

# An Imaging Architecture Based on Derivative Estimation Sensors

Maria Petrou and Flore Faille

Imperial College London, SW7 2AZ, UK

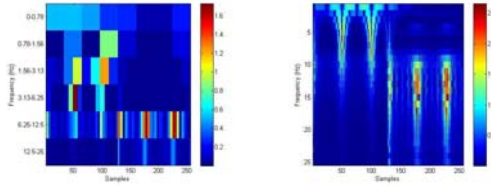
**Abstract.** An imaging architecture is proposed, where the first and second derivatives of the image are directly computed from the scene. Such an architecture bypasses the problems of estimating derivatives from sampled and digitised data. It, therefore, allows one to perform more accurate image processing and create more detailed image representations than conventional imaging. This paper examines the feasibility of such an architecture from the hardware point of view.

## 1 Introduction

In many image processing and computer vision operations we have to use the brightness derivatives of the observed scene. Examples of such operations include all methods that rely on gradient estimation, Laplacian estimation or estimation of higher order fluctuations. Applications include edge detection, multiresolution image representation using the Laplacian pyramid, all methods that rely on anisotropic (or isotropic) diffusion and even wavelet-based methods. In all these cases the derivatives needed are calculated from the discrete data. Discretisation, however, introduces significant errors in the calculation of differentials. An example is shown in figure 1, taken from [26], where the discrete and the continuous wavelet transform of a signal are shown. We can easily appreciate how gross a frequency representation the wavelet transform computed from the sampled version of the signal is.

The importance of processing in the analogue domain has become more evident in the recent years. Splines [28] is a very versatile and powerful tool for representing the discrete data in the continuous domain. Joshi [24] has shown that much improved histogram estimates of the data may be obtained by upsampling and interpolating the data before calculating the histograms. There have also been cases where people try to go back to the continuous domain by emulating “continuous” sensors. In [27] virtual cameras were introduced, with spectral responses in between the discrete spectral responses of actual cameras, in order to improve colour segmentation. In [23] virtual sensors measuring the potential at border points of a 2D vector field were introduced in order to improve the vector field reconstruction using the inverse Radon transform.

It is not necessary, however, for the extra sensors introduced to measure the same quantity as the existing sensors. It is true, that a much denser array of CCDs will sample the brightness of the scene much better than a not so dense array, and it will make the estimated derivatives approach more the true ones



**Fig. 1.** The discrete (left) and continuous (right) wavelet transform of a signal

that refer to the continuous scene. There is another, option, however: it may be possible to use extra sensors that measure the desired quantities directly from the continuous scene. We started by mentioning the significant role differentiation plays in image processing. We would suggest that we may incorporate sensors in our imaging devices that measure the first and second derivatives of the scene directly, as they measure the brightness of the scene. This may be done densely and for different colour bands. The information from the derivative sensors may be used for subsequent processing, as extra information alongside the brightness information, for example in the form of extra image bands, or it may be used to construct a very accurate representation of the scene, much more accurate than a single layer of brightness sensors may do on their own. This interlacing of sensors of different types and different sensitivities appears to sound too complicated, but it may be what nature has implemented for us. In [25], a connection of such a scheme to the architecture of the human retina was made. Some aspects of the retina structure could be explained by the scheme and some predictions were made concerning the photosensitivity of the retinal cells.

In this paper, we attempt to answer the question concerning the feasibility of such an imaging system, from the hardware point of view. We present an overview of existing hardware systems for estimating first and second order derivatives directly from continuous signals. This paper is organised as follows. In Section 2 we present the theoretical underpinnings of the proposed scheme. In Section 3 we consider how such a device might be realised in hardware. We conclude in Section 4.

## 2 Theoretical Considerations

Reconstructing an image from its derivatives requires the integration of the field of derivative values. In this section we consider the process of integration in the continuous and the digital domain and the role the constants of integration play in the process.

### 2.1 Constants of Integration

Assume that we know the second derivative  $d^2 f(x)/dx^2$  of a function  $f(x)$ . What are the values of function  $f(x)$ ? We have to integrate  $d^2 f(x)/dx^2$  twice: first we find the first derivative of the function

$$\frac{df(x)}{dx} = \int \frac{d^2 f(x)}{dx^2} dx + c_1 \quad (1)$$

where  $c_1$  is a constant of integration. Then we have to integrate  $df(x)/dx$  once more to derive function  $f(x)$

$$\begin{aligned} f(x) &= \int \left( \int \frac{d^2 f(x)}{dx^2} dx + c_1 \right) dx + c_2 \\ &= \int \left( \int \frac{d^2 f(x)}{dx^2} dx \right) dx + c_1 x + c_2 \end{aligned} \quad (2)$$

where  $c_2$  is another constant of integration.

Note that the constants of integration appear because we perform an indefinite integral. When we perform definite integrals between pre-specified lower and upper limits, say  $a$  and  $b$ , the result we get is a numerical value of the area under the curve of the integrand between these two limits.

Now, let us consider digital integration. In the digital domain, differentiation is replaced by differencing and integration by summation. The summation, however, is between specific values of the summation index, and so it really corresponds to the *definite* integration of the continuous domain. What in the digital domain corresponds to the *indefinite* integration of the continuous domain is the recovery of the values of the differenced function at all sample positions, by propagating a starting value. We shall explain this with a specific example.

Assume that the true values of a function in a succession of sampling points are:

$$x_1, x_2, x_3, x_4, \dots, x_N \quad (3)$$

Assume that we are given only the first difference values at each of the sampling points, defined as  $d_i \equiv x_i - x_{i-1}$ :

$$\begin{aligned} &?, d_2, d_3, d_4, \dots, d_N \equiv \\ &?, x_2 - x_1, x_3 - x_2, x_4 - x_3, \dots, x_N - x_{N-1} \end{aligned} \quad (4)$$

Here the question mark means that we do not have the value at the first point due to the definition we used for  $d_i$ . To recover the values of the original sequence, from the knowledge of the  $d$  values, we hypothesise that the first value of the sequence is  $c_1$ . Then, the recovered values are:

$$\begin{aligned} &c_1, c_1 + d_1, c_1 + d_1 + d_2, c_1 + d_1 + d_2 + d_3, \\ &c_1 + d_1 + d_2 + d_3 + d_4, \dots, c_1 + d_1 + d_2 + \dots + d_N \end{aligned} \quad (5)$$

This process corresponds to the indefinite integration of the continuous case, with constant of integration the guessed original value  $c_1$ .

There are three important observations to make.

- Without the knowledge of  $c_1$  it is impossible to reconstruct the sequence.
- To recover the value at a single point we need to add the values of several input points.
- As the sequence is built sample by sample, any error in any of the samples is carried forward and is accumulated to the subsequent samples, so the  $N$ th sample will be the one with the most erroneous value.

There are two conclusions that can be drawn from the above observations.

- Such reconstructions cannot be too long, as very quickly the error of reconstruction accumulates and the reconstruction becomes useless. So, for the reconstruction of a long sequence, one has to consider many small sequences in succession, and possibly with overlapping parts.
- If one has a series of sensors that return the local difference value of the observed scene, one needs another series of sensors that return the value of  $c_1$  every so often in the sequence, ie at the beginning of every small reconstruction sequence.

Next, assume that the array of sensors we have does not measure the first difference of the sequence, but the second difference,  $dd_i \equiv d_i - d_{i-1}$ . Then we must apply the above process of reconstruction once in order to get the sequence  $d_i$  and then once more to get the  $x_i$  values. Note that this implies that we must have a series of sensors that every so often in the long sequence of  $dd_i$  will supply the starting constant we need, which in this case is denoted by  $c_2$ . This constant is actually a first difference, so these sensors should measure the first difference at several locations.

## 2.2 The Basic Idea of the Imaging Device in 1D

A device that functions according to the principles discussed above, has to consist of five layers, as shown in figure 2.

The function of this structure effectively repeats twice: below the dashed line we have the first integration, outputting above the dashed line the values of the first difference it computes, and above the dashed line we have the second integration, integrating the first differences it receives and outputting the signal values.

## 2.3 Extension to 2D

The analysis done in the previous two sections is in 1D. However, images are 2D. This has some serious implications, particularly for the  $c_2$  sensors.

From the mathematical point of view, once we move to 2D, we are dealing with 2D integrals, not 1D. A 2D integration implies spatially dependant constants of integration. For a start, a 2D function  $f(x, y)$  has two spatial derivatives,  $\partial f/\partial x$  and  $\partial f/\partial y$ . Let us assume that we know both of them and we wish to recover function  $f(x, y)$  by integration. Integrating the first one of them will yield

$$f(x, y) = \int \frac{\partial f}{\partial x} dx + c_x(y) \quad (6)$$

where  $c_x(y)$  is a function of  $y$ , which, as far as integration over  $x$  is concerned, is a constant. Differentiating result (6) with respect to  $y$  should yield  $\partial f/\partial y$ , which is known, and this can help us work out constant  $c_x(y)$  as a function of  $y$ .

There is an alternative route to work out  $f(x, y)$ . Integrating the partial derivative with respect to  $y$  we get

$$f(x, y) = \int \frac{\partial f}{\partial y} dy + c_y(x) \quad (7)$$

where  $c_y(x)$  is a function of  $x$ , which as far as integration over  $y$  is concerned, is a constant. Differentiating result (7) with respect to  $x$  should yield  $\partial f / \partial x$ , which is known, and this can help us work out constant  $c_y(x)$  as a function of  $x$ .

Obviously, both routes should yield the same answer. In the digital domain, this corresponds to reconstruction of the 2D signal either line by line or column by column. So, let us assume that the true values of the 2D digital signal are  $g_{ij}$ . However, we do not have these values, but we are given instead the first differences of the digital signal along both directions. So, we assume that we have  $dx_{ij} \equiv g_{ij} - g_{i-1,j}$  and  $dy_{ij} \equiv g_{ij} - g_{i,j-1}$ . We can construct the signal column by column as follows. First column:

$$\begin{aligned} g_{12} &= dy_{12} + c_y(1) \\ g_{13} &= dy_{13} + dy_{12} + c_y(1) \\ &\dots \end{aligned}$$

Second column:

$$\begin{aligned} g_{22} &= dy_{22} + c_y(2) \\ g_{23} &= dy_{23} + dy_{22} + c_y(2) \\ &\dots \end{aligned}$$

And similarly for the rest of the columns. This is shown in figure 3a. In a similar way, the signal may be reconstructed along rows. First row:

$$\begin{aligned} g_{21} &= dx_{21} + c_x(1) \\ g_{31} &= dx_{31} + dx_{21} + c_x(1) \\ &\dots \end{aligned}$$

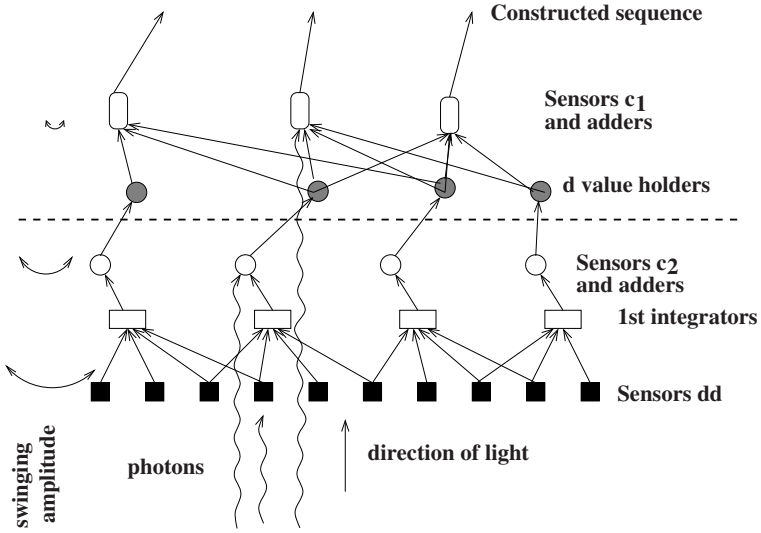
Second row:

$$\begin{aligned} g_{22} &= dx_{22} + c_x(2) \\ g_{32} &= dx_{32} + dx_{22} + c_x(2) \\ &\dots \end{aligned}$$

And similarly for the rest of the rows. This is shown in figure 3b.

Of course, these reconstructions should be equivalent, ie one expects that  $g_{22} = dy_{22} + c_y(2) = dx_{22} + c_x(2)$ . One may also reconstruct the signal by using a combination of rows and columns, and again, the reconstruction should be the same irrespective of the path followed. This is shown in figure 3c.

There are two problems with the above analysis: in practise the alternative reconstructions are never identical due to noise. This is something well known



**Fig. 2.** Sensors *dd* measure the second derivative, while sensors *c<sub>2</sub>* the first derivative (the constant of integration for the first integration) and sensors *c<sub>1</sub>* the value of the function (the constant of integration for the second integration)

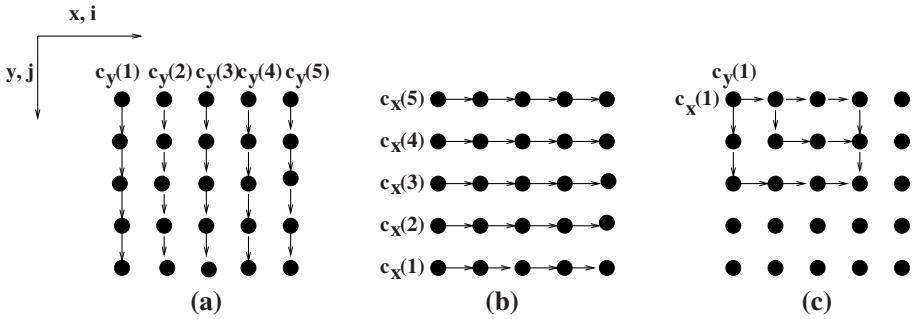
from digital image processing. The other problem is the use of two directions which creates an anisotropic grid, as there are two preferred orientations. Along these two orientations, the samples used are at a fixed distance from each other. However, if we consider samples that are aligned along the diagonal of these two orientations, their distance is  $\sqrt{2}$  times that of the samples along the preferred orientations. This anisotropy is not desirable.

#### 2.4 The Basic Idea of the Proposed Architecture in 2D

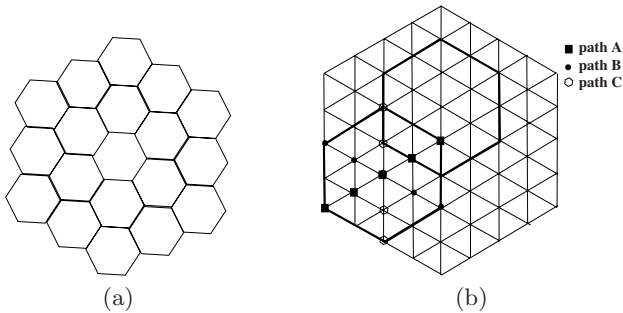
The above approach is compatible with the conventional CCD sensors that consist of rectangular cells, ie rectangular pixels. The cones in the fovea region of the retina, however, have a hexagonal structure, as shown in figure 4a. At first sight this does not look very useful. However, instead of considering the cells, consider their centres as the sampling points of a grid. The nodes in the grid shown in figure 4b are the points where the reconstruction has to take place.

This sampling grid at first sight does not appear hexagonal, but rather based on equilateral triangles. However, several hexagons of various scales can be perceived here.

Imagine now, that we have a device centred at the centre of one of these hexagons. Imagine that the device vibrates along the paths shown. Imagine that this device hangs from a vertical nail above the centre of the hexagon, and consists of three types of sensor hanging from the same string: the bottom one measures second differences, the middle one first differences, and the top one just values. As the string swings like a pendulum, the bottom sensor swings



**Fig. 3.** Reconstruction from first difference values in 2D can proceed along columns (a), or rows (b), or along any path (c). The answers should be equivalent.



**Fig. 4.** (a) The arrangement of cells in the mammalian retina. (b) The hexagonal sampling grid.

more, the middle less and the top not at all (see left of figure 2). This will be consistent with the notion that the second difference sensor needs to see larger part of the scene to do its job than the first difference sensor, while the fixed sensor does not need to swing at all to do its job. Note: it is mathematically impossible to calculate any derivative if you consider only a single sample. So, a device like the one shown in figure 2 swinging along one direction, will allow the reconstruction of the signal along that direction for several sampling points. The amplitude of the swing and the range of reconstruction performed by each single set of sensors are two different things. The amplitude of the swing is for measuring locally what is needed for the reconstruction. Swinging along another direction, will measure the first and second differences along that direction, and the signal will be reconstructed along that direction, by using the propagation techniques we discussed in the 1D case.

There are many advantages of this approach: the reconstruction grid is isotropic; we have no preferred directions; the hexagons fit nicely with each other at all scales; the reconstruction along the lines of one hexagon can be complemented by the reconstruction along the lines of other hexagons that may be directly underneath other sets of sensors hanging from our swinging strings; overlapping

reconstructions are expected to add robustness and super-acuity (ie resolution higher than the sampling distance as determined by the spacing of the sensors); the reconstruction is expected to be complete and very accurate.

### 3 Hardware Considerations

The proposed imaging framework consists of three types of sensor: brightness sensors and sensors that measure spatial first and second order derivatives. These spatial derivatives are obtained by moving the photosensors and computing temporal first and second order derivatives. The photosensor motion is fast and of small amplitude. This is inspired from the microsaccades in the human visual system. After the data acquisition, the image may be reconstructed by using the values of the derivatives, as well as the brightness values.

In a hardware realisation of such an imaging device, brightness values would be acquired with photodiodes, which transform the energy received by photons into electrical charges. Detector motion could be realised by moving the optic or a masking element in front of the photoreceptors, using micro-sized actuators, like the piezoelectric actuator presented in the micro-scanning device in [8]. As far as imaging technology is concerned, CMOS should be used because charges are transformed into voltages or currents directly at the pixel. This allows one to start the processing (the derivative estimation in our case) before any read-out is performed. However, CMOS technology suffers from higher noise sensitivity than CCD cameras. Typical problems are fixed pattern noise, which results from transistor mismatches on the chip, and higher noise sensitivity in dark lighting conditions when the chip has a low fillfactor (ratio of the photosensitive chip area over the total chip area). Image quality is typically improved using processing like the double sampling technique against fixed pattern noise and by keeping the fillfactor as high as possible [16].

In our application, brightness values, as well as first and second order derivatives must be measured with a good precision in order to achieve good image reconstruction quality. Brightness sensors can be implemented like typical pixels in a CMOS camera, with double sampling to reduce fixed pattern noise (see e.g. [16]). For the first and second order derivative sensors, the photodiode signal must be processed to estimate the temporal derivatives. To keep the fillfactor high and to allow a high density of derivative sensors, the circuit for derivative estimation should be as small as possible if processing is performed in situ. Derivatives will not be affected by fixed pattern noise.

Three possible technologies can be used to estimate temporal derivatives from the photodiode signals: fully digital processing, discrete time analog processing and fully analog processing. Digital processing offers the best robustness against electronic noise but it requires a large circuit size and the power consumption is high. On the contrary, a fully analog solution allows a small and energy-efficient circuit, at the cost of a higher sensitivity to electronic noise and parasitic effects caused e.g. by transistor mismatches. Discrete time analog processing offers an intermediate solution: photodiode values are sampled at regular time intervals



but they are not digitised. Processing is performed with analog circuits. The three subsections that follow describe existing systems that estimate derivatives using one of these three technologies. Most of these systems deal with imaging sensors or audio applications. In particular, many VLSI based motion sensors can be found, because the high volume of the data generated by such applications can be handled more easily by parallel hardware, like a VLSI chip, than by standard serial processors. We could not find any existing systems computing temporal second order derivatives. However, they can be built similarly to the first order derivatives for fully digital processing and for discrete time analog processing. For analog processing, they can be obtained by putting two first order derivative circuits in series.

### 3.1 Digital Circuits

In fully digital circuits, signals are sampled in time and digitised during acquisition. As a result, signals are affected by aliasing and discretisation noise. Digital circuits are however a lot less sensitive to electric noise and distortions because values are encoded as binary and because transistors are only used in their saturated and blocked states. Digital circuits have a high power consumption and are most of the time large circuits. Digital processing is the method of choice when data acquisition and processing can be performed at different time instants and on different systems (e.g. the processing is performed on an external processor).

Here, however, processing (derivative estimation) should be performed on the chip, if possible near the photodiode. Thanks to the constant size reduction of the electronics (Moore's law), the first intelligent cameras (also named artificial retinas) with in situ digital processing have become a reality. One such system is described in [20]. Each pixel is composed of a photodiode, a simple analog to digital converter and a tiny digital programmable processing element. It comprises about 50 transistors. The processing element has a local memory and can perform boolean operations. The chip has a fillfactor of 30%. Due to the limited number of transistors, brightness values can only be encoded with few grey levels (8 grey levels in the examples given in the paper), resulting in low precision. A few simple applications have been implemented on the system: motion detection, segmentation and shape recognition. However, our application requires high precision. Even though the control signals, necessary for programmability, can be avoided in our case, it would be difficult to obtain the necessary precision in the circuit space available near each pixel. Therefore, fully digital processing seems not appropriate for our application at the moment, except if derivative computations are performed externally (on a separate processor). However, this defeats the purpose of our approach, which is the direct and in situ estimation of the derivatives.

### 3.2 Discrete Time Circuits

For these systems, the photodiode signals are sampled at given time instants and they are stored in local analog memories or sample-and-hold units. Processing is

performed using the stored pixel values on analog circuits. Like for the fully digital circuits, the system operates in two phases: sampling, followed by processing. Such circuits represent an intermediate solution between fully analog and fully digital systems. Like digital systems, they might be affected by temporal aliasing (which corresponds to spatial aliasing in our case, because spatial derivatives are estimated using sensor motion and temporal derivatives). However, brightness values are not digitised, so there is no discretisation noise. Control signals for the sampling should be generated and synchronised to the sensor motion. Like in analog systems, computations are less robust to electronic noise and distortions than in fully digital systems. Noise and distortions are caused, for example, by active filter elements, like amplifiers, by transistor mismatches and other element inaccuracies, by parasitic capacitors, by leakage currents, etc.

Discrete time intelligent camera systems started being designed when CMOS cameras became popular. Early systems only added local memories to pixels (analog memories or sample-and-hold units, composed, for example, of two transistors, used as switches, and a capacitor to store the value [16]). Processing was performed during the image read out phase, just before analog to digital conversion [6,16]. This is used in [6] to perform spatial convolutions with  $5 \times 5$  kernels with coefficients in  $\{-1, 0, 1\}$ . In [16], the designed camera can deliver either normal images or differences between consecutive frames. More recent systems, based on the same organisation (separated photosensitive and processing areas), can perform more elaborate processing, like optical flow computation in [18] or saliency detection in [9]. The separation of photosensitive and processing areas allows high fillfactors (e.g. 40% in [9]). However, processing is performed row by row, which limits the achievable frame rate. These last two systems [18,9] compute temporal derivatives by subtracting pixel values at two time instants,  $t$  and  $t + \Delta t$ :  $d(t) = I(t) - I(t - \Delta t)$ , where  $\Delta t$  is different from the sampling period of the camera. In [18], a calibration scheme is included to suppress the distortions caused by mismatches between the p and n elements.

A more recent system, where temporal derivatives are estimated by differentiating the sensor values at two different time instants, is given in [21]. Sample-and-hold units are implemented with switched capacitor technology, which allows high accuracy and programmability of the capacitances. In addition to differentiation, sample values are also amplified with a simple inverting amplifier. This system implements an audio application, with which sound sources can be localised. Space is therefore not an issue for them, unlike for imaging systems. A recent intelligent camera is presented in [4], in which processing is performed in situ. Each pixel is composed of a photodiode and 38 transistors. It contains two analog memories (to allow the acquisition of the next image during the processing of the current image) and a simple analog arithmetic unit (to perform spatial convolution with integer based kernels). As a result of this organisation (in situ processing), a high framerate of several thousands of images per second can be achieved even when image processing is performed. The fillfactor is 25%.

This last example shows that analog computation of derivatives could be performed in situ if a discrete time circuit is chosen for the implementation of

our imaging sensor. None of the papers gives any indication about noise level or whether distortions influence the precision of the results. However, the time discretisation allows one to reduce the circuit complexity: derivation is replaced by a simple subtraction of two sampled values. The results should therefore be quite accurate. Second order derivatives can be implemented similarly by sampling data at three time instants and by performing two additions and one subtraction:  $dd(t - \Delta t) = I(t - 2\Delta t) + I(t) - (I(t - \Delta t) + I(t - \Delta t))$ . The signals required to control the sample-and-hold units could be generated from the signals used to control the actuators, therefore synchronising sampling and sensor motion.

### 3.3 Analog Circuits

In fully analog systems, analog signals are processed in continuous time. The circuit usually consumes low power and is of small size. However, analog processing is the most sensitive to electronic noise and distortions, which are caused by active elements, like amplifiers, element inaccuracies, transistor mismatches, frequency response of the circuits, parasitic capacitances, current leakage, etc. The difference between idealised model and practice is the biggest for analog circuits due to the complexity of the underlying physical phenomena. Analog circuits have the advantage of not being limited by the camera frame rate, as they are working in continuous time. Therefore they are not affected by aliasing or discretisation noise. The classical method of computing derivatives in an analog circuit is through the use of a capacitor for which  $I = C \frac{dV}{dt}$ . However, an ideal differentiator would require high power for high frequency signals, as the ideal transfer function of a capacitor is:

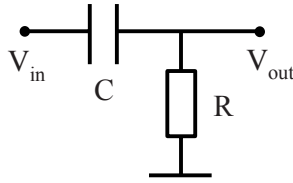
$$H(j\omega) = \frac{I_{out}}{V_{in}} = j\omega C. \quad (8)$$

This is physically unrealisable and in a real system, some resistance will always limit the current. This results in the well-known RC filter shown in figure 5. The derivative of a signal can be measured as the voltage through the resistor. The transfer function for this system is:

$$H(j\omega) = \frac{V_{out}}{V_{in}} = \frac{jRC\omega}{1 + jRC\omega} = \frac{j\tau\omega}{1 + j\tau\omega}. \quad (9)$$

For low frequencies ( $\omega \ll 1/\tau$ ), the RC filter is a good approximation of the ideal differentiator:  $H(j\omega) \approx j\tau\omega$ . For high frequencies ( $\omega \gg 1/\tau$ ), the transfer function becomes approximately 1. This simple RC filter is the basis of all analog circuits used to estimate derivatives.

In practice, amplification may be required or active resistors (i.e. non-linear resistance circuits based on transistors) may be necessary to obtain the desirable time constant  $\tau = RC$  with the manufacturable circuit elements (see e.g. [19,17]). Therefore, real circuits are more complicated than a simple RC filter. Many applications do not require an accurate estimation of the derivatives. Therefore,



**Fig. 5.** RC filter used to estimate first order derivatives

many systems only detect changes in the input signal using high pass filters which are easier to realise than a good differentiator. This solution is used for example in [17,2,1,19,5]. In [17], it is shown that such an approximation can increase the robustness of a circuit (see e.g. the developed hysteretic differentiator). This approach is enough for applications aiming only at detecting motion in visual input. It is, however, not appropriate in our case, because precise derivatives are necessary for image reconstruction.

Another approach to estimate the current going through a capacitor using transistors is proposed in [12]. Instead of measuring the voltage through a resistor, the current flowing through a capacitor is copied using transistors. The photosensitive circuit and the temporal derivative circuit proposed in [12] have been extended and used in applications by the authors in [10,13] and by others in [7,3]. The temporal derivative circuit contains an amplifier and a feedback loop with a capacitor and the circuit for measuring the current going through it. As the current through the capacitor is measured, the output is, in theory, proportional to the derivatives of the light energy hitting the photodiode, as shown in [13]. The system delivers two currents: one representing negative changes in the input and one representing positive changes in the input. It is based on a functional model of the auditory hair cells that sense the motion of the basilar membrane in the cochlea. The results are quite noisy, mainly due to the amplifier [13,3]. The non-linearities and inaccuracies in the transistors also cause an asymmetric response: the signal representing negative changes reacts slightly differently from the signal representing positive changes for changes of the same amplitude [7]. Despite the noise level, the results allow one to estimate velocity in 1D in [10,13], in 2D [7] and to compute optical flow at each pixel in [3].

The same principle (measuring the current through the capacitor using transistors) is used in [14,11,15,22]. In these more recent papers, much simpler circuits are used. In particular, the amplifier is realised with fewer transistors, probably because less gain is necessary or because the transistors are of better quality. The simplest circuits are used in [14,15]. These two applications do not aim at obtaining precise derivatives but only at detecting changes in the input signal. The authors hence did not pay much attention to noise and distortions during circuit design. In [14] the resulting noise level is too high even for their application, so the authors conclude that their differentiator circuit must be improved. Therefore, the circuits in [14,15] are not appropriate for our application. The circuits in [11,22] are designed to estimate the temporal derivative of the photodiode signal as accurately as possible. Both systems use the same circuit element

to read out the current through a capacitor, but the feedback loop to the amplifier is designed slightly differently in the two circuits. The results in [22] seem to be more symmetrical (the responses for positive and negative changes are more similar). However, this might be due to the fact that the article shows fewer results. The results in [11] are of much better quality than the previous papers [13,3]: the noise level is significantly reduced. However, the output signals are still far from ideal. In particular, the asymmetry between outputs for positive and for negative changes is problematic. None of the papers gives any estimation of the noise level, probably because it is influenced by many factors. Both systems in [11,22] can be implemented with one capacitor and less than 10 transistors, which is a very small circuit size.

### 3.4 Discussion

Three different kinds of system can be used to estimate first and second order temporal derivatives from the output of a photodiode: fully digital systems, discrete time analog systems and fully analog systems. Here we gave an overview of existing systems that estimate derivatives using any of these methods. The goal is to find a system which would be suitable to use in order to implement the new derivative based imaging sensor in hardware.

Fully digital systems offer the best robustness to electronic noise and distortions at the cost of high power consumption and large circuit areas. Only one fully digital imaging system with in situ processing could be found in the literature. It had a very limited number of grey values, which results in low precision. As a result, a fully digital in situ derivative estimation cannot be realised with today's technology. So, fully digital systems could only be used if processing would be performed on an external processor like a DSP system.

Fully analog systems have a low power consumption and more importantly they can be implemented with few transistors (in [11,22] first order derivatives are estimated with one capacitor and less than 10 transistors). On the other hand, the signals are sensitive to electronic noise and distortions caused by the non-linearities and parasitic effects in the circuits. The circuits in [11,22] allow in theory to estimate the first order derivatives accurately. The results shown are encouraging in comparison with previous works. However, they are still far from accurate. In addition to the moderate noise level, the circuit responses to positive and to negative input changes of the same amplitude are slightly different. It is therefore not certain whether a fully analog system would be accurate enough to allow a good image reconstruction. Another problem is that no circuit could be found to estimate second order temporal derivatives. These could be estimated by putting two differentiators in series, but this would also amplify noise and distortions, reducing even more the accuracy of the estimated second order derivatives.

Discrete time analog systems represent an intermediate solution. The photodiode signals are sampled at given time instants and stored in an analog memory or sample-and-hold unit. The first and second order derivatives can be estimated by subtracting the stored values. These operations are performed with analog circuits.

The resulting analog circuit is much simpler than a fully analog system, reducing problems like electronic noise and distortions. Such systems have been used recently to develop intelligent cameras (or artificial retinas) which perform not only image acquisition but simple image processing operations as well. The system in [4] shows that as many as 38 transistors can be integrated in each pixel for processing, while keeping the fillfactor at a reasonable value. This would be enough to allow the computations of first or second order derivatives. The signals used to control the actuators could be used to synchronise data sampling and sensor motion. As a result, we conclude that a discrete time analog system would be the most appropriate to be used in order to implement the proposed imaging system.

## 4 Conclusions

The proposed imaging architecture has several advantages over conventional architectures that measure only scene brightness:

- (i) it allows the direct and accurate estimate of the first and second image derivatives directly from the scene;
- (ii) it allows the increase of sensor resolution if the image is upsampled with the use of its local derivative values.

The viability of such a device rests on two fundamental questions.

1) Can we develop sensors that can estimate the first and second derivatives directly from the scene? In this paper we reviewed the current technology and concluded that discrete time analog systems are a promising direction for developing such a device. There are already sensors that can estimate the first spatial derivative of the scene, and although there are no sensors that can estimate the second spatial derivative, we do not think that such a development is too difficult or beyond the state of the art of current sensor technology.

2) Will the outputs of these sensors be more accurate and resilient to noise than the calculations of the derivatives from the sampled data? This question cannot be answered until such a device has actually been realised in hardware.

Actually, both the above questions have to be answered by sensor scientists, as they cannot be answered theoretically. There is no doubt that if the answer is “yes” to both these questions, the image processing that we shall be able to do with such devices will be much more reliable and accurate than the image processing we are doing now.

**Acknowledgements.** This work was supported by a UK Research Councils translation grant.

## References

1. Chong, C.P., Salama, C.A.T., Smith, K.C.: Image-motion detection using analog VLSI. *IEEE Journal of Solid-State Circuits* 27(1), 93–96 (1992)
2. Delbrück, T., Mead, C.A.: Time-derivative adaptive silicon photoreceptor array. In: *Proceedings SPIE*, vol. 1541, pp. 92–99 (1991)

3. Deutschmann, R.A., Koch, C.: An analog VLSI velocity sensor using the gradient method. In: Proceedings of the 1998 International Symposium on Circuits and Systems (ISCAS 1998), pp. VI 649–VI 652 (1998)
4. Dubois, J., Ginhac, D., Paindavoine, M.: Design of a 10000 frames/s CMOS sensor with in situ image processing. In: Proceedings of the 2nd International Workshop on Reconfigurable Communication-centric Systems-on-Chip (ReCoSoc 2006), pp. 177–182 (2006)
5. Etienne-Cummings, R., Van der Spiegel, J., Mueller, P.: A focal plane visual motion measurement sensor. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* 44(1), 55–66 (1997)
6. Funatsu, E., Nitta, Y., Kyuma, K.: A 128 x 128-pixel artificial retina LSI with two-dimensional filtering functions. *Japanese Journal of Applied Physics* 38(8B), L938–L940 (1999)
7. Higgins, C.M., Deutschmann, R.A., Koch, C.: Pulse-based 2-D motion sensors. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing* 46(6), 677–687 (1999)
8. Hoshino, K., Mura, F., Shimoyama, I.: Design and performance of a micro-sized biomorphic compound eye with a scanning retina. *Journal of Microelectromechanical Systems* 9(1), 32–37 (2000)
9. Kimura, H., Shibata, T.: A motion-based analog VLSI saliency detector using quasi-two-dimensional hardware algorithm. In: Proceedings of the 2002 International Symposium on Circuits and Systems (ISCAS 2002), pp. III 333–III 336 (2002)
10. Kramer, J.: Compact integrated motion sensor with three-pixel interaction. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18(4), 455–460 (1996)
11. Kramer, J.: An integrated optical transient sensor. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing* 49(9), 612–628 (2002)
12. Kramer, J., Sarpeshkar, R., Koch, C.: An analog VLSI velocity sensor. In: Proceedings of the 1995 International Symposium on Circuits and Systems (ISCAS 1995), pp. 413–416 (1995)
13. Kramer, J., Sarpeshkar, R., Koch, C.: Pulse-based analog VLSI velocity sensors. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing* 44(2), 86–101 (1997)
14. Landolt, O., Mitros, A., Koch, C.: Visual sensor with resolution enhancement by mechanical vibrations. In: Proc. of the 2001 Conference on Advanced Research in VLSI (ARVLSI 2001), pp. 249–264 (2001)
15. Lichsteiner, P., Posch, C., Delbruck, T.: A 128x128 120dB 30mW asynchronous vision sensor that responds to relative intensity change. In: Proc. of the 2006 IEEE International Solid-State Circuits Conference (2006)
16. Ma, S.-Y., Chen, L.-G.: A single-chip CMOS APS camera with direct frame difference output. *IEEE Journal of Solid-State Circuits* 34(10), 1415–1418 (1999)
17. Mead, C.: *Analog VLSI and Neural Systems*. Addison-Wesley Publishing Company, Reading (1989)
18. Mehta, S., Etienne-Cummings, R.: Normal optical flow chip. In: Proceedings of the 2003 International Symposium on Circuits and Systems (ISCAS 2003), pp. IV 784–IV 787 (2003)
19. Moini, A., Bouzerdoum, A., Yakovleff, A., Abbott, D., Kim, O., Eshraghian, K., Bogner, R.E.: An analog implementation of early visual processing in insects. In: Proceedings of the 1993 International Symposium on VLSI Technology, Systems and Applications (VLSITSA 1993), pp. 283–287 (1993)

20. Paillet, F., Mercier, D., Bernard, T.M.: Second generation programmable artificial retina. In: Proceedings of the 12th annual IEEE International ASIC/SOC Conference, pp. 304–309 (1999)
21. Stanacevic, M., Cauwenberghs, G.: Micropower gradient flow acoustic localizer. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* 52(10), 2148–2157 (2005)
22. Stocker, A.A.: Compact integrated transconductance amplifier circuit for temporal differentiation. In: Proceedings of the 2003 International Symposium on Circuits and Systems (ISCAS 2003), pp. I 25–I 28 (2003)
23. Giannakidis, A., Kotoulas, L., Petrou, M.: Improved 2D Vector Field Reconstruction using Virtual Sensors and the Radon Transform. In: International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, Canada, August 20–24 (2008)
24. Joshi, N.B.: Non-parametric probability density function estimation for medical images, PhD thesis, University of Oxford (2007)
25. Petrou, M.: A new imaging architecture and an alternative interpretation of the structure of the human retina. In: Zaman, H.B., Sembok, T.M.T., van Rijsbergen, K., Zadeh, L., Bruza, P., Shih, T., Taib, M.N. (eds.) Proceedings of the International Symposium on Information Technology, Kuala Lumpur Convention Centre, Malaysia, August 26–29, vol. 1, pp. 9–17, IEEE Cat. No CFP0833E-PRT, ISBN 978-1-4244-2327-9
26. Varnavas, A.: Signal processing methods for EEG data classification, PhD thesis, Imperial College London (2008)
27. Verges-Llahi, J.: Colour Constancy and Image Segmentation Techniques for Applications to Mobile Robotics, PhD thesis, University Politecnica de Catalunya, Barcelona, Spain (2005)
28. Unser, M.: A Guided Tour of Splines for Medical Imaging. In: Plenary talk, Twelfth Annual Meeting on Medical Image Understanding and Analysis (MIUA 2008), Dundee UK, Scotland, July 2-3 (2008)