Cox et al., 1996; Belhumeur, 1996) although the models on which the expression is based differ.

2.2 Marginalization of the Gains

In order to neutralize the unknown gains we marginalize over α_1 and α_2:

$$p(\mathbf{z}_1,\mathbf{z}_2|x) = \int_{\alpha_1}\int_{\alpha_2} p(\mathbf{z}_1,\mathbf{z}_2|x,\alpha_1,\alpha_2)p(\alpha_1)p(\alpha_2)d\alpha_2 d\alpha_1 \quad (6)$$

The prior pdfs $p(\alpha_k)$ should reflect the prior knowledge about the gains α_k. Usually, the gain factors do not deviate too much from 1. For that reason, we chose for $p(\alpha_k)$ a normal distribution, centred around 1, and with standard deviations σ_α. In order to make the analytical integration of eq. (5) possible, we approximate the term $1/(\alpha_1^2+\alpha_2^2)$ by its value at $\alpha_k=1$, that is $\frac{1}{2}$. This approximation is rough, but not too rough. For $\alpha_k<1$, the factor $1/(\alpha_1^2+\alpha_2^2)$ is underestimated, but for $\alpha_k>1$ it is overestimated. Since the integration takes place on both side of $\alpha_k=1$, the error is partly compensated for.

Under the assumption $\alpha_k \sim N(1,\sigma_\alpha)$, the approximation leads to the following result:

$$p(\mathbf{z}_1,\mathbf{z}_2|x) \propto \frac{\exp\left(-\frac{\sigma_\alpha^2(\rho_{11}\rho_{22}-\rho_{12}^2)+\rho_{11}+\rho_{22}-2\rho_{12}}{\sigma_\alpha^4(\rho_{11}\rho_{22}-\rho_{12}^2)+2\sigma_\alpha^2(\rho_{11}+\rho_{22})+4}\right)}{\sqrt{\sigma_\alpha^4(\rho_{11}\rho_{22}-\rho_{12}^2)+2\sigma_\alpha^2(\rho_{11}+\rho_{22})+4}} \quad (7)$$

where:

$$\rho_{k\ell} = \mathbf{z}_k^T\mathbf{C_n}^{-1}\mathbf{z}_\ell \quad \text{with:} \quad k,\ell=1,2 \quad (8)$$

In the limiting case, as $\sigma_\alpha \to 0$, we have

$$p(\mathbf{z}_1,\mathbf{z}_2|x) \propto \exp\left(-\tfrac{1}{4}(\rho_{11}+\rho_{22}-2\rho_{12})\right) \quad (9)$$

which coincides with eq. (1). Intuitively, this is correct since the uncertainties about α_1 and α_2 is zero then. In the other limiting case, as $\sigma_\alpha \to \infty$, the likelihood becomes:

$$p(\mathbf{z}_1,\mathbf{z}_2|x) \propto \frac{1}{\sqrt{\rho_{11}\rho_{22}-\rho_{12}^2}} \quad (10)$$

We will analyse these expressions further in Section 3.

2.3 Neutralizing the Unknown Offsets

We assume that the offsets β_k have a normal distribution with zero mean, and standard deviation σ_β. The vectors $\beta_k\mathbf{e}$ have a covariance matrix $\sigma_\beta^2\mathbf{e}\mathbf{e}^T$. Since the random vectors are additive, we may absorb them in the noise vectors $\mathbf{n}_k$. Effectively this implies that the covariance matrix $\mathbf{C_n}$ now becomes $\mathbf{C_n}+\sigma_\beta^2\mathbf{e}\mathbf{e}^T$. Consequently, the variable $\rho_{k\ell}$ in eq. (8) should be redefined by $\rho_{k\ell} = \mathbf{z}_k^T(\mathbf{C_n}+\sigma_\beta^2\mathbf{e}\mathbf{e}^T)^{-1}\mathbf{z}_\ell$ This can be rewritten in:

$$\rho_{k\ell} = \mathbf{z}_k^T\left(\sum_{n=1}^{N}\mathbf{v}_n\lambda_n^{-1}\mathbf{v}_n^T\right)\mathbf{z}_\ell \quad (11)$$

λ_n are the eigenvalues of the covariance matrix. $\mathbf{v}_n$ are the corresponding eigenvectors. Suppose that $N\sigma_\beta^2$ is large relative to all other eigenvalues of $\mathbf{C_n}$ (N is the dimension of $\mathbf{z}_k$). In case of white noise, the equivalent assumption is $N\sigma_\beta^2 \gg \sigma_n^2$). Then one of the eigenvalues of $\mathbf{C_n}+\sigma_\beta^2\mathbf{e}\mathbf{e}^{\mathbf{T}}$ is close to $N\sigma_\beta^2$, while all other eigenvalues are considerably smaller. The eigenvector that corresponds to σ_β^2 is close to $\mathbf{e}$. The contribution of this particular eigenvalue/eigenvector to $\rho_{k\ell}$ in eq. (11) is about:

$$\frac{\mathbf{z}_k^T\mathbf{e}\mathbf{e}^T\mathbf{z}_\ell}{N\sigma_\beta^2}. \quad (12)$$

The limit case, $\sigma_\beta \to \infty$, represents the situation of full lack of prior knowledge of the offsets. In this circumstance, the approximations above become exact. Thus, the full contribution in eq. (12) becomes zero.

There is no need to embed $\sigma_\beta^2\mathbf{e}\mathbf{e}^{\mathbf{T}}$ in $\mathbf{C_n}$. The factor $\mathbf{z}_k^T\mathbf{e}$ is the projection of $\mathbf{z}_k$ on $\mathbf{e}$. We just need to remove this projection from $\mathbf{z}_k$ beforehand, and then its contribution is zero anyhow. This can be obtained by subtracting the average of the elements of the vector. Thus, if $\overline{z}_k$ is the average of the elements of the vector $\mathbf{z}_k$, then:

$$\rho_{k\ell} = (\mathbf{z}_k-\overline{z}_k\mathbf{e})^T\mathbf{C_n}^{-1}(\mathbf{z}_\ell-\overline{z}_\ell\mathbf{e}) \quad (13)$$

Note that this approach to cope with unknown offsets is equivalent to the normalization of the mean, just as in the NCC procedure.

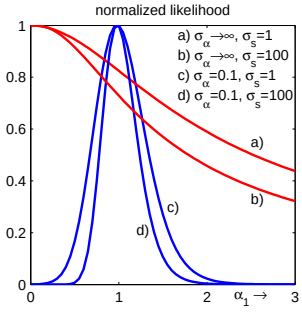


Figure 1: The likelihood function with varying α_1. Other parameters are: $\alpha_2 = 1$, $\sigma_n = 1$, $N = 225$

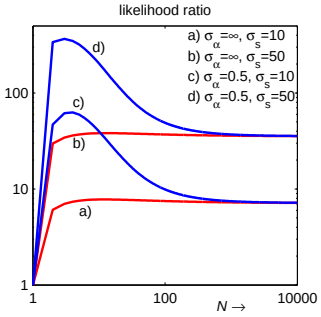


Figure 2: The likelihood ratio with varying N. Other parameters are: $\alpha_1 = 1$, $\alpha_2 = 1$, $\sigma_n = 1$

3 LIKELIHOOD ANALYSIS

In this section, we examine the behaviour of the proposed likelihood in different circumstances. For simplicity, we consider only the white noise case, $\mathbf{C_n} = \sigma_n^2\mathbf{I}$. First we examine the behaviour of the likelihood function under the null hypothesis with varying α_1. Other parameters are kept constant. Substitution of eq. (2) in eq. (8) yields:

$$\begin{aligned}\rho_{kk} &= (\alpha_k^2\mathbf{s}^T\mathbf{s} + 2\alpha_k\mathbf{s}^T\mathbf{n}_k + \mathbf{n}_k^T\mathbf{n}_k)/\sigma_n^2 \\ \rho_{12} &= (\alpha_1\alpha_2\mathbf{s}^T\mathbf{s} + \alpha_1\mathbf{s}^T\mathbf{n}_1 + \alpha_2\mathbf{s}^T\mathbf{n}_2 + \mathbf{n}_1^T\mathbf{n}_2)/\sigma_n^2\end{aligned} \tag{14}$$

We regard $\mathbf{s}$ as a nonrandom signal. The energy σ_s is defined as $\sigma_s^2 \equiv \mathbf{s}^T\mathbf{s}/N$. We examine the behaviour by replacing the inner products in eq. (14) by their root mean squares. That is:

$$\begin{aligned}\mathbf{s}^T\mathbf{n}_k &\sim \sqrt{\mathrm{E}\left[\left(\mathbf{s}^T\mathbf{n}_k\right)^2\right]} = \sqrt{N}\sigma_s\sigma_n \\ \mathbf{n}_k^T\mathbf{n}_k &\sim \sqrt{\mathrm{E}\left[\left(\mathbf{n}_k^T\mathbf{n}_k\right)^2\right]} = \sqrt{N^2+2N}\sigma_n^2 \\ \mathbf{n}_1^T\mathbf{n}_2 &\sim \sqrt{\mathrm{E}\left[\left(\mathbf{n}_1^T\mathbf{n}_2\right)^2\right]} = \sqrt{N}\sigma_n^2\end{aligned} \tag{15}$$

Figure 1 shows the likelihood function $p(\mathbf{z}_1, \mathbf{z}_2|x)$ for $\sigma_\alpha = \infty$, conform eq. (10), and for $\sigma_\alpha = 0.1$, conform eq. (7) for varying α_1. Of course, a substitution by RMSs is not exact, but nevertheless, the resulting figure gives a good impression of the behaviour. As expected, if σ_α is very large, the likelihood function covers a wide range of α_1. If σ_α is small, then the function is narrowly peaked around $\alpha_1 = 1$.

In order to check whether the new likelihood function is able to distinguish between similar textures and dissimilar textures, we also examined the ratio of the likelihood function under these two different cases. For that purpose, we also considered the alternative model:

$$\mathbf{z}_k = \alpha_k\mathbf{s}_k + \mathbf{n}_k + \beta_k\mathbf{e} \qquad k = 1,2 \tag{16}$$

In this situation, $\mathbf{s}_1$ and $\mathbf{s}_2$ are two different textures, but with the same signal energy σ_s. If we model $\mathbf{s}_1$ and $\mathbf{s}_2$ as realizations from two independent random signals, then $\mathrm{E}[(\mathbf{s}_1^T\mathbf{s}_2)^2]^{1/2} = \sigma_s^2\sqrt{N}$. Thus, if the textures are dissimilar, the RMS of the factor $\mathbf{s}^T\mathbf{s}$ in ρ_{12} in eq. (14) should be replaced accordingly. The ratio between the likelihoods in the two cases is:

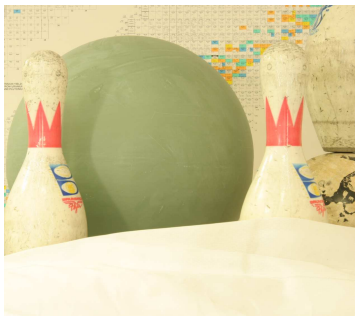
$$\Lambda(\mathbf{z}_1, \mathbf{z}_2) \equiv \frac{p(\mathbf{z}_1, \mathbf{z}_2|x, \text{similar textures})}{p(\mathbf{z}_1, \mathbf{z}_2|x, \text{dissimilar textures})} \tag{17}$$

Figure 2 shows this ratio for varying N. We see that the ratio's with $\sigma_\alpha = 0.5$ are always larger than the one with $\sigma_\alpha = \infty$, but for large N the ratio's with $\sigma_\alpha = 0.5$ approaches the other one and becomes constant on the long run. The reason for this typical behaviour is that in the factor $\rho_{11}\rho_{22} - \rho_{12}^2$ the contribution of the signal $\alpha_1\alpha_2\mathbf{s}^T\mathbf{s}$ is cancelled out, while the contribution of the noise, i.e. $\mathbf{n}_k^T\mathbf{n}_k$, is proportional to N, and thus keeps growing as N increases.

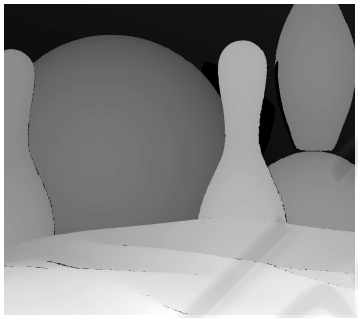
4 EXPERIMENTS

A preliminary experiment is conducted to demonstrate the abilities of our newly proposed likelihood. For that purpose, two rectified stereo images were selected. See Figure 3. In order to embed the likelihood function within a probabilistic framework, we treat the stereo correspondence along the epipolar line as a state estimation problem using a HMM (Hidden Markov Model). The reconstruction is done using the FwBw (forward-backward) algorithm (van der Heijden et al., 2004). The Viterbi algorithm is also applicable, but in our experiments, FwBw outperformed Viterbi. We calculated the disparity map using the new likelihood function as the observation probability, and compared this map with a map obtained from the same HMM, but with an other likelihood function plugged in.

(a) right image



(b) left image



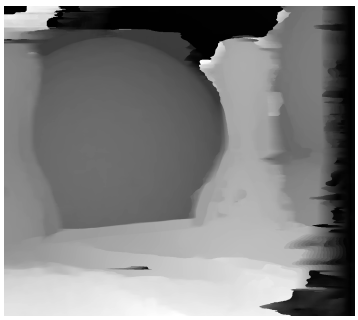
(c) disparity map of right image

Figure 3: Two stereo images and the corresponding reference disparity map.

4.1 The Hidden Markov Model

Each row of the left image is considered as a HMM. Thus, the running variable is the row index i. The state variable of the HMM is taken to be the disparity x_i. The set of allowable states is: $x_i \in \Omega = \{K_{min}, ..., K_{max}\}$. K_{min} and K_{max} are the minimum and the maximum disparities between the two images. The number of different states is $K = K_{max} - K_{min} + 1$. The transition probabilities between consecutive states are given by the transition probability $P_t(x_{i+1} = n | x_i = m)$.

The disparity x_i along an image row is a piecewise continuous function of i. Sudden jumps are caused by occlusions and boundaries between adjacent objects of different depth in the scene, but for the remaining part the depth tends to be smooth. We can model this prior knowledge by selecting $P_t(x_{i+1} = n | x_i = m)$ such that the next state x_{i+1} is likely to be close to the current state x_i. The variable has the highest probability to stay in the same state. The probability should decrease as the absolute difference $\Delta \equiv |x_{i+1} - x_i|$ increases. However, the probability should also allow the large jumps that are caused by occlusions and object boundaries.



(a) the new likelihood function

(b) NCC based likelihood

Figure 4: Reconstructed depth maps.

In our experiments, P_t consists of two modes. Large jumps are modelled with an overall probability $P_{outlier}$ uniformly distributed over the range $x_i - J_{max}, \cdots, x_i + J_{max}$. In this mode, each state within this range is reached with a probability $P_{outlier}/(2J_{max} + 1)$. Inliers are modelled with an overall probability of $1 - P_{outlier}$. Here, the transition probability linearly decreases with Δ up to where Δ is larger than a threshold T_{max}. We chose $P_{outlier} = 0.05$, $J_{max} = 8$, and $T_{max} = 3$. Note, however, that the choice of P_t could be refined by, for instance, using the uniqueness constraint on the disparities (Faugeras, 1993).

4.2 Reconstruction

The selected rectified stereo pair is shown in Figure 3. These images are taken from (Hirschmller and Scharstein, 2007). The scene, 'bowling1', is chosen because our intention is to apply the algorithm for the reconstruction of textureless and smooth surfaces

so that later the application can be extended to the 3D reconstruction of faces. The minimum and maximum disparities of these images are $K_{min} = 374$ and $K_{max} = 446$, which means that the state-space model has $K = 73$ states.

The reconstruction is done by applying the forward-backward algorithm to an HMM with the transition probability described above and with the observation probability given by eq. (7). For the calculation of the likelihood expression, we consider that the noise variance is $\sigma_n^2 = 0.05$, the gain variances $\sigma_{\alpha 1}^2 = \sigma_{\alpha 1}^2 = 0.25$. We performed the calculations on the pixels within 31x31 windows. Thus, $N = 961$.

The reconstruction is also performed using the NCC as similarity measure. Since this measure is not a probability density, it possibly should undergo a rescaling to make it more suitable for a substitute of the observation probability. After some experimentation, we found that the following mapping of the NCC

$$\left(\tfrac{1}{2}(1+NCC)\right)^{\gamma} \tag{18}$$

is a suitable choice. The best reconstruction was obtained with $\gamma = 6$. We applied this expression within a HMM with the transition probability described above. The windows that were used are also 31x31.

4.3 Results

The reconstructed disparity maps are shown in Figure 4. A comparison with the ground truth (Figure 3) shows that the reconstruction based on the new likelihood function is more accurate and more robust than the one based on the NCC measure. The new likelihood expression is better able to deal with, especially, the steplike transitions due to occlusion. The NCC-based result is oversmoothed, and cannot locate this transitions accurately. Note that the large error on the right-hand side of the disparity maps are caused by missing data in the left image.

5 CONCLUSIONS

We have found an expression for a likelihood function that can cope with unknown textures, uncertain gain factors and uncertain offsets. In contrast to the classical approaches this likelihood is not based on some arbitrary selected heuristics, but on a sound probabilistic model. As such it can be used within a probabilistic framework. The likelihood can be fine-tuned by setting a limited range of allowable gain factors rather than just any gain factor.

Using the model we were able to show that coping with unknown offsets can safely be done by normalizing the means of the data, as done in other approaches such as the normalized correlation coefficient. Unknown gain factors and unknown textures are dealt with in a way that differs a lot from other approaches. Yet, the computational complexity of the proposed metric is quite comparable with, for instance, the computational load of the NCC.

We demonstrated stereo reconstruction within the probabilistic framework by the forward-backward algorithm with a suitably chosen HMM and showed that it is a resourceful approach. We showed that the newly proposed likelihood is more suitable for stereo reconstruction within the probabilistic framework than the NCC. The reconstruction using the new likelihood deals better with occlusion, while the NCC tends to oversmooth the area with greater abrupt change in depth.

REFERENCES

Belhumeur, P. N. (1996). A bayesian approach to binocular stereopsis. *Int. J. Comput. Vision*, 19(3):237–260.

Cox, I. J., Hingorani, S. L., Rao, S. B., and Maggs, B. M. (1996). A maximum likelihood stereo algorithm. *Comput. Vis. Image Underst.*, 63(3):542–567.

Egnal, G. (2000). mutual information as a stereo correspondence measure. Technical Report Technical Report MS-CIS-00-20, Comp. and Inf. Science, U. of Pennsylvania.

Faugeras, O. (1993). *Three-Dimensional Computer Vision - A Geometric Viewpoint*. The MIT Press.

Hirschmller, H. and Scharstein, D. (2007). Evaluation of cost functions for stereo matching. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007)*. http://vision.middlebury.edu/stereo/data/.

Roy, S. and Cox, I. J. (1998). A maximum-flow formulation of the n-camera stereo correspondence problem. In *ICCV '98: Proceedings of the Sixth International Conference on Computer Vision*, page 492, Washington, DC, USA. IEEE Computer Society.

Spreeuwers, L. (2008). Multi-view passive acquisition device for 3d face recognition. In *BIOSIG 2008: Biometrik und elektronische Signaturen*.

van der Heijden, F., Duin, R., de Ridder, D., and Tax, D. (2004). *Classification, Parameter Estimation and State Estimation An Engineering Approach Using MATLAB*. WILEY.