16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

# IMPLEMENTATION OF HARDWARE-OPTIMIZED 3-D WAVE DIGITAL FILTERS FOR MOTION-BASED OBJECT DETECTION IN VIDEO SCENES

*Sam Schauland, Joerg Velten, Anton Kummert*

Department of Electrical, Information and Media Engineering, University of Wuppertal
Rainer-Gruenter-Str. 21, 42119 Wuppertal, Germany
phone: + (49) 202 4391808
email: {schauland|velten|kummert}@uni-wuppertal.de
web: http://www.tnt.uni-wuppertal.de

## ABSTRACT

*Even though filtering objects in video scenes using linear filters has been an active research field for many years, the presented results have been more focused on theory than on application, which is mainly due to the high computational complexity of the developed filter structures that former computer systems could not process in real-time. Particularly due to high product quantities and high competition, computation costs of applications like object detection in driver assistance systems, for instance, still have to be as small as possible. This work is focused on 3-D linear shift invariant (LSI) filters targeting the separation of signals due to different directional components. We have implemented wave digital filters with low complexity and therefore low hardware requirements with the aim of developing a cost efficient real-time object detection system. The developed filters have been both simulated using Simulink Fixed Point and implemented on a Xilinx Spartan 3E FPGA Board. Experiments have shown promising results underlining the potential of motion-based object detection to lead to new and effective methods in driver assistance system development.*

## 1. INTRODUCTION

This paper is focused on the separation of three-dimensional signals due to different velocity or directional components, more precisely detecting moving objects in visual scenes by means of linear multidimensional filters. While soundness of results obtained by similar systems developed in the past has been convincing, interest in those systems declined to the reduced computational power of contemporary computers.

Particularly due to high product quantities and high competition, computation costs of applications like object detection in driver assistance systems, for instance, still have to be as small as possible.

This constraint and recent trends in automotive applications made us choose an FPGA (field-programmable gate array) as a final platform. As size, capabilities, and speed of FPGAs are increasing with prices decreasing simultaneously, they are capable of taking over larger and larger functions and can even be marketed as full systems on chips (SoC).

The work presented in this paper is focused on 3-D linear shift invariant (LSI) filters [1] targeting the separation of signals due to different directional components. We implement wave digital filters with low complexity and therefore low hardware requirements with the aim of developing cost efficient real-time object detection systems.

Advanced theoretical knowledge on system description and handling (using state-space methods like the Givone-Roesser model, for example [2]) available for these filters is a big advantage compared to most recently presented approaches in the wide field of object detection applications. In most cases nonlinear filters and methods originating in the field of computer vision rather than multidimensional signal processing are employed.

The input video sequence for our system is captured by a CMOS camera for automotive purposes, and the application example we give in this paper is highway traffic surveillance and therefore the scene background is almost constant (aside from changing weather conditions). However, this work is not limited to a static background: the proposed system is able to detect objects moving in a certain direction independent from the background and without any velocity restrictions, i.e. every object in the image sequence not moving in the specified direction is oppressed, independent from its individual velocity.

The relation between movement paths in the spatial domain to their spectrum correspondents is described and exemplary transfer functions as well as test sequences are presented.

The remainder of this paper proceeds as follows. First, the following section gives theoretical background on "linear trajectory signals" and wave digital filters. Next, we describe design and implementation of the filter used for object detection in section 3. Section 4 presents experimental results on real video images. Finally, we conclude the paper and discuss future work in section 5.

## 2. THEORETICAL BACKGROUND

### 2.1 Linear trajectory signals

A very good signal representation for detection of two-dimensional objects homogeneously moving in a video sequence has been proposed by Bruton et al. [3][4]. The corresponding signal class is called "linear trajectory signals" and will be described subsequently.

In this paper, we focus on signals that are constant on straight lines in the spatio-temporal domain, i.e. the objects' velocity is assumed to be constant, or, in other words, their acceleration is zero.

A descriptive introduction to the Fourier description of multidimensional visual scene representations is given subsequently by investigation of a temporally moving 1-D object. Let $f(x)$ be a spatially limited 1-D signal describing the object, i.e. $f(x) \neq 0$ for $a < x < b$, with $a, b, x \in R$, and let $F(jk_x)$ be its respective Fourier transform. A linear movement of this object in positive direction of $x$ by constant speed

$v$ is characterized by the 2-D signal

$$s_c(t,x) \quad = \quad f(x-vt). \tag{1}$$

The Fourier transform of $s_c(x,t)$ with respect to the spatial variable $x$ can be evaluated by the Fourier shift theorem as

$$\mathscr{F}_x\{s_c(t,x)\} \quad = \quad F(jk_x)\cdot e^{-jvtk_x}, \tag{2}$$

where $\mathscr{F}_x$ denotes the Fourier transform with respect to the variable $x$.
The Fourier transform with respect to the temporal variable $t$ is evaluated by using the modulation theorem, which leads to the 2-D Fourier transform of $s_c(t,x)$

$$\begin{aligned} S_c(j\omega_t, jk_x) &= \mathscr{F}_t\{F(jk_x)\cdot e^{-jvk_x t}\} \\ &= F(jk_x)\cdot 2\pi\cdot\delta(\omega_t + vk_x). \end{aligned} \tag{3}$$

The defined movement can thus be decomposed into complex exponential waves for which the relation between $k_x$ and $\omega_t$ is fixed, namely $k_x = -\frac{1}{v}\omega_t$, i.e. the respective Fourier components are located on a line including the origin. The latter is clarified by fig. 1: fig. 1(a) shows the locus of $f(x)$ and fig. 1(b) the respective nonzero points of $S_c(j\omega_t, jk_x)$.

This description is easily generalized to the n-D case by replacing scalars by vectors. The object is now an (n-1)-D object $f(\mathbf{r}), \mathbf{r}\in R^{n-1}$ with its (n-1)-D Fourier transform $F(j\mathbf{k})$, where $j\mathbf{k} = (jk_x, jk_y, jk_z)'$ in the case of three spatial coordinates. $\mathbf{v} = (v_x, v_y, v_z)'$ contains velocity information corresponding to each component of $\mathbf{r}$. The Fourier transform of the n-D signal

$$s_c(t,\boldsymbol{r}) = f(\mathbf{r}-\mathbf{v}t) \tag{4}$$

is given by

$$\mathscr{F}\{f(\mathbf{r}-\mathbf{v}t)\} = F(j\mathbf{k})\cdot 2\pi\cdot\delta(\omega_t + \mathbf{v}'\mathbf{k}). \tag{5}$$

If the object defined by eq. (1) is transferred into a higher dimensional space, e.g.

$$s_c(t,\mathbf{r}) \quad = \quad f(x-vt), \tag{6}$$

the resulting wave is a combination of plane waves with propagation speed $v$ and identical propagation direction. The shape of the wave in this direction is defined by Fourier components along the respective line, i.e. the respective composition of plane waves. While in eq. (6) this wave is propagating in direction of the first space coordinate, combinations of plane waves propagating in direction of the normed vector $\mathbf{r}_0$ can be described by

$$s_c(t,\mathbf{r}) \quad = \quad f(\mathbf{r}_0'\mathbf{r}-vt), \tag{7}$$

where $\mathbf{v} = \mathbf{r}_0\cdot|\mathbf{v}| = \mathbf{r}_0\cdot v$. In the n-D Fourier description, such waves are thus defined by spectral components that are located on a plane including the origin. The angle between $\omega_t$-axis and the normal vector of this plane determines the propagation speed of the wave.

### 2.2 Short Introduction to Wave Digital Filters (WDF)

Before we proceed to the next step, namely filter implementation, we give a short introduction on the type of filters used in this context: WDFs were introduced by Fettweis [5] in the 1970s. As they are directly derived from classic analog circuits (reference circuits) every delay element in the WDF can be interpreted physically as holding the current state of an inductor or capacitor, which distinguishes them from most other digital filter types.
Instead of the voltages and currents of the reference circuit linear combinations thereof are used as signals. These linear combinations are called "‘wave variables"’. In this way a class of digital filters is obtained whose passivity and therefore stability constraints can be easily assured–even if using finite arithmetic [6].
Multidimensional wave digital filters (n-D WDF) are the result of the generalization of the one-dimensional concept based on multiple independent variables. N-D WDFs are derived from n-D reference circuits depending on $n$ complex frequency variables $\psi_1,...,\psi_n$. The correspondence between the frequency domain of the WDF and the reference filter is defined by the bilinear transform

$$\psi_i = \frac{z_i-1}{z_i+1}, \quad i = 1,...,n. \tag{8}$$

A detailed treatment of WDFs can be found in [5], for example.

## 3. FILTER IMPLEMENTATION

The platform for the final object detection system will be an FPGA. Thus, the transfer function presented above has been implemented as a WDF using fixed-point two's-complement arithmetic. In this section, we present details on the resulting WDF-structure and aspects that have to be taken into account to guarantee filter stability.

### 3.1 Transfer function

Obviously, the more complicated the transfer function the more complex the filter structure and the more arithmetic operations are necessary. That is why the complexity of the transfer functions has to be as high as necessary but at the same time as low as possible. For example, in the case of our traffic surveillance application it is assumed that every object moving from top to bottom in the image plane is a vehicle. Other objects moving in that direction are not expected, which reduces the requirements for the chosen transfer function.

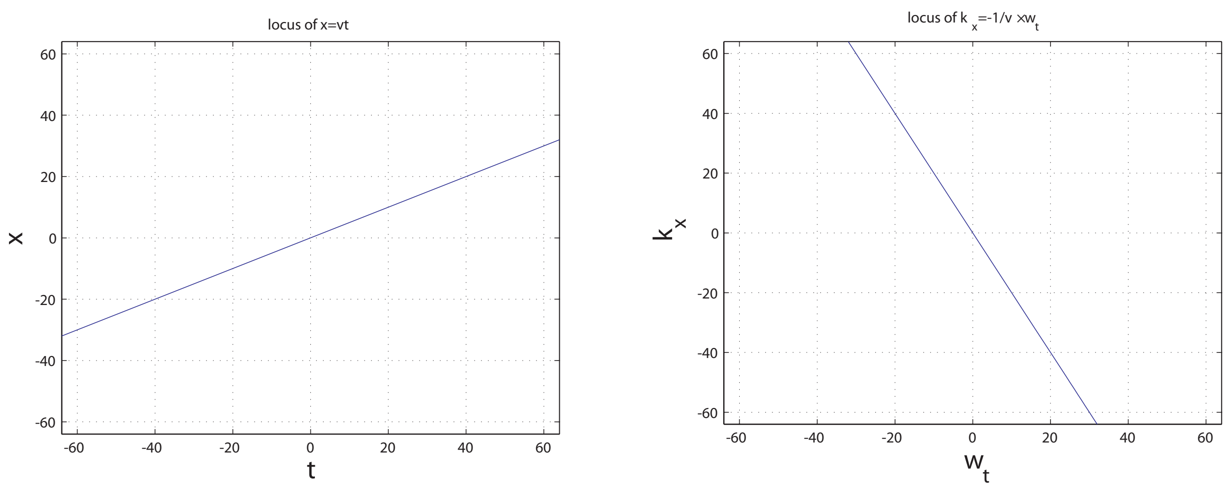
In the following we assume a coordinate system containing one temporal and two spatial axes.

Figure 2 shows a 3-D transfer function suitable to filter objects moving from top to bottom in the image plane (detailed information on the corresponding 3-D transfer function design can be found in [7]). Note that no velocity restrictions exist in this case.
We define this transfer function as $H(j\boldsymbol{\omega})$, with

$$\boldsymbol{\omega} = \begin{pmatrix}\omega_t\\ \boldsymbol{k}\end{pmatrix} = \begin{pmatrix}\omega_t\\ k_x\\ k_y\end{pmatrix}. \tag{9}$$

$H(j\boldsymbol{\omega})$ can be described using the unit step function

16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP




(a) (b)

Figure 1: (a) the locus of $f(x)$ and (b) the right side the respective nonzero points of $S_c(j\omega_t, jk_x)$.

(Heaviside function) $u(\omega)$:

$$u(\omega) = \begin{cases} 1\, if\ \omega > 0 \\ \frac{1}{2}\, if\ \omega = 0 \\ 0\, else, \end{cases} \tag{10}$$

leading to

$$H(j\boldsymbol{\omega}) = u(-k_y)\cdot u(\omega_t) + u(k_y)\cdot u(-\omega_t). \tag{11}$$

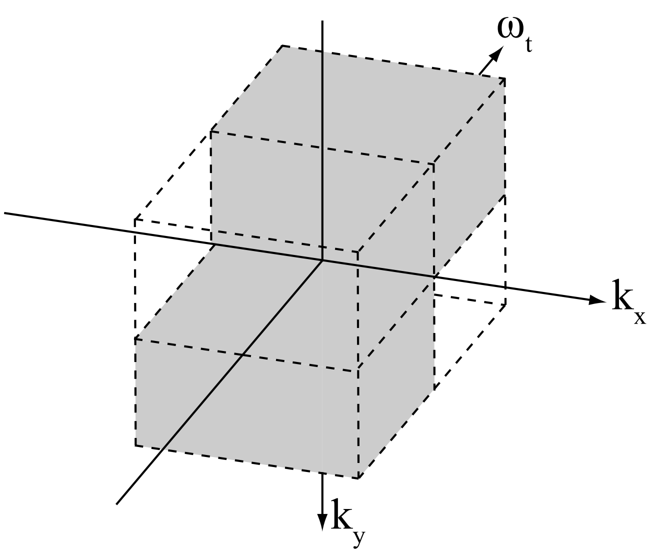



Figure 2: 3-D transfer function suitable to filter objects moving from top to bottom in the image plane.

### 3.2 Quadrant fan filter structure

It can be easily observed from eq. 11 that $H(j\boldsymbol{\omega})$ is independent from $k_x$. Therefore, we can write

$$H(j\boldsymbol{\omega})) = H(j\begin{pmatrix}\omega_t \\ k_y\end{pmatrix}). \tag{12}$$

Thus a 2D quadrant fan filter in y-t-direction can be used to approximate the 3-D transfer function. Figure 3 shows the wave flow diagram of the WDF we use for this purpose with the corresponding coefficients and a plot of the magnitude of its transfer function. A detailed treatment on (discrete) fan filter design can be found in and [8][9].
The resulting filter consists of five three-port adaptors, and as the depiction of the inner structure of this type of adaptor (see figure 3(c)) shows, each of the five adaptors consists of two multipliers and four adders, leading to a total number of 10 multipliers and 20 adders for the whole filter.

## 4. EXPERIMENTS

The filter presented in section 3 has been both simulated using Simulink Fixed Point and implemented on a Xilinx Spartan 3E FPGA Board. The input image sequence for our system has been captured by a CMOS camera for automotive purposes installed on a pedestrian bridge over a highway in Germany. Vehicles approach the camera and stay in the camera's field of view starting from the horizon until they almost reach the bridge (figure 4(a) shows an example for a recorded image). With reference to the image plane, the vehicle objects appear at the top of the image and move on a straight line to the bottom. Naturally, their size on the image plane increases while they come closer.
Figure 4(b) shows the reconstructed filter output of the sample frame. It can be easily observed that most of the non-moving objects in the scene have disappeared. After thresholding (figure 4(c)) we perform simple morphological closing to complete the object detection process. As shown in figure 4(d), the detection result is very promising and the change of object size does not have a negative effect on the detection result.

When it comes to optimizing the FPGA-implementation, the coefficient-robustness of the used filter (and WDFs in general) turns out to be a big advantage: the coefficient-wordlength affecting the 10 multipliers can be cut down extensively with only minor loss of detection performance. In the case of the presented example, it has been possible to cut

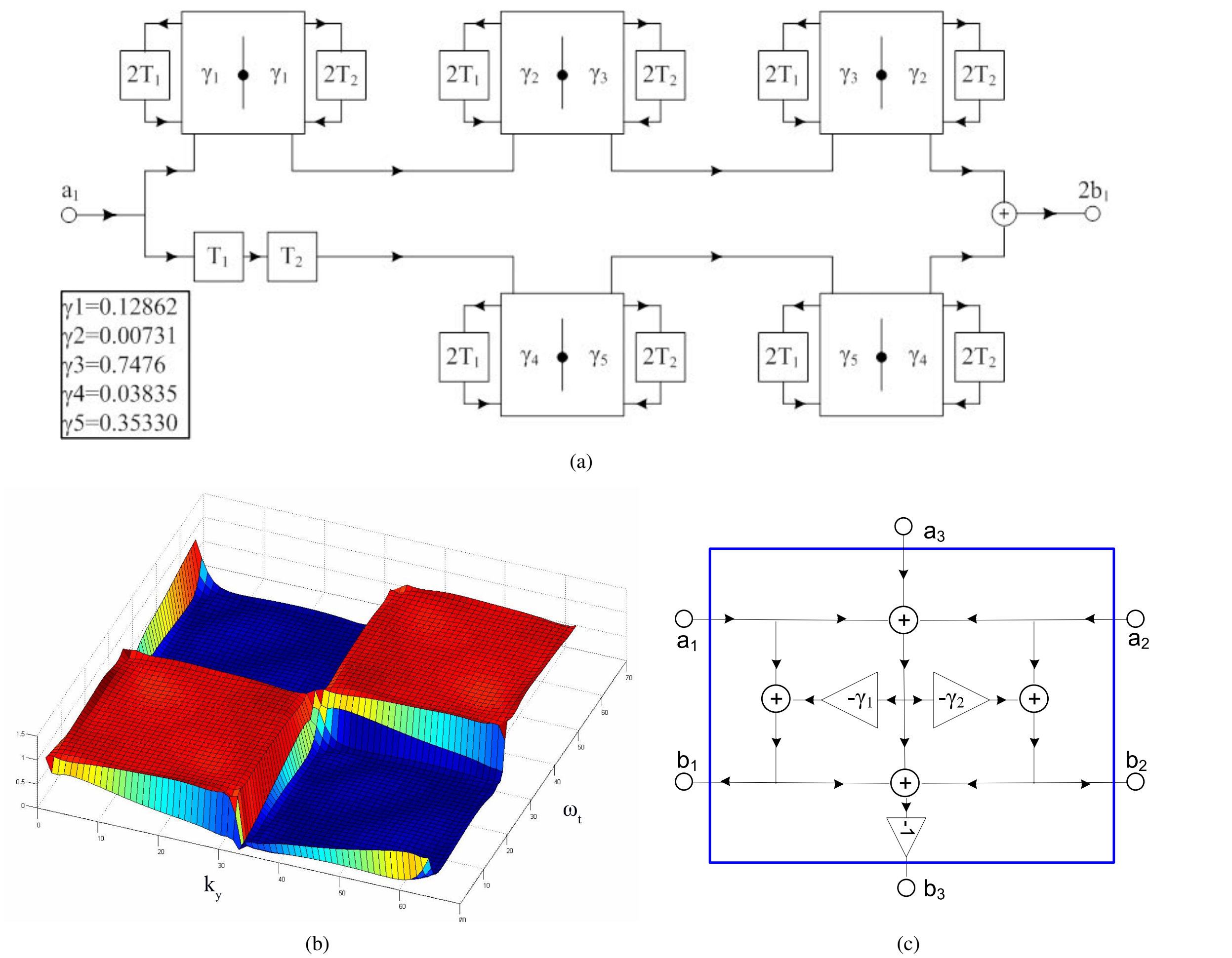



Figure 3: (a) Wave flow diagram of the used 2D quadrant fan filter, (b) the magnitude of its transfer function, (c) a serial three-port adaptor.

down the coefficient-wordlength to 6 Bits (including 3 fractional Bits) and still receive an acceptable result.

## 5. CONCLUSIONS AND FUTURE WORK

In this paper, a detection system for vertically moving objects in video sequences showing realworld traffic scenes using multidimensional linear filters has been presented. A 3-D transfer function suitable for this task has been designed based on the theoretical work of Runze on the design of three-dimensional digital velocity filters [1] and the considerations we have presented in [7] and [10]. It has been implemented using 2D quadrant fan filters.

Experiments prove velocity filtering by means of these filters to be an adequate approach for extraction of movement information from visual scenes. Implemented on an FPGA-platform, the presented system meets real-time requirements.
These results form the basis for further research on the applicability and potential of linear multidimensional filters in real-time driver assistance applications. The objective is to benefit from the advantages multidimensional systems theory offers with respect to system description, design, stability etc.

In the next step, our main focus will be set on the development of a complete object detection SoC. Optimization and hardware-effective implementation of more specialized transfer functions capable to accomplish more complex detection task will be another future research topic. Yet another interesting aim is the solution of one of the most challenging problems in the field of automotive applications using single camera vision: distinguishing between object movement and movement solely caused by ego motion (i.e. host vehicle motion)–without additional sensors.
Overall, the velocity filters presented in this paper show good performance for motion-based object detection with the potential of leading to new and effective methods in driver assistance system development.

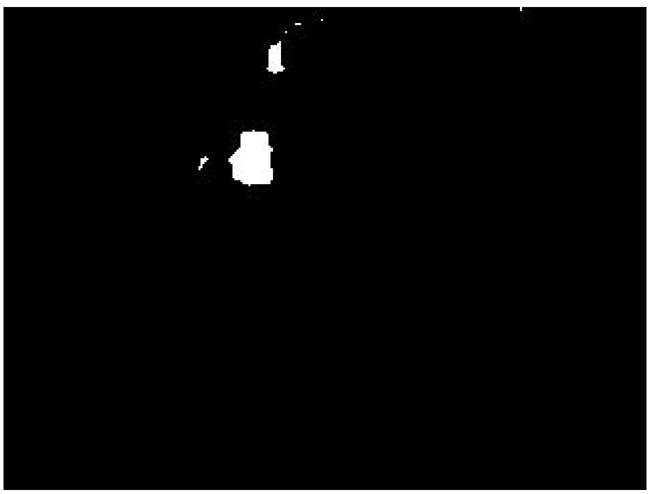
## REFERENCES

[1] G. Runze, *Entwurf dreidimensionaler digitaler Geschwindigkeitsfilter zur Separation unterschiedlicher Geschwindigkeitskomponenten eines Signals*. PhD thesis, Universitt Erlangen-Nuernberg, 2005.

[2] W.-S. Lu and A. Antoniou, *Two-dimensional Digital Filters*. CRC, 1992.

(a)

(b)



(c)

(d)

Figure 4: (a) Frame of an image sequence recorded by traffic surveillance camera on a highway bridge in Germany. (b) Reconstructed filter output of the sample frame. (c) Filter result after thresholding. (d) Final detection result.

[3] N. R. Bartley and L. T. Bruton, “The enhancement and tracking of moving objects in digital imagesusing adaptive three dimensional recursive cone filters,” *IEEE Transactions on Circuits and Systems*, vol. CAS-33, pp. 604–611, 1986.

[4] N. R. Bartley and L. T. Bruton, “The design of highly selective adaptive three dimensional recursive cone filters,” *IEEE Transactions on Circuits and Systems*, vol. CAS-34, pp. 775–781, 1987.

[5] A. Fettweis, “Wave digital filters: theory and practice,” *Proceedings of the IEEE*, vol. 74, no. 2, pp. 270–327, 1986.

[6] A. Fettweis and K. Meerkoetter, “Suppression of parasitic oscillations in wave digital filters,” *IEEE Transactions on Circuits Systems*, vol. CAS-22, pp. 239–246, 1975.

[7] S. Schauland, J. Velten, and A. Kummert, “3d velocity filters for the detection of moving objects in image sequences,” in *Proc. 15th Conference on European Signal Processing Conference, EUSIPCO, Poznan, Poland*, pp. 326–329, September 2007.

[8] A. Fettweis, “Design of recursive quadrant fan filters,” *Archiv Elektronik und Uebertragungstechnik*, vol. 34, pp. 97–103, March 1980.

[9] M. Bolle, *Entwurfsverfahren fuer mehrdimensionale diskrete Faecherfilter*. PhD thesis, Ruhr-Universitaet Bochum, 1992.

[10] S. Schauland, J. Velten, and A. Kummert, “3d velocity filters for improved object detection in automotive applications,” in *Proc. 10th International IEEE Intelligent Transportation Systems Conference, ITSC, Seattle, WA, USA*, pp. 391–395, September 2007.