Multiresolution reliability scheme for range image filtering

Thierry Zamofing and Heinz Hügli

University of Neuchâtel - Institute of Microtechnology, 2000 Neuchâtel, Switzerland

ABSTRACT

A large number of 3D cameras suffer from so-called holes in the data, i.e. the measurement lattice is affected by invalid measurements and the range image has undefined values. Conventional image filters used for removing the holes perform not well in presence of holes with large varying hole sizes. The novel hole-filling method presented in this paper operates on reliability attributed range images featuring unwanted holes with wide varying sizes. The method operates according to a multi resolution scheme where the image resolution is decreased at the same time as the range reliability is successively increased until sufficient confidence is reached. It builds on three main components. First, the described process performs a weighted local neighbourhood filter where the contribution of each pixel stands for its reliability. Second, the filtering combines filters with different kernel sizes and implements therefore the multi resolution schema. Third, the processing requires a complete travel from high resolution down to the resolution of satisfactory confidence and back again to the highest resolution. The algorithm for the described method was implemented in a efficient way and was widely applied in the hole-filling of range images from a depth from focus process where reliability is obtainable non-linearly from the local sharpness measurement. The method is valid in a very general way for all range imagers providing reliability information. It seems therefore well suited to depth cameras like time-of-flight, stereo and other similar rangers.

Keywords: Range images, depth images, multi resolution scheme, reliability based filtering, range interpolation, range image filtering, hole-filling, 3D data filtering

1. INTRODUCTION

A large number of 3D cameras suffer from so-called holes in the data, i.e. the measurement lattice is affected by invalid measurements and the range image has undefined values. Assuming signal continuity, image filters can be used to suppress the holes. Various solutions have been proposed. Median filter,[1] weighted median filter[2] and weighted myriad filters[3] have been widely used to remove undesired outliers from an image. Such filters often do a better job than the mean filter because they preserve useful details in the image. A median filter is therefore an appropriate solution to remove small holes from an image without affecting the image structure. Unfortunately it leads to bad results if the size of holes is too large. In fact it cannot interpolate regions, when only few measured values are known. "It can't fill large holes.". Further when the filter is limited to a fixed kernel size for the whole image it is always too small for some large hole and too large for regions without holes. Therefore a novel filtering algorithm has been developed, which considers the data with associated reliability information and uses filters with different kernel sizes. Note that various camera and acquisition systems generate such images with associated reliability. The reliability value associated to each pixel informs if the pixel value is precise or not.

Section 2 gives a general description of the filter which can be implemented directly as in Section 3 or according to a multi resolution schema as in Section 4. An application is provided in Section 5.

2. FILTER USING RELIABILITY INFORMATION

The idea of the filter is to use the reliability information as a weight in the filtering process. Given the measurement values $V(x,y)$ and the associated reliability information $W(x,y)$ as input, the task is to develop a filter that estimates the real measurement values $\widehat{V}(x,y)$. The filter assumes that the measurement values of neighbouring values are similar and produces as a byproduct a reliability value $\widehat{W}(x,y)$ associated to the filtered values.

In practice, V is typically a range image and W the associated reliability as delivered for example by the camera. Note that a reliability attributed measurement can be considered as generalisation of a hole affected data. A hole affected data can be simply represented by a measurement and associated reliability value set to a zero.
The proposed filtering assumes that measurement points with a high reliability should contribute more than those with a low reliability to the final measurement value. Further the locality of the point should also be taken into consideration, so that distant points have a lower influence than near points. The filter used for this task is called *weighted local neighbourhood filter* and is described in section 3.
The size of the neighbourhood is a problem. Small size for high reliability and large size for low one are desired. Therefore we propose to use an adaptive size. The idea is to increase the size until a given reliability is reached. In order to perform this idea in a efficient way, we propose the multi resolution approach described in section 4.

3. WEIGHTED LOCAL NEIGHBOURHOOD FILTER

Assuming similar V values for neighbouring pixels, the idea is to perform a local spatial filter where each pixel contributes in a way proportional to its weight value.
Let $W(x,y)$ be the weight assigned to the measured values $V(x,y)$. Thus a good estimation $\widehat{V}$ for the real value V is the *uniform weighted average on local neighbourhood* filter(eq. 1). The neighbourhood is typically a square:

$$\widehat{V}(x_0,y_0) = \left.\frac{\sum W(x,y)\cdot V(x,y)}{\sum W(x,y)}\right| x,y \in \text{Neighbourhood}(x_0,y_0) \tag{1}$$

An alternative filter assumes that the values contribute in a different way depending on their distance from the center (x_0,y_0). In fact values near the center (x_0,y_0) should be taken more into account than distant values. This can be achieved by multiplying the values with a filter that also considerers the locality.

Assuming the locality is defined in term of a pseudo 2-dimensional shaped Gaussian function $= G(x,y)$ centered on (x_0,y_0) with a kernel size of $K \times K$:

$$G(x,y) = \begin{cases} \sqrt{2\pi}\sigma \cdot e^{\frac{-(x^2+y^2)}{2\sigma^2}}, & |x| < K \text{ and } |y| < K; \\ 0, & |x| \geq K \text{ or } |y| \geq K. \end{cases} \tag{2}$$

we obtain the *Gaussian weighted average on local neighbourhood* filter:

$$\widehat{V}(x_0,y_0) = \frac{\Big(W(x,y)\cdot V(x,y)\Big) **G(x,y)}{W(x,y)**G(x,y)} \tag{3}$$

To speed up the computation, the separability of $G(x,y)$ can be used. The 2 dimensional bounded Gaussian filter G(x,y) can be calculated as a convolution of two one dimensional Gaussian filter g(x):

$$g(x) = \begin{cases} \sqrt{2\pi}\sigma \cdot e^{\frac{-(x^2)}{2\sigma^2}}, & |x| < K; \\ 0, & |x| \geq K. \end{cases} \tag{4}$$

$$G(x,y) = \sqrt{2\pi}\sigma \cdot e^{\frac{-(x^2+y^2)}{2\sigma^2}} = \sqrt{2\pi}\sigma \cdot e^{\frac{-(x^2)}{2\sigma^2}} \cdot e^{\frac{-(y^2)}{2\sigma^2}} = g(x)\cdot g(y) = g^T * g \tag{5}$$

The result $\widehat{I}$ of the image I filtered with a Gaussian can be rewritten as:

$$\widehat{I}(x,y) = G(x,y)**I(x,y) = (g^T * g) * I(x,y) = g^T * (g * I(x,y)) \tag{6}$$

Applying this to equation 3 results in

$$\widehat{V} = \frac{g^T * \left(g * (W \cdot V)\right)}{g^T * (g * W)} \qquad \widehat{W} = g^T * (g * W) \tag{7}$$

which are the estimated values $\widehat{V}$ and the estimated reliabilities $\widehat{W}$.

A good tradeoff for the kernel size must be found. On one hand large kernels with large σ are desired on regions with large holes and low reliability. On the other hand small kernels with small σ are desired on regions with high reliability were the structure of the image should not be affected by the filter. This can be obtained with following multi resolution filter.

4. MULTI RESOLUTION FILTER

The basic idea is to compute the *weighted local neighbourhood filter* of each image pair in a multi resolution scheme that averages the values and their reliabilities on progressively larger neighbourhoods. The estimated measurement value is then the measurement value of the level which maximises the reliability.

A multi resolution pyramid is a stack of the same image with different resolutions like depicted in figure 1. The different levels are created by successively down sampling the image by a factor 2.

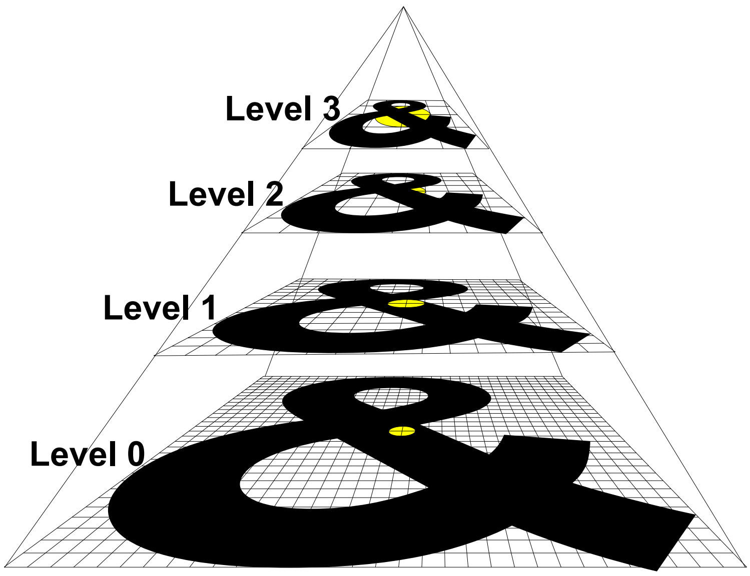


Figure 1. multi resolution pyramid

In the present case, there are two input images for which a multi resolution pyramid must be build. For the image W a *normal pyramid* is build and for the image V a *weighted pyramid* is build. After creation of the weight and measurement value pyramid, the image are up sampled and recombined to the final, filtered result image.

4.1. Down Sampling

Down sampling reduces the resolution of a discrete signal.[4,5] Down sampling consists in decimation of the lowpass filtered signal. The used symbol for decimation that decreased the samples by a factor of L is $\boxed{\downarrow\mathrm{L}}$. Formally, down sampling the image I with an anti aliasing filter F is expressed by:

$$I_d = \boxed{\downarrow\mathrm{L}}(F * * I)$$

The incident down sampling of the image pyramid decimates the image with a factor $L = 2$ in the x and y dimension and uses a 3×3 Gauss kernel as anti alias filter.

4.1.1. Down Sampling of W

Normal down sampling is used to create an image pyramid of the weight image W. The used anti alias filter G ia a 3×3 Gauss filter. Thus the down sampled image W_{i+1} of the input image W_i is:

$$\begin{aligned} W_{i+1} &= \boxed{\downarrow 2}(W_i ** G) \\ \text{with: } G(x,y) &= \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} \cdot \frac{1}{16} \end{aligned} \tag{8}$$

Efficient down sampling (direct 3x3 filtering and down sampling) is described in.[6]

4.1.2. Weighted Down Sampling of V

This weighted down sampling process is used to create an image pyramid of the measurement image V. It uses the *3x3 Gaussian weighted spatial filter* as anti aliasing filter (equation 3).
Thus the down sampled image V_{i+1} of the input image V_i is:

$$V_{i+1} = \boxed{\downarrow 2}\left(\frac{(W_i \cdot V_i) ** G}{W_i ** G}\right) \tag{9}$$

4.2. Up Sampling

Up sampling is the inverse operation of down sampling. It is used to create an image with a higher resolution. Up sampling consists in lowpass filtering of the interpolated signal. Interpolation is the process of inserting additional samples between the original low-resolution samples. The used symbol for interpolation that increases the samples by a factor of L is $\boxed{\uparrow \text{L}}$.
The incident up-sampling interpolates the image with a factor $L = 2$ in the x and y dimension and filters the image with a 3x3 Gauss kernel.

4.2.1. Up Sampling of W

The weight image W has been down sampled with a *normal down sampling* and must be up sampled with a *normal up sampling*. After interpolation, the image is filtered with a 3x3 Gaussian kernel. Thus the up sampled image W_{i-1} of the input image W_i is:

$$\begin{aligned} W_{i-1}^u &= (\boxed{\uparrow 2} W_i) ** H \\ \text{with: } H(x,y) &= \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} \cdot \frac{1}{8} \end{aligned} \tag{10}$$

4.2.2. Weighted Up Sampling of V

The measurement image V has been down sampled with a *weighted down sampling* and must be up sampled with a *weighted up sampling*. Thus the up sampled image V_{i-1} of the input image V_i is:

$$V_{i-1}^u = \left(\frac{\boxed{\uparrow 2}(W_i \cdot V_i) ** H}{\boxed{\uparrow 2}(W_i) ** H}\right) \tag{11}$$

4.3. Combination

After up sampling an image pair, there are 2 image pairs $\{W_n,V_n\}$ and $\{W_n^u,V_n^u\}$ with the same resolution, that have to be combined to a new filtered image pair $\{W_n^f,V_n^f\}$.
The difference between the image pair $\{W_n^u,V_n^u\}$ and $\{W_n,V_n\}$ are the following: On one hand the image pair $\{W_n,V_n\}$ has a higher fluctuation of the reliability and therefore also pixels with a lower reliability than $\{W_n^u,V_n^u\}$. On the other hand if the reliability is good enough, $\{W_n,V_n\}$ has more details of the structure than $\{W_n^u,V_n^u\}$. The task is to optimise the structure of the image hence to maximise the reliability. Therefore the *combination* compares the reliability images of these two image pairs and takes the values with the best reliability.
Mathematically the combination $\{W_n^f, V_n^f\}$ of the images $\{W_n,V_n\}$ and $\{W_n^u,V_n^u\}$ takes the measurement and weight values $\{W_n,V_n\}$, if $k_n \cdot W_n > W_n^u$ and otherwise it takes the values $\{W_n^u,V_n^u\}$.

$$\{W_n^f(x,y), V_n^f(x,y)\} = \begin{cases} \{W_n(x,y), V_n(x,y)\}, & k_n \cdot W_n > W_n^u; \\ \{W_n^u(x,y), V_n^u(x,y)\}, & \text{otherwise.} \end{cases} \tag{12}$$

The compare factor k_n introduces a mechanism for tuning the method and adjusting the comparison to peculiarities of the structure in the source images.

4.4. Whole Filtering Process

This section describes the whole filtering process which involves *down sampling* (section 4.1) *up sampling* (section 4.2) and *combination* (section 4.3).

The filter consists of 2 main steps. In a first step the image is filtered and down sampled with normal and weighted down sampling. The second step consists in up sampling and combination of the image pairs. After up sampling of an image pair, there are two image pairs with the same resolution, for example: $\{W_n,\ V_n\}$ and $\{W_n^u,\ V_n^u\}$. The combination process fuses these two image pairs.
The whole process is showed in details in figure 2.

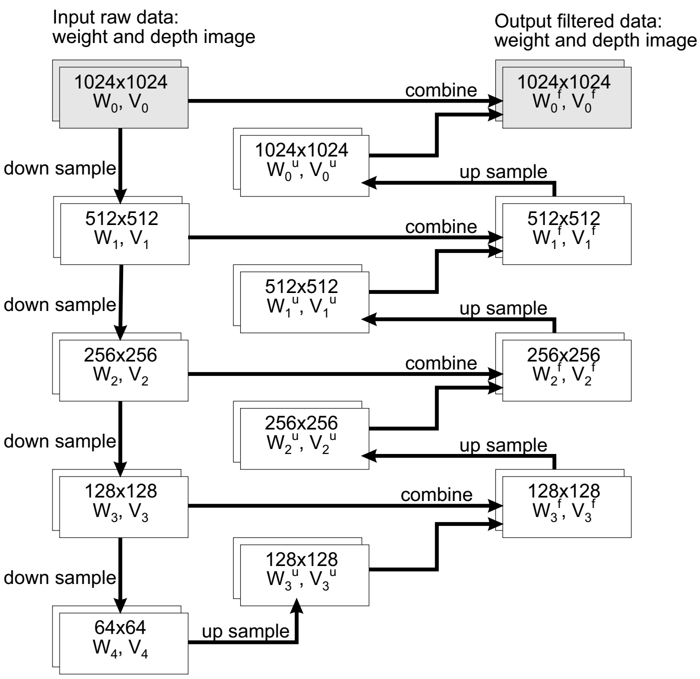
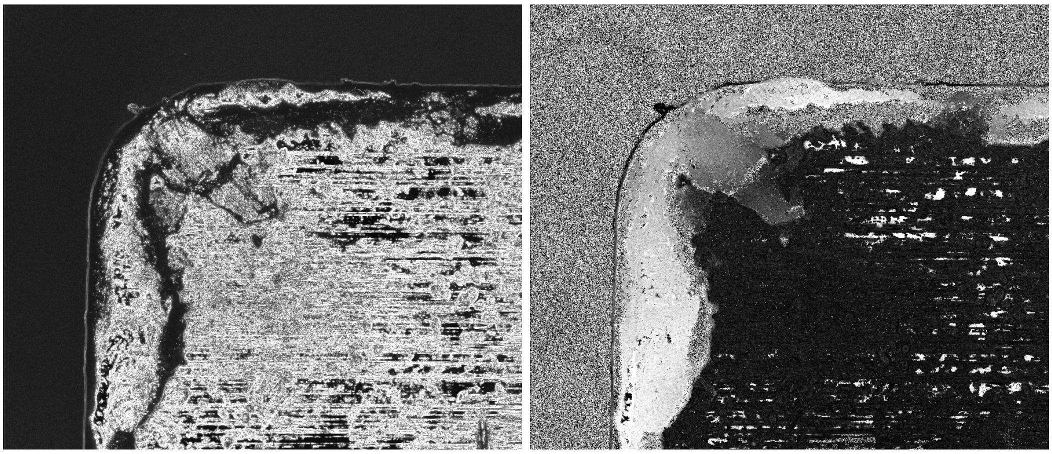


Figure 2. multi resolution filter for 5 levels

5. APPLICATION

This section describes the application of the method to a range image. Figure 3 shows a typical input image pair stemming from a *depth from focus*[7] process. The images show a corner of a punched metal part with a typical burr, the left image is a range image and the right image contains the quality information. A quality value of 0 is equivalent to a hole.



(a) quality values $Q(x, y)$ (b) measurement values $V(x, y)$

Figure 3. input quality values and measurement values

The quality information is the sharpness of the acquisition process and has a nonlinear relation to the reliability. In such cases the first processing step is to map the quality to a weight proportional to its reliability.

$$W_i = f_Q(Q_i) \tag{13}$$

Our application exhibits following behaviour: A small variation on large quality value does not affect enormously the reliability. On the other side the same variation on a small quality value makes a high variation of the reliability. In this case, the quality values is mapped with a exponential function to *weight values.* Equation 14 shows the quality-to-weight-mapping function that was used for *depth from focus.* The function is illustrated in figure 4 using following parameters: $u = 7$, $v = 255$, $r = 0.02$

$$f_Q(q) = \begin{cases} 0, & s \leq u; \\ v \cdot \frac{1-\exp(-(q-u)\cdot r)}{(-\exp(-(v-u)\cdot r)}, & u < s < v; \\ v, & s \geq v. \end{cases} \tag{14}$$

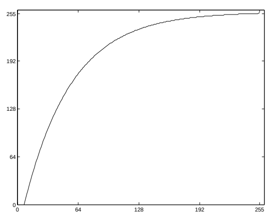


Figure 4. quality to weight mapping function

Figure 5 shows the used comparison factors k_n (eq. 12) to filter images acquired with *depth from focus* method for 8 level pyramid. These values have been selected heuristically for the specification of the application. The factor of level 0 and 1 are low, because these levels have strong reliability fluctuations and therefore a lot of noised values with low reliability. The factor decreases for levels higher than 3 in order to demote somehow the lower frequencies of higher levels.

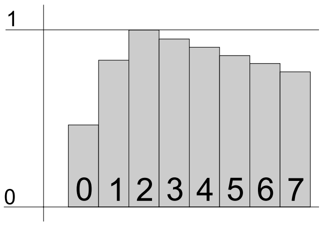


Figure 5. weights k_n for 8 different levels used in *depth from focus*

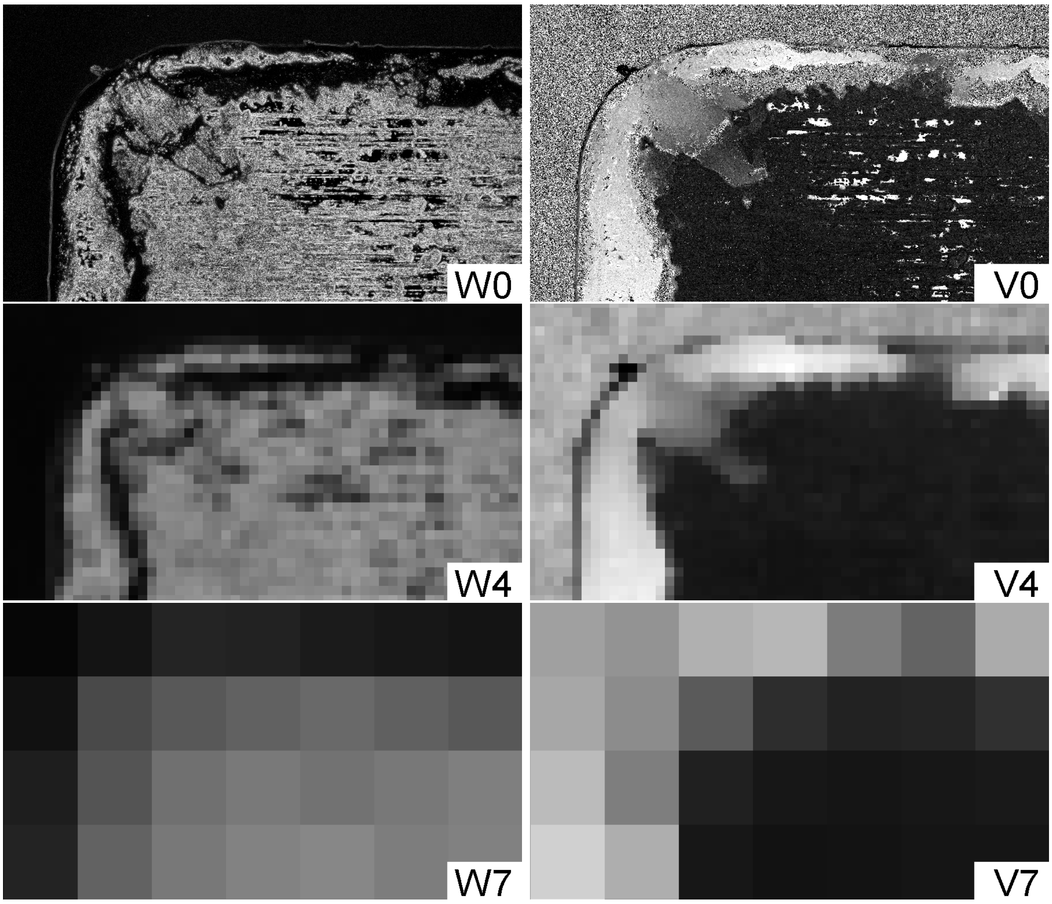
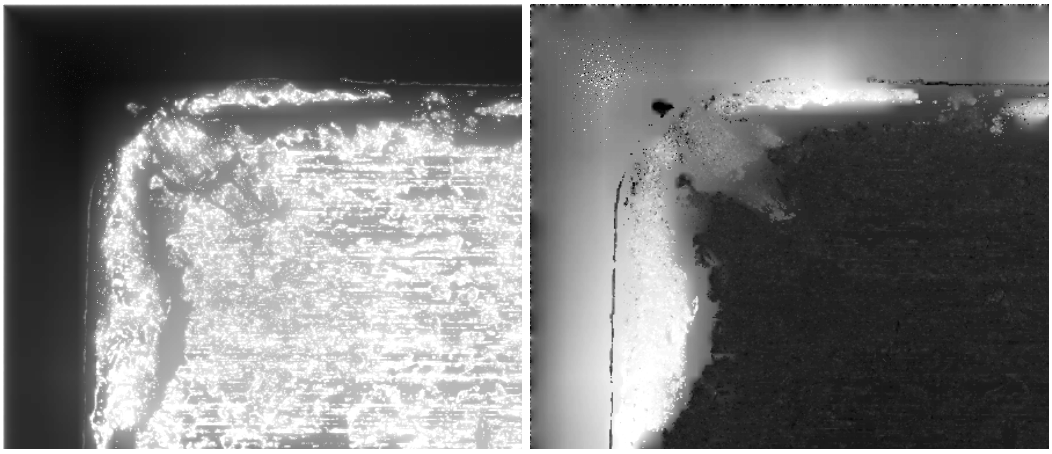


Figure 6. some selected images (W_0, V_0), (W_4, V_4) and (W_7, V_7) for an 8 level pyramid

Figure 6 shows the selected levels of the weighted down sampled image pyramid and figure 7 shows the result of the filtering process. It clearly illustrates the advantage of the method: On one hand regions with a good quality information will keep their details and on the other hand regions with a bad quality information and therefore a lot of noise will be smoothed.

Concerning the computation time, the method is quite efficient, due to the multi scale approach. Non highly optimised C code processes the filter on a Pentium 4, 2.6GHz of a 1024×1024 image with 5 pyramid levels in about 100ms. With code optimisation and use of the processor specific MMX/SSE/SSE2/SSE3 instructions,



(a) reliability values $\widehat{W}(x, y)$ (b) measurement values $\widehat{V}(x, y)$

Figure 7. output estimated reliability values and measurement values

huge additional speedup is possible and speeds below 5ms can be expected.

6. CONCLUSION

The paper presents a filtering method that applies to reliability attributed images and is well suited for the filtering of range images. Specifically, the method solves the hole-filling problem often encountered with range images provided by the various kinds of 3D cameras. Faced with the problem to deal with large varying hole sizes, the method provides a solution by adapting its filter size accordingly.

The method operates according to a multi resolution scheme where the image resolution is successively decreased at the same time as the range reliability is increased until sufficient confidence is reached. The final image is reconstructed from the multiresolution pyramid by choosing the values with best confidence. A main advantage of the multiresolution approach is its computational efficiency, a characteristic benefit of a scheme that reduces the data size instead of increasing the filter size.

The method was applied to several kinds of range images. The presented example stems from a depth from focus camera and is typical of range images with large varying hole sizes. It illustrates the adaptive capability of the method to keep details in regions with high reliability and reconstruct data by weighted filtering elsewhere. Also the excellent computing efficiency is confirmed.

REFERENCES

1. A. Bovic and T. Huang, "The effect of median filtering on edge estimation and detection," **9**, pp. 181–194, IEEE Trans. Pattern Analysis and Machine Intelligence, 1987.
2. D. Brownrigg, "The weighted median filter," **27**, pp. 807–818, Commun. Assoc. Comput. Mach., 1984.
3. S. Kalluri and G. Arce, "Adaptive weighted myriad filter optimization for robust signal processing," CISS, (Princeton, NJ), 1996.
4. DSP-Day, "Ece 4213/5213 digital signal processing day 36: Decimation and interpolation," 2001. URL `www.ecn.ou.edu/vdebrunn/www/ece4213/day36.pdf`.
5. S. Vaseghi, "Changing sampling rate: Decimation and interpolation." URL `www.brunel.ac.uk/depts/ee/COM/Home_Saeed_Vaseghi/Decimation%20and%20Interpolation.pdf`.
6. V. Cantoni and M. Ferretti, *Pyramidal Architectures for Computer Vision*, ch. 2, pp. 35–48. Kluwer Academic/Plenum Pub. Coorp. New York, 1994.
7. T.Zamofing and H.Hügli, "Applied multifocus 3d microscopy," **5265**, pp. 134–144, SPIE Two- and Three-Dimensional Vision Systems for Inspection, 2003.