3D SUPER-RESOLUTION USING GENERALIZED SAMPLING EXPANSION

H. Shekarforoush *M. Berthod* *J. Zerubia*

INRIA - 2004 Route des Lucioles
06902 Sophia Antipolis Cedex
France
email: hshekar@sophia.inria.fr
Fax: (+33) 93 65 76 43

ABSTRACT

A 3D super-resolution algorithm is proposed below, based on a probabilistic interpretation of the n-dimensional version of Papoulis' generalized sampling theorem. The algorithm is devised for recovering the albedo and the height map of a Lambertian surface in a Bayesian framework, using Markov Random Fields for modeling the *a priori* knowledge.

1. INTRODUCTION

Using a set of low resolution images it is possible to reconstruct high resolution information by merging low resolution data on a finer grid. Several approaches have been adopted for low resolution data fusion in 2D cases [6][12]. In [1][11] the idea was extended to 3D visual reconstruction of a Lambertian surface, where it was shown how one can reconstruct high resolution albedo with the knowledge of high resolution height and vice versa. In this paper, we will extend this method to simultaneous reconstruction of the albedo and height.

Using a probabilistic interpretation of n-dimensional version of Papoulis' Generalized Sampling theorem [3][8], one can formulate the problem of 3D super-resolution as an optimization one. A Bayesian framework has been proposed for the formulation of the problem which leads in turn to Markov Random Field (MRF) modeling of the *a priori* knowledge. A generalized simplex algorithm (see [10] for details) has been proposed for optimizing a cost function iteratively. Initial conditions for the optimization have been provided by the low resolution images with subpixel interframe overlap and a low resolution noisy/sparse depth map.

The algorithm can be applied in the area of aerial and satellite image processing.

2. A PROBABILISTIC VIEW OF THE GENERALIZED SAMPLING EXPANSION

Various generalizations of the sampling theorem for non-uniform sampling have been brought together in a single theorem by Papoulis.

We start by recalling the generalized sampling expansion in n dimensions. Unless otherwise mentioned, all functions are n-dimensional vector functions and $f(\vec{t}) \leftrightarrow \mathcal{F}(\vec{\omega})$ denote a Fourier transform pair.

Definition: *A finite energy function $f(\vec{t})$ [9] is said to be σ-band limited if its Fourier transform $\mathcal{F}(\vec{\omega}) = 0$ outside the finite size hypercube $| \omega_i | \geq \sigma_i$, $i = 1 \ldots n$.*

Theorem: *nD Generalised Sampling Theorem*
We apply a σ-band limited function $f(\vec{t})$ as a common input to m independent linear shift invariant systems with transfer functions $\mathcal{H}_1(\vec{\omega}) \ldots \mathcal{H}_m(\vec{\omega})$. The resulting outputs are:

$$\phi_r(\vec{t}) = \frac{1}{(2\pi)^n} \int_{-\sigma_1}^{\sigma_1} \ldots \int_{-\sigma_n}^{\sigma_n} \mathcal{F}(\vec{\omega}) \mathcal{H}_r(\vec{\omega}) exp(j\vec{\omega}^T \vec{t}) d\omega_1 \ldots d\omega_n \quad (1)$$

where $r = 1 \ldots m$, $\omega_1 \ldots \omega_n$ are the components of $\vec{\omega}$ and T denotes the transposition. Next, we sample these outputs at $\frac{1}{m}$th of the Nyquist rate along each dimension, ie. with a sampling matrix S whose diagonal terms are $\frac{m\pi}{\sigma_i}$:

$$S = [s_{ab}] \, , \; s_{ab} = \frac{m\pi}{\sigma_i} \text{ if } a = b = i \quad (2)$$

It can be shown that [3] *:*

$$f(\vec{t}) = \sum_{r=1}^{m} \sum_{k_1=-\infty}^{\infty} \ldots \sum_{k_n=-\infty}^{\infty} \phi_r(S\vec{k}) y_r(\vec{t} - S\vec{k}) \quad (3)$$

0-8186-7310-9/95 $4.00 © 1995 IEEE

where $\vec{k} = [k_1 \dots k_n]$ *is an integer valued vector and:*

$$y_r(\vec{t}) = \frac{|S|}{(2\pi)^n}\int_{-\sigma_1}^{\sigma_1+c_1} \cdots \int_{-\sigma_n}^{\sigma_n+c_n} Y_r(\vec{\omega},\vec{t}) exp(j\vec{\omega}^T\vec{t}) d\omega_1 \dots d\omega_n \quad (4)$$

$Y_r(\vec{\omega},\vec{t})$ *are given by the following set of simultaneous equations:*

$$\sum_{r=1}^{m} \mathcal{H}_r(\vec{\omega}+(\ell-1)\vec{c}) Y_r(\vec{\omega},\vec{t}) = exp(j(\ell-1)\vec{c}^T\vec{t}) \ , \ \ell = 1 \dots m \quad (5)$$

where $\vec{c} = [\frac{2\sigma_1}{m} \dots \frac{2\sigma_n}{m}]$ *and* $-\sigma_i \leq \omega_i \leq -\sigma_i + \frac{2\sigma_i}{m}$.

Note that the n-dimensional expansion in (3) is valid iff the sampling matrix S is non-singular. In what follows, we will assume that the matrix S is diagonal. In other words, we will assume a rectangular sampling. This, merely, implies that the sampling is performed independently along each dimension, in which case, S will be automatically non-singular, since the sampling density $|S| = \prod_{i=1}^{n} \frac{m\pi}{\sigma_i} \neq 0$.

Now, let the number of available low resolution frames be $\ell < m$ then by simply applying the principle of superposition, one can attempt to reconstruct a sample of $f(\vec{t})$ at the resolution $\frac{\ell}{m}$th of the Nyquist rate by minimizing the following error:

$$\epsilon^2 = \mathcal{E}\{(f_\ell(\vec{t}) - \sum_{r=1}^{\ell} \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_n=-\infty}^{\infty} \phi_r(S\vec{k}) y_r(\vec{t}-S\vec{k}))^2\} \quad (6)$$

where $f_\ell(\vec{t})$ is a sample of $f(\vec{t})$ at $\frac{\ell}{m}$th of the Nyquist rate along each dimension and $\mathcal{E}$ represents the expected value. Alternatively, we can minimize the error after sampling. Therefore if $\hat{\phi}_r(S\vec{k})$ denotes an estimate of $\phi_r(S\vec{k})$:

$$\epsilon_\phi^2 = \mathcal{E}\{\sum_{r=1}^{\ell}(\phi_r(S\vec{k}) - \hat{\phi}_r(S\vec{k}))^2\} \quad (7)$$

An interesting situation arises when $f_\ell(\vec{t})$ is expressed in terms of two (or more) variables (eg. albedo and height): $f_\ell(\vec{t}) = f_\ell(\vec{s}_1(\vec{t}), s_2(\vec{t}))$. We can then seek high resolution information for $s_1(\vec{t})$ and $s_2(\vec{t})$. This is obviously an inverse problem and can be tackled using regularization or equivalently MRF's.

3. PROBLEM STATEMENT AND PROPOSED METHOD

The n-dimensional extension of Papoulis' sampling expansion is an ideal tool for our purpose. In this context, in a sequence of low resolution images, each frame can be assumed to contain the recuring samples of a nonuniform sampling sequence obtained by applying a common input function to a set of linear shift invariant systems. We, obviously, need to register the recuring samples so that the theorem can be applied correctly. This implies a preprocessing, using registration algorithms such as [4][12].

Now, let g and z denote the vectors of unknown variables ie. the vectors of the albedo $g_{(x,y)}$ and the height $z_{(x,y)}$ on the super-resolution grid at $\frac{\ell}{m}$th of the Nyquist rate. Here, the input is a function of g and z: $f(g,z)$ (see Appendix A). Let also I denote the vector of all observed pixel values in our low resolution frames, ie. the vector of all $\phi_r(S\vec{k})$. Then using Bayes law and assuming that g and z are independent, we can write the following equation between the posterior and prior probability distributions:

$$p(g,z \mid I) = K\ p(I \mid g,z)\ p(g)p(z) \quad (K = \texttt{Const.}) \quad (8)$$

Applying the Hammersley-Clifford theorem [7] and assuming a Gaussian distribution for $p(I \mid g,z)$ we can calculate the Maximum a Posteriori (MAP) estimate by minimizing the following cost function:

$$\begin{aligned} E(g,z) &= -ln\ p((g,z) \mid I) \\ &= e^T C_e^{-1} e + u_g^T C_g^{-1} u_g + u_z^T C_z^{-1} u_z + const. \end{aligned} \quad (9)$$

where C_e, C_g and C_z are the corresponding covariance matrices, and u_g and u_z are the vectors of local potentials associated with the MRF's of g and z. Using a membrane model for $p(g)$ and $p(z)$ [2], since $e^T C_e^{-1} e = \epsilon_\phi^2$, we can write the following cost function at any point (x,y) of our sampled array:

$$\begin{aligned} E_{(x,y)} = &\sum_{\ell} \sum_{(k',l') \in \nu_{(\mathrm{k,l})}} \frac{(I^\ell_{(k',l')} - \hat{I}^\ell_{(k',l')})^2}{2\sigma_e^2} \\ &+ \sum_{(x',y') \in \nu_{(\mathrm{x,y})}} \frac{(\hat{g}_{(x,y)} - \hat{g}_{(x',y')})^2}{2\sigma_g^2} \\ &+ \sum_{(x',y') \in \nu_{(\mathrm{x,y})}} \frac{(\hat{z}_{(x,y)} - \hat{z}_{(x',y')})^2}{2\sigma_z^2} \end{aligned} \quad (10)$$

where $\nu_{(\mathrm{k,l})}$ is the neighbouring pixels to (k,l) whose intensities are affected by the irradiance at (x,y) (depending on the support of H, see Appendix A), $\nu_{(\mathrm{x,y})}$ is the neighbourhood structure depending on the order of the MRF associated with $\hat{g}$ and $\hat{z}$ ($\hat{}$ depicts estimated values), σ_e^2, σ_g^2 and σ_z^2 are the variances of the error vector, g and z, respectively. Therefore, the MAP estimator amounts to minimizing (10) for which we have employed a generalized simplex algorithm [10].

4. EXPERIMENTAL RESULTS

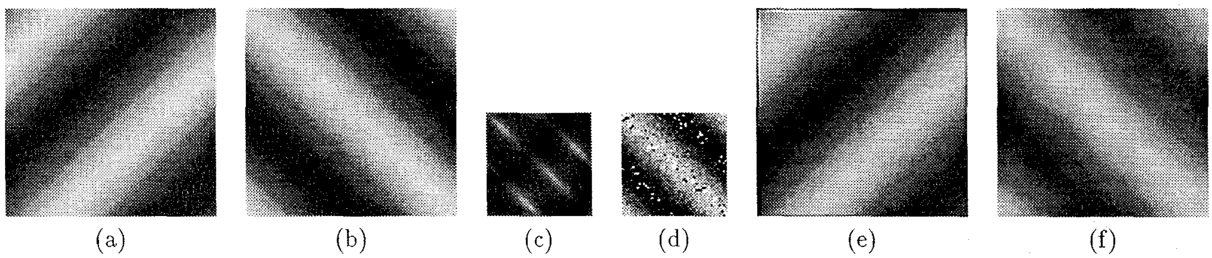

Figure 1: (a), (b) albedo & height, (c) one of the simulated low resolution camera images, (d) a noise corrupted low resolution height (SNR=5 dB), (c) reconstructed albedo (SNR=34 dB), (d) reconstructed height (SNR=27 dB)

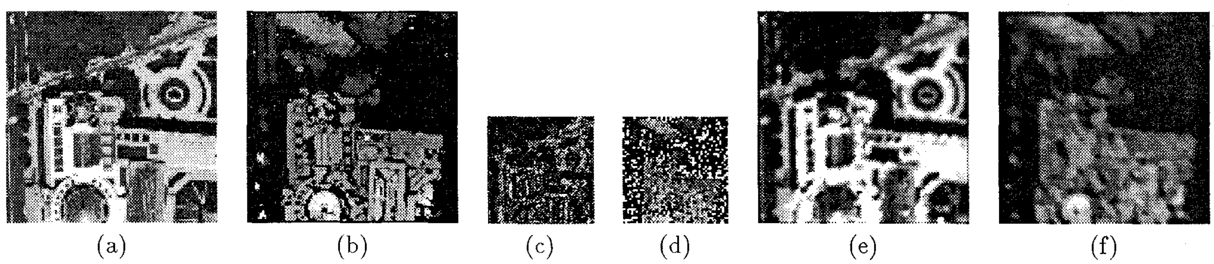

Figure 2: (a) albedo, (b) height, (c) one of the low resolution camera images, (d) a noise corrupted low resolution height (SNR=5 db), (e) reconstructed albedo (SNR=25 dB), (f) reconstructed height (SNR=22 dB)

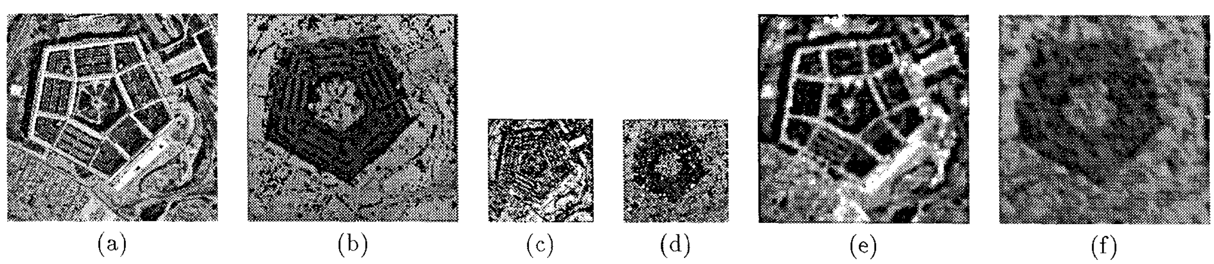

Figure 3: (a) albedo, (b) height, (c) one of the simulated low resolution camera images, (d) a noise corrupted low resolution height (SNR=5 db), (e) reconstructed albedo (SNR=26 dB), (f) reconstructed height (SNR=24 dB)

5. DISCUSSION ON RESULTS

The main problem encountered was the trade-off between the regularization terms and the error term in the cost function (10), which proved to be not so easy when dealing with real images. The algorithm could be improved by including discontinuity fields in our *a priori.*

The generalized simplex algorithm seems to slow down as we approach the solution. A possible remedy would be to switch to a gradient method with fast convergence properties in the vicinity of the solution where the cost function is likely to exhibit convexity.

Since the problem has been formulated as a constrained optimization one, further extensions of the proposed method can be considered along the same direction. Therefore, we might consider, for example, optimizing with respect to camera parameters in matrix A (see below), or the variances in equation (10).

Appendix A

The input to our linear systems is, merely, the Lambertian imaging function defined by:

$$f(g(\vec{t}), z(\vec{t})) = g(\vec{t})R(\vec{t}) \qquad (11)$$

where $\vec{t} = (x, y)$, g is the albedo of the surface which is assumed to be varying and R is the reflection function of the surface which depends on the direction of the local normal to the surface. The latter is given by local variations of height, ie. the partial derivatives of $z(\vec{t})$ along each dimension (see [1][5] for details). In our case, for example, the local normal will be given by $(\frac{\partial z}{x}, \frac{\partial z}{y}, -1)$.

We have assumed a Gaussian kernel for the linear shift invariant systems and therefore the imaging model prior to discretization is given by a convolution integral:

$$\phi_r(\vec{t}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(\vec{t} - \vec{\tau})g(\vec{t})R(\vec{t})\, dx\, dy \qquad (12)$$

where τ is a shift vector. In practice, the support of H is approximated by a finite size window. Fourier transforming the right hand side and using the convolution theorem, we can easily verify that this equation corresponds to equation (1). Note also that $\phi_r(\vec{t})$ are continuous versions of our discrete images given by $I_{(k,l)}$, where camera coordinates (k, l) are related to the world coordinate frame by: $[k, l]^T = A\ [x, y, z, 1]^T$ wher A is a 2×4 matrix containing camera parameters. The last column of A contains the interframe shift parameters, specifying the coregistration of image frames. Other parameters specify the sampling rate along the two axes and the rotation of the camera around the origin (for further details see [4][11]).

6. REFERENCES

[1] M. Berthod, H. Shekarforoush, M. Werman, and J. Zerubia. Reconstruction of high resolution 3d visual information. In *Proc. IEEE CVPR*, pages 654–657, Seattle, Washington, 1994.

[2] A. Blake and A. Zisserman. *Visual Reconstruction.* MIT Press, 1987.

[3] K. F. Cheung and R. J. Marks. Papoulis' generalization of the sampling theorem in higher dimensions and its applications to sample density reduction. In *Proc. Int. Conf. on circuits and systems*, Nanjing, China, 1989.

[4] E. De Castro and C. Morandi. Registration of translated and rotated images using finite fourier transforms. In *IEEE PAMI*, volume 9, No 5, pages 700–703, 1987.

[5] B. K. P. Horn. Understanding image intensities. *Artificial Intelligence*, 8(2):201–231, 1977.

[6] D. Keren, S. Peleg, and R. Brada. Image sequence enhancement using sub-pixel displacement. In *Proc. CVPR*, pages 742–746, Ann Arbor, Michigan, June 1988.

[7] J. Moussouris. Gibbs and markov random systems with constraints. *Journal of Statistical Physics*, 10(1), 1974.

[8] A. Papoulis. Generalized sampling expansion. In *IEEE Trans. on Circuits and Systems*, volume 24, No 11, Nov. 1977.

[9] A. Papoulis. *Signal Analysis.* McGraw-Hill Book Company, 1977.

[10] H. Shekarforoush, M. Berthod, and J. Zerubia. Direct search generalized simplex algorithm for optimizing non-linear functions. research report N^o 2535, INRIA - France, 1995.

[11] H. Shekarforoush, M. Berthod, J. Zerubia, and M. Werman. Sub-pixel bayesian estimation of albedo and height. *International Journal of Computer Vision*, to appear.

[12] R. Tsai and T. Huang. Multiframe image restoration and registration. *Adv. Comp. Vis. Im. Proc.*, 1, 1984.