A Systolic-Array Architecture for First-Order 4-D IIR Frequency-Planar Digital Filters

Randeel Wimalagunarathne
and Arjuna Madanayake*
Electrical and Computer Engineering
University of Akron
Akron, OH, 44325-3904
Email: arjuna@uakron.edu

Donald G. Dansereau
Australian Centre for Field Robotics
School of Aerospace, Mechanical and
Mechatronic Engineering
University of Sydney, NSW, Australia
Email: d.dansereau@acfr.usyd.edu.au

Len T. Bruton
*Electrical and Computer Engineering
University of Calgary
Calgary, AB, Canada

Abstract—A novel parallel semi-systolic semi-scanned array architecture is proposed for the implementation of four-dimensional (4-D) IIR filters. These filters have emerging applications in computed tomography (CT), volumetric ultrasound, and light field processing for computer vision. The proposed architecture can be applied to a broad class of 4-D IIR filters, and we show results for a frequency-planar depth-selective filter. Our implementation is on a Xilinx Virtex-6 xc6vsx315t-3ff1156 FPGA, and is suitable for filtering of a $N_1 \times N_2 = 4 \times 4$ aperture light field camera input. Results compare favourably with ideal and FPGA-hardware measured outputs, with an $N_1 N_2$ factor increase in throughput compared to a corresponding fully raster-scanned design clocked at the same clock frequency.

I. INTRODUCTION

Light fields [1] are gaining popularity as a rich means of describing the light permeating a scene. A light field is a four-dimensional (4-D) subset of the more complete Plenoptic function [2] – it describes the behaviour of light in terms of both ray directions and positions. A light field camera measures a subset of the rays passing through a scene, and can be implemented as a planar array of cameras/apertures [3] – a generalization on two-aperture stereo and three-aperture trinocular cameras – or by placing a lens array in the optical path of an otherwise normal camera [4]. Some applications of multi-view camera arrays are found in [5], [6].

Four-dimensional digital filters are a fundamental building block in light field processing, with applications in depth estimation, depth filtering, distractor isolation and visual odometry [7]–[11]. This paper is focused on the efficient implementation of 4-D infinite impulse response (IIR) filters, drawing on the specific example of frequency-planar depth filters (see [12] for examples). The input to our filter is a light field camera – our development discusses an aperture array, but a lens array-based camera can be directly substituted. Denoting $\mathbf{n} = (n_1, n_2, n_3, n_4)$, the 4-D quantized and sampled light field is given as $w(\mathbf{n}) = Q[w_a(n_1 \Delta s, n_2 \Delta t, n_3 \Delta u, n_4 \Delta v)]$, where $w_a(s, t, u, v) \in \mathbb{R}$ is the 4-D continuous domain light field, distances $(\Delta s, \Delta t) \in \mathbb{R}^2$ correspond to spacing between apertures on the array plane, $(\Delta u, \Delta v) \in \mathbb{R}^2$ correspond to spacing between pixel elements in the focal plane, and Q denotes quantization, typically 8 bits per color field. We propose a VLSI architecture for the computation of $y(\mathbf{n}) = h_d(\mathbf{n}) \overset{4-D}{*} w(\mathbf{n})$
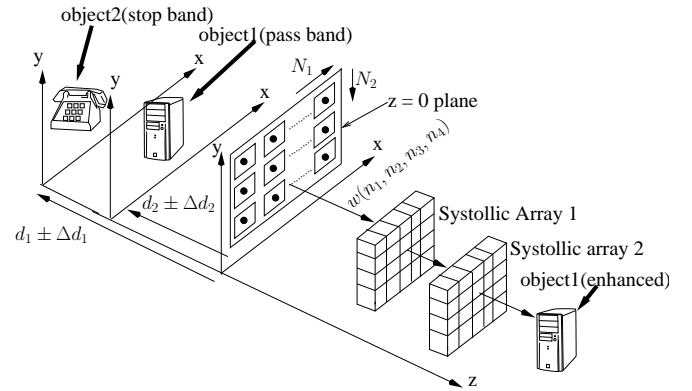

Fig. 1. Depth-selective filtering of a light field camera input using a cascade of two 4-D IIR frequency-hyperplanar filters.
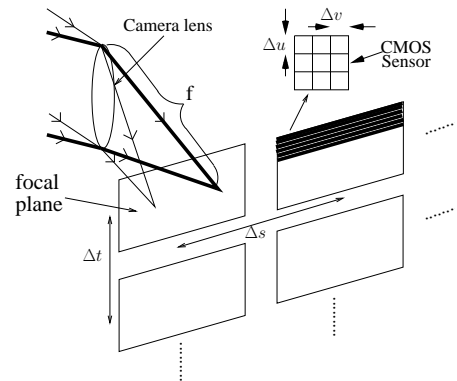

Fig. 2. Focal plane u, v and aperture plane s, t.

assuming zero initial conditions (ZICs) where $\overset{4-D}{*}$ denotes 4-D convolution, and $h(\mathbf{n})$ is the 4-D unit impulse response of the frequency-planar depth filter. In the $\mathbf{z}$-transform domain, the filtering operation becomes $Y(\mathbf{z}) = H(\mathbf{z})W(\mathbf{z})$ where $(p, P) \in \{(w, W), (h_d, H_d), (y, Y)\}$, $p(\mathbf{n}) \overset{4-D}{\Leftrightarrow} P(\mathbf{z})$ and where $\mathbf{z} \equiv (z_1, z_2, z_3, z_4) \in \mathbb{C}^4$.

II. 4-D IIR DEPTH FILTERS

A common parameterization of the 4-D light field is the 2-plane parametrization (Figs. 1, 2), in which apertures are

indexed on an s, t plane ($z = 0$), and pixels on a second u, v plane ($z = f$). Under this scheme, a ray in the scene is uniquely identified by its points of intersection with the s, t and u, v planes [8]. The common light field camera implementations can all yield two-plane parameterized light fields, and we utilize this parameterization for our input. We implement a class of 1st-order practical bounded-input bounded-output (p-BIBO) stable [13] 4-D IIR filters having transfer functions of the form

$$H(\mathbf{z}) = \frac{\prod\limits_{k=1}^{4} \left(1 + z_k^{-1}\right)}{1 + \underbrace{\sum\limits_{i=0}^{1}\sum\limits_{h=0}^{1}\sum\limits_{k=0}^{1}\sum\limits_{l=0}^{1} b_{ihkl} z_1^{-i} z_2^{-h} z_3^{-k} z_4^{-l}}_{i+h+k+l \neq 0}} \equiv \frac{Y(\mathbf{z})}{W(\mathbf{z})}. \tag{1}$$

A depth-selective frequency-hyperplanar filter can be implemented in this form by assigning the filter coefficients

$$b_{ihkl} \equiv \frac{R + (-1)^i L_1 + (-1)^h L_2 + (-1)^k L_3 + (-1)^l L_4}{R + \sum\limits_{K=1}^{4} L_K}, \tag{2}$$

where $L_k \geq 0$, $k = 1, 2, 3, 4$, and $R > 0$. A highly selective frequency-planar filter is formed as a cascade of these hyperplanar filters, $H_d(\mathbf{z}) = H_1(\mathbf{z})H_2(\mathbf{z})$, where each $H_k(\mathbf{z})$ is of the form (1) [8]. From the transfer function (1), we obtain the following 4-D $\mathbf{z}$-domain input-output relationship $Y(\mathbf{z}) =$

$$\frac{1}{\psi_O(z_3^{-1}, z_4^{-1})} \cdot (W(\mathbf{z})(1 + z_1^{-1})(1 + z_2^{-1})\psi_L(z_3^{-1}, z_4^{-1})$$
$$- Y(\mathbf{z})z_1^{-1}\psi_M(z_3^{-1}, z_4^{-1}) - Y(\mathbf{z})z_2^{-1}\psi_N(z_3^{-1}, z_4^{-1})$$
$$- Y(\mathbf{z})z_2^{-1}z_1^{-1}\psi_P(z_3^{-1}, z_4^{-1})) \tag{3}$$

$$\psi_L(z_3^{-1}, z_4^{-1}) = 1 + z_3^{-1} + z_4^{-1} + z_3^{-1}z_4^{-1} \tag{4}$$
$$\psi_M(z_3^{-1}, z_4^{-1}) = b_{1000} + b_{1001}z_4^{-1} + z_3^{-1}(b_{1010} + b_{1011}z_4^{-1}) \tag{5}$$
$$\psi_N(z_3^{-1}, z_4^{-1}) = b_{0100} + b_{0101}z_4^{-1} + z_3^{-1}(b_{0110} + b_{0111}z_4^{-1}) \tag{6}$$
$$\psi_P(z_3^{-1}, z_4^{-1}) = b_{1100} + b_{1101}z_4^{-1} + z_3^{-1}(b_{1110} + b_{1111}z_4^{-1}) \tag{7}$$
$$\psi_O(z_3^{-1}, z_4^{-1}) = 1 + b_{0001}z_4^{-1} + z_3^{-1}(b_{0010} + b_{0011}z_4^{-1}) \tag{8}$$

III. PROPOSED MASSIVELY-PARALLEL ARCHITECTURE

We now describe the proposed massively-parallel systolic-array architecture for the computation of $y_k(\mathbf{n}) = h_k(\mathbf{n}) \overset{4-D}{*} w_k(\mathbf{n})$, where w_k, y_k, h_k correspond to the input, output, and impulse response of the frequency hyperplanar digital filter stage $k = 1, 2$. Convolution is achieved by iteratively computing a 4-D difference-equation, found as the inverse $\mathbf{z}$-transform of (1).

We assume an aperture array of size $N_1 \times N_2$, with $N_3 \times N_4$ pixels per aperture, where $N_1, N_2, N_3, N_4 \in \mathbb{Z}^+$. The light field is raster scanned along columns $n_3 < N_3$ and rows $n_4 <$
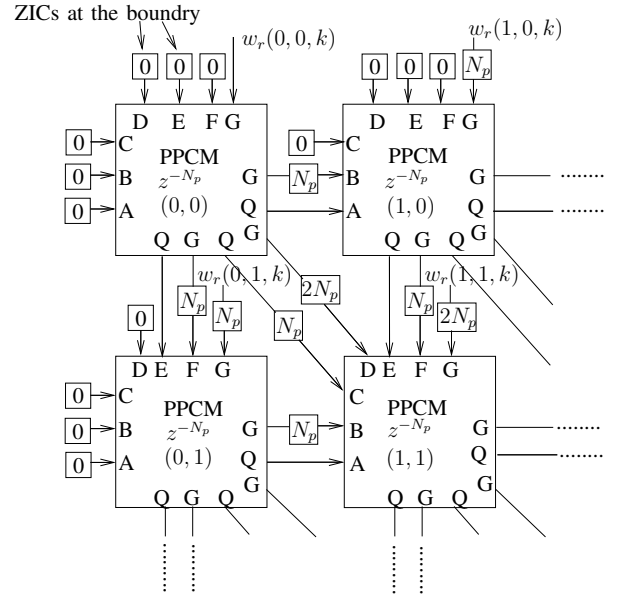

Fig. 3. Systolic array architecture having 7-inputs $A, B, ..., G$ and 1-output Q. Note, for diagrammatic clarity, multiple labelling on the PPCM denotes fan-out connections where signals are fed-through unchanged.

N_4, leading to a time-multiplexed sequence from each aperture: $w_r(n_1, n_2, k), k = N_3 n_4 + n_3$. Each such sequence will be processed by a parallel processing core module (PPCM). This leads to an array of locally interconnected PPCMs – one for each aperture – resulting in an $N_1 \times N_2$ array of PPCMs that form a systolic array of processors. Time synchronous operation is assumed for all PPCMs in the systolic array as well as the $N_1 N_2$ parallel channels of serial data derived from the aperture array.

A. PPCM Signal Flow Graph (SFG)

The inverse $\mathbf{z}$-transform of equation (3) is realized on each PPCM resulting in a 7-input and 1-output circuit. The systolic array of PPCMs that perform the 4-D filtering operation is depicted in Fig. 3. ZICs are required for practical-BIBO stability and are achieved by zeroing all data registers and FIFOs at start up, and connecting ZICs at boundaries of the systolic array architecture.

The spatial delays along the columns n_3 and row n_4 of each focal plane image is represented by the corresponding $\mathbf{z}$-transform z_3^{-1} and z_4^{-1}. These spatial delays require ZICs which are realized in VLSI hardware circuits using spatial delay processors that can either delay a serial stream of data by 1 time sample, for a column delay, or by N_3 time samples, for a row delay, while ensuring the initial conditions for both row and columns are zero [14]. The PPCM interconnections along the n_1 and n_2 dimensions are fully systolic. However, the operation along n_3 and n_4 is raster-scanned, making the architecture a hybrid between fully-systolic and raster-scanned methods. The PPCMs are fully parallel, employing 15 multipliers and 24 adders following Fig. 4, where, $\psi_k(z_3^{-1}, z_4^{-1}), k \in \{L, M, N, O, P\}$ are defined in (3)-(8).

TABLE I

INTER PPCM CONNECTION DESCRIPTION.

Port	Fed by/Feeds	Description
A	$y_r(n_1-1, n_2, k)$	In [fb]
B	$w_r(n_1-1, n_2, k-N_p)$	In [ff]
C	$y_r(n_1-1, n_2-1, k-N_p)$	In [fb]
D	$w_r(n_1-1, n_2-1, k-2N_p)$	In [ff]
E	$y_r(n_1, n_2-1, k)$	In [fb]
F	$w_r(n_1, n_2-1, k-N_p)$	In [ff]
G	$w(n_1, n_2, k-(n_1+n_2-1)N_p)$	In [ff]
G	port B at (n_1+1, n_2)	ft
Q	port A at (n_1+1, n_2)	Out [fb]
G	port D at (n_1+1, n_2)	ft
Q	port C at (n_1+1, n_2)	Out [fb]
G	port F at (n_1, n_2+1)	ft
Q	port E at (n_1, n_2+1)	Out [fb]

From Fig. 4, it is clear that the SFG of the PPCM has 4 inputs containing input signals, and 3 inputs connected to *outputs* of neighbouring PPCMs. For example, each $y_r(n_1, n_2, k)$ output stream at aperture array grid point (n_1, n_2) is connected to the spatial feedback input ports requiring b_{10kl} terms for the PPCM at n_1+1, n_2, b_{01kl} terms for PPCM at n_1, n_2+1, and b_{11kl} terms for the PPCM at n_1+1, n_2+1, respectively.

We apply a fine-grained pipelining to the PPCMs, resulting in lower critical path delays at the cost of latency. The additional clocked latency, in periods, for each of the 7 input ports $A, B, ..., G$ to the output port Q, is $N_p \in \mathbb{Z}^+$ clock cycles. Additional clocked first-in-first-out (FIFO) buffer memories are inserted between PPCMs to compensate for pipelining. Further, the input signal at (n_1, n_2) camera array location needs to be delayed by $N_p(n_1+n_2)$ clock cycles using FIFO buffers in order to compensate for pipeline latency inside the PPCMs. The 4-D filtered output at port G of the PPCM at n_1, n_2 must in turn be delayed by clocked FIFOs of depth $N_p(N_1+N_2-n_1-n_2-2)$. The inter-PPCM connections are summarized in Table I, where ff denotes feed-forward, fb denotes feedback, and ft denotes a direct feed-through connection.

TABLE II

DESIGN PARAMETERS AND MEASURED RESULTS.

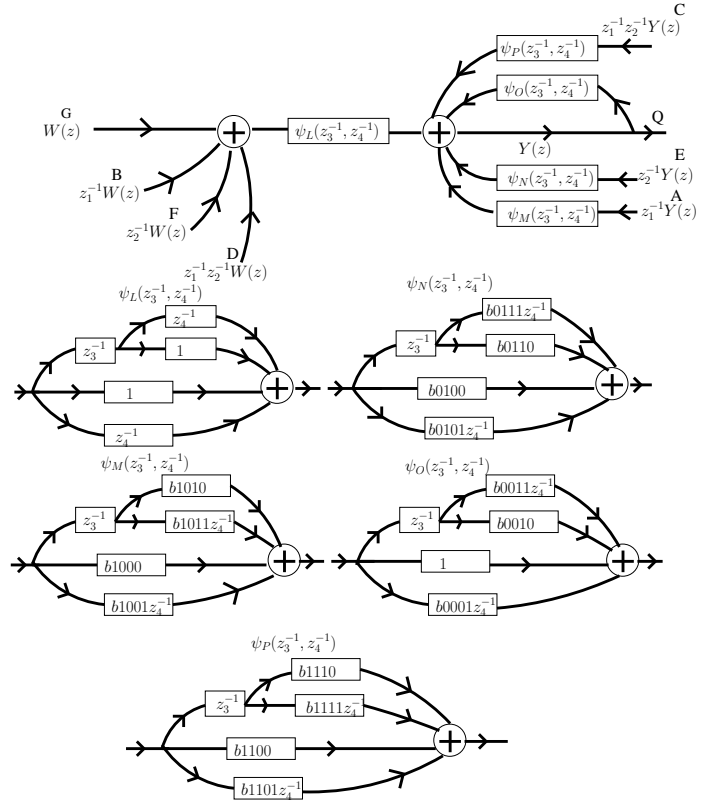|	Example Design
W_{in}	8
D_{in}	0
W_{mul}	18
D_{mul}	12
N_1	4
N_2	4
N_3	256
N_4	256
W_{out}	23
D_{out}	15
Max frequency	$39.42 MHz$
Slices	$17,784$
Slice LUTs	$59,250$
Slice flip-flops	$34,994$
LUT flip-Flop pairs	$61,895$
DSP48s	240
Bonded IOBs	481
BUFGs	1

Fig. 4.   Signal flow graph of PPCM corresponding to (3)-(8).

IV. RESULTS

A. FPGA Architecture

We provide an example implementation on a Xilinx Virtex-6 xc6vsx315t-3ff1156 field-programmable gate array (FPGA) device installed on a Xilinx ML605 rapid prototyping system. Our example assumes a light field camera input with 4×4 apertures, where each aperture yields a 256×256-pixel image with 8 bits/pixel of precision. For the purposes of verification, a 4-D input signal comprising of a unit impulse $w(\mathbf{n}) = \delta(\mathbf{n})$ is used to excite the 4×4 systolic array, leading to the 4-D unit impulse response tensor $h_1(\mathbf{n})$ of bounded size $(4 \times 4 \times 256 \times 256)$. We measure the fixed-point 4-D frequency response by computing the 4-D fast Fourier transform of $h_1(\mathbf{n})$. In Fig. 5, we compare our measured on-chip FPGA-circuit response with the impulse response of an infinite-precision simulation, and a fixed-point simulink design. Shown is a 2-D slice of the 4-D transfer-function, for $\omega_{3,4} = 0$.

The design parameters are summarized in Table II, in which W_{in} and D_{in} define precision as input word size and binary point position, W_{out}, D_{out} define output precision and W_{mul}, D_{mul} define coefficient precision. The minimum critical path delay (CPD) of the circuit $T_{CPD} \approx T_M + T_{A/S} + T_{MUX}$, where T_M, $T_{A/S}$ and T_{MUX} are the propagation delays of the multiplier, adder/subtractor, and two-input W-bit multiplexer, resulting in the maximum clock frequency $F_{CLK,Max} = 1/T_{CPD}$. If the VLSI real-estate requirements for W_{mul}-bit multipliers and adders are γ_M, and $\gamma_{A/S}$, respec-
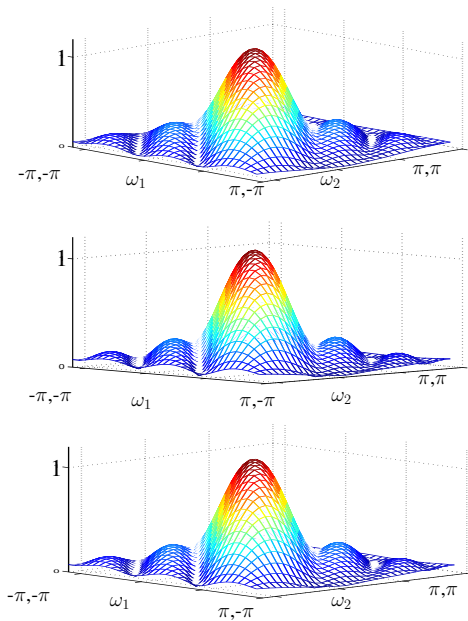
Fig. 5. 2-D slices of the 4-D magnitude frequency response $|H(e^{j\omega_1}, e^{j\omega_2}, e^{j\omega_3}, e^{j\omega_4})|$ for $\omega_3 = \omega_4 = 0$: ideal (top), Simulink simulation (middle), and on-chip measured results from FPGA physical implementation using the stepped hardware simulation method provided by Xilinx and Matlab (bottom); all plots are shown over the Nyquist square $-\pi \leq \omega_{1,2} < \pi$.

tively, then the total VLSI resource consumption of the circuit is approximated by $\gamma_T \approx \gamma_M (15N_1N_2 - 8(N_1 + N_2) - 4) + \gamma_{A/S}(24N_1N_2 - 9(N_1 + N_2 + 1))$.

B. Computational Complexity and Throughput

The throughput of the proposed design is N_1N_2 parallel frames every N_3N_4 clock cycles. The 4-D light field frame rate is therefore $F_{4D} = \frac{F_{clock}}{N_3N_4}$. For $F_{clock} = 39.42$ MHz, this means a light field frame rate of $F_{4D} = 601.50$ Hz. That is, more than 600 light fields can be filtered in real-time assuming a 4×4 array of apertures having resolution 256×256. When compared to a corresponding fully raster-scanned architecture [14], the proposed architecture achieves an N_1N_2-fold increase in image frame rate (throughput) at the cost of an $\approx N_1N_2$-fold increase in multiplier complexity, with no change in system clock frequency or memory requirement.

C. Fixed-point Effects

The port width of the output block for the PPCM is set to W_{out}-bits with binary point at location D_{out}. The precision level at each input port is $w_{in,k}$, $k \in \{A, B, ..., G\} = W_{out} - n_k$, where $n_k \in \mathbb{Z}^+$ are the number of adder stages from each input port k, $k \in \{A, B, ..., G\}$ to the filter output port Q.

That is, for each path from input to output having n_k adders, the input precision is correspondingly truncated to $w_{out} - n_k$, leading to additional quantization noise due to loss of n_k least significant bits.

V. CONCLUSIONS

We proposed a massively-parallel systolic-array VLSI architecture for the real-time implementation of 1st−order p-BIBO stable 4-D IIR digital filters having transfer functions of form (1). A single-FPGA example was shown realizing a frequency-planar depth-selective light field filter, operating on light fields with 4×4 apertures and 256×256 pixels per aperture. The prototype 4-D IIR filter is raster-scanned along the image frames but fully systolic along the aperture array dimensions, resulting in an array of PPCMs each dedicated to a corresponding aperture. The FPGA realization was based on a Xilinx Virtex-6 xc6vsx315t-3ff1156 FPGA device and is operational on-chip.

Measured 4-D unit impulse responses were transformed to the 4-D frequency domain, and shown to conform closely to the ideal frequency response resulting from infinite-precision simulation. The prototype FPGA physical implementation is capable of maintaining a light field frame rate of 601.5 Hz making it suitable for the most demanding of light field based computer vision applications. To our knowledge, apart from raster-scanned 4-D IIR filters [14], the proposed architecture is the *only available* architecture for real-time 4-D IIR filtering.

REFERENCES

[1] M. Levoy and P. Hanrahan, "Light field rendering," in *Proceedings of the 23rd annual conf. on Computer graphics and interactive techniques*. ACM, 1996, pp. 31–42.

[2] E. Adelson and J. Bergen, "The plenoptic function and the elements of early vision," *Computational models of visual processing*, vol. 1, pp. 3–20, 1991.

[3] B. Wilburn, N. Joshi, V. Vaish, E. V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy, "High Performance Imaging Using Large Camera Arrays," vol. 24, 2005, pp. 765–776.

[4] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," *Computer Science Technical Report CSTR*, vol. 2, 2005.

[5] B. Wilburn, N. Joshi, V. Vaish, M. Levoy, and M. Horowitz, "High speed videography using a dense camera array," in *In IEEE Society Conference on Pattern Recognition*, 2004, pp. 294–301.

[6] N. Joshi, S. Avidan, W. Matusik, and Kriegman, "Tracking through Occlusions using Multiple Cameras," in *IEEE ICCV*, 2007.

[7] D. Dansereau, "4D light field processing and its application to computer vision," Master's thesis, University of Calgary, 2003.

[8] D. Dansereau and L. T. Bruton, "A 4D frequency-planar IIR filter and its application to light field processing," vol. 4, ISCAS 2003, pp. 476–479.

[9] D. G. Dansereau and L. T. Bruton, "Gradient-based depth estimation from 4D light fields," in *Proceedings of the Intl. Symp. on Circuits and Systems*, vol. 3. IEEE, May 2004, pp. 549–552.

[10] D. G. Dansereau and S. B. Williams, "Seabed modeling and distractor extraction for mobile auvs using light field filtering," in *Accepted to Robotics and Automation (ICRA), 2011 IEEE Intl. Conf. on*, May 2011.

[11] D. G. Dansereau, I. Mahon, O. Pizarro, and S. B. Williams, "Plenoptic Flow: Closed-Form Visual Odometry for Light Field Cameras," in *Intelligent Robots and Systems (IROS), Proceedings, 2011 IEEE/RSJ International Conference on*. IEEE, Sept 2011, pp. 4455–4462.

[12] "$http://dl.dropbox.com/u/4457078/ver1/light\_fields/welcome.html$," 2011.

[13] P. Agathoklis and L. Bruton, "Practical-BIBO stability of N-dimensional discrete systems," *Proc. Inst. Elec. Eng.*, vol. 130, Pt. G, no. 6, pp. 236–242, 1983.

[14] A. Madanayake, R. Wimalagunaratne, D. G. Dansereau, and L. T. Bruton, "Design and FPGA-implementation of 1st-order 4D IIR frequency-hyperplanar digital filters," in *Circuits and Systems (MWSCAS), 2011 IEEE 54th International Midwest Symposium on*, 2011, pp. 1–4.