A device and an algorithm for the separation of visible and near infrared signals in a monolithic Silicon sensor

G. Langfelder*[a], T. Malzbender[b], A. F. Longoni[a], F. Zaraga[a]

[a] Electronics and Information Technology Department, Politecnico di Milano,
via Ponzio 34/5, I-20133, Milano, Italy;
[b] Hewlett Packard Labs, 1501 Page Mill Road, Palo Alto, CA 94304
* mail: giacomo.langfelder@polimi.it; phone +39 02 2399 3425; fax +39 02 2367604

ABSTRACT

The Transverse Field Detector (TFD) is a filter-less and demosaicking-less color sensitive device that easily allows the design of more than three color acquisition channels at each pixel site. The separation of light into different wavelength bands is based on the generation of transverse electric fields inside the device depleted region, and exploits the properties of the Silicon absorption coefficient. In this work we propose such a device for the joint capture of visible and near infrared (NIR) radiation, for possible applications in videoconferencing and 3D imaging. In these applications the detector is used in combination with suitably generated NIR structured light. The information of the fourth acquisition channel, mainly capturing NIR signals, can be used both for sampling NIR light intensity and for subtracting unwanted NIR crosstalk from visible channels thus avoiding the need for the IR-blocking filter. Together with the presentation of a 4-channel sensor, a suitable algorithm for the processing of signals generated in the visible and infrared bands is described. The goal of the algorithm is to minimize the crosstalk of NIR radiation inside the visible channels and, simultaneously, to maintain good color reproduction and noise performance for the sensor, while holding a good sensitivity of the NIR channel up to 900 nm. The analysis indicates that the algorithm reduces the crosstalk of infrared signals inside R, G and B channels from 31%, 12% and 5% respectively to less than 2%. Concerning noise propagation, the worst coefficient of the color conversion matrix (CCM) is -2.1, comparable to those obtained for CCM of Bayer Color Filter Arrays.

Keywords: CMOS sensors, infrared imaging, visible imaging, 3D imaging, digital photography.

1. INTRODUCTION

The combined acquisition of images in the visible (VIS) and infrared (IR) bands of the spectrum has several advantages for a large number of applications. In digital imaging, concurrent infrared acquisition has been proposed as a mean to enhance photographs both in terms of dynamic range (improving contrast and details)[1], and in terms of image quality, avoiding the degradation caused by atmospheric haze[2]. In 3D imaging systems, it is possible to use the IR channel to recover the depth information of the scene or part of it[3,4]. In surveillance systems, using the IR channel allows to differentiate the patterns of objects at different temperatures down to approximately 600 K. Night-and-Day vision systems obtained combining visible and infrared signals have been proposed also for automotive cameras[23]. Nevertheless, the use of two separate imaging sensors and optics for the VIS and IR band (see e.g. ref [3,17]) has so far made unpractical the implementation of a low-cost VIS-IR imaging device.

A solution to the need for two separate sensors can be the use of a single Silicon sensor in which the IR filter (or "hot mirror") is removed, so that near infrared (NIR) and VIS radiation can be captured simultaneously and in some way distinguished (Silicon is indeed inherently sensitive up to 1107 nm). In 2008 Koyama[23] proposed a modified CFA with a fourth channel for infrared acquisition: a subtraction of the NIR signal from visible channels was made to reduce the crosstalk, that however remained still large (9%, 4% and 21% for R, G and B channels respectively). Their device used photonic crystal based filters instead of commonly available organic CFA. In 2009 Lu[5] proposed an optimization algorithm for the design of a Color Filter Array for the joint capture of VIS and NIR radiation: the CFA optimization was done in the Fourier domain, considering the demosaicking issues, and resulted in an uncommon 4x4 CFA including 15 visible/NIR filters and one NIR-only filter. In a recent work[24], Kawada demonstrated the selective deposition of an on-chip IR filter onto RGB channels only of an RGBW CFA. In this way the W channel remained sensitive to NIR radiation

while RGB channels showed usual spectral responses of visible imaging devices. However the selective deposition was not obtained with a standard process but with a dedicated process step of deposition of multiple layers of SiO_2 and TiO_2.

Due to the physical properties of Silicon, the absorption of NIR radiation happens with a deeper penetration length (several μm) than VIS radiation (fractions of μm for blue light, few μm for red light). Similarly to what has been proposed for RGB sensors[7,8,9] using layered junctions, one can thus think to collect radiation from different depths in Silicon to distinguish the different contributions. Following this approach, in 2007 Lyu[6] proposed a device combining a Bayer CFA with a vertically integrated photodiode structure for NIR capture. However, due again to the physical properties of the absorption (exponential decay of the radiation flux propagating into a mean), a complete physical separation of photo-carriers generated by different wavelength bands inside a monolithic Silicon detector is not possible. While this issue is not critical for visible light in RGB sensors (where some overlap of the spectral responses in the R, G and B sub-bands is on the contrary required), it represents a source of crosstalk for a detection system whose aim is to separate information from VIS and NIR bands. No information on the effects of the crosstalk between VIS and NIR channels was however provided in [6] and the combination of RGB and NIR images, having different resolution, was not discussed.

In this work we present a 4-channel implementation of the Transverse Field Detector (TFD), a previously presented Silicon sensor for visible imaging based on the Silicon absorption properties (i.e. without color filters). Carrier separation as a function of the generation depth is obtained through a suitable transverse electric field configuration obtained in the device depleted region. Carriers generated at different depths are in this way driven to different R, G and B surface contacts. Beside these three channels, in an improved version we have built a fourth channel[9]. In this work we show that through a geometrical optimization and a new device design, this fourth channel can be used for NIR acquisition. With respect to other sensors based on the photon absorption properties in Silicon, the TFD has a great advantage in that it requires no specific technology to provide 4 channels. On the contrary, passing from three to four channels for detectors based on layered junctions would require further process steps to design a very deep implant. While this latter idea was proposed in [10], as far as the authors know there has been no published scientific retinue so far. With respect to the device proposed by Lyu[6], in the TFD the resolution of the NIR channel is the same as for the VIS channels.

Together with the presentation of a simulated 4-channel sensor – based on experimental results obtained on a device built for the validation of the working principle – a suitable algorithm for the processing of the information generated in the visible and infrared bands is described. The algorithm takes into account the final quality of the images: its goal is to minimize the unwanted crosstalk of NIR radiation inside the visible channels and, simultaneously, to maintain good color reproduction and noise performance for the sensor, while holding a good sensitivity of the NIR channel up to 900 nm. All these requests depend on the shape of the TFD spectral responses and are often in contrast, so that the best solution is necessarily a compromise.

In particular, the algorithm is based on the mathematical analysis of the absorption of infrared light in the four channels of the device (R, G, B and NIR). Tested on a *Matlab*® simulated TFD device, the algorithm recovers, from the original four channels, three spectral functions (R', G' and B') that well match the ones that would be obtained using an IR blocking filter. The correction is linear and requires a minimum negative coefficient value around -1. In this way the crosstalk of infrared radiation inside the visible channels is corrected, without significant propagation of noise. A further Color Conversion Matrix (CCM) is used for transformation to the standard CIE *XYZ* color space. In particular, the device simulation is optimized for a final application in 3D imaging and videoconferencing, where the infrared light is provided by a structured light path generated by a led with a peak wavelength around 850 nm (see Fig. 1: coded structured IR radiation is considered one of the most reliable techniques for recovering depth information and it is based on the correspondence between coded projected points and imaged points[12,13]). At this wavelength the sensitivity of the IR channel is still larger than the 30% of its peak sensitivity.

Finally a possible implementation of a real TFD detector to be used with this algorithm is shown. The device is designed exploiting the features and respecting the rules of a 130 nm CMOS standard technology, and it is simulated using an electron device simulator, *Dessis* from *ISE-TCAD*®. The device biasing respects the critical technological issues in terms of leakage and punch-through currents. The obtained photoresponses well match the previous simulations and demonstrate the feasibility of such a device.
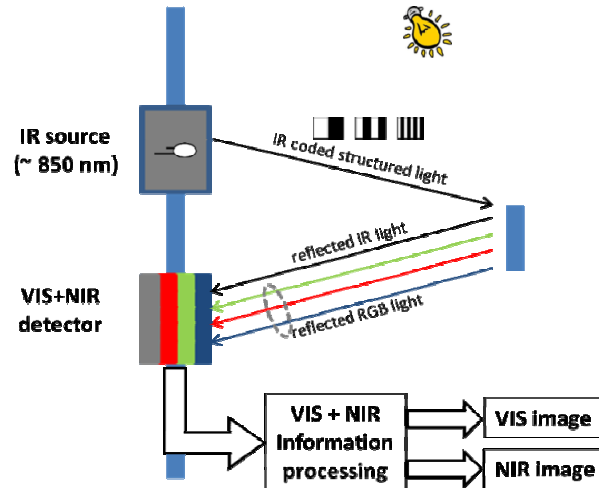
Figure 1. Schematic representation of the proposed system.

2. APPLICATIONS

2.1 3D scene depth imaging

Jointly imaging infrared as well as visible light has numerous applications as well as a rich history. In one method, Structured Light[12,15] 3D scene depth is inferred from a pattern that is projected onto the scene and imaged with a camera. Typically such a pattern is projected using infrared illumination to avoid being visible and therefore distracting to a user. Often a sequence of separate projection patterns are projected and imaged on a static scene, yielding good robustness[13]. However, it is possible to project and decode using single, static image as well, resulting in the capability of real-time inference of depth in a potentially dynamic scene[14]. Such structured light techniques can infer 3D scene structure and are often assisted by a separate visible light camera for applying color and texture to the geometry. Since separate imaging chips and optics are employed for the color and infrared imaging, alignment between the two datasets is necessary. Since such alignment depends on the 3D depth of the scene, it is problematic and ill-defined for regions suffering from occlusion in one of the images but not the other. Using the same sensor for both the structured light and visible imagery makes this merging process trivial.

2.2 Foreground/background segmentation

The use of infrared illumination and imaging combined with visible imaging also has application to real time foreground/background segmentation[16,17]. In this application, infrared illumination is provided to cover the foreground of a scene, such as the participants of a videoconferencing session, but not their background. This allows algorithms such as Graph-Cuts or Relaxation Labeling to robustly segment foreground of the IR image, which then can be applied to the visible light image. As in the case with structured light, alignment of the IR and visible imaging can be done either optically, where it is expensive and results in a loss of brightness, or algorithmically when co-axial IR and visible cameras are not practical. Deploying an integrated IR/visible sensor avoids these merging difficulties.

2.3 Videoconferencing

IR imaging has also been demonstrated to be useful for videoconferencing situations where lighting is less than ideal[18]. In this application, NIR LEDs placed around the bezel of a monitor provide NIR illumination from several known but varying lighting directions. Using photometric stereo[19], surface normals are computed in real time on a GPU from the IR sequence, and these are used for relighting simultaneous video streams from a coaxial visible-light camera. The requirement of having separate cameras that are coaxial can be avoided by having a joint IR/visible sensor.

2.4 Time-of-flight approach

Lastly, several 3D cameras are available that recover real-time 3D depth using time-of-flight methods as opposed to the structured light methods described above. Canesta[19], 3DV, and PMD[20] are all examples of manufactures of such devices.

By measuring the time it takes for light to travel from its source, reflect off an object and return to the detector, per-pixel measurements of depth can be made. These measurements are made at infrared wavelengths to avoid distracting the user. Although making such precise timing measurements require that the sensor pixels be somewhat larger and more complex than in the conventional case of simple brightness imaging, these approaches would also benefit for having an integrated on board visible light mosaic. The task of real-time merging such depth camera data with a visible light camera stream for background removal in teleconferencing applications[22] is particularly challenging and would also be rendered trivial with an integrated approach.

3. COMBINED VISIBLE AND INFRARED PIXEL DESIGN

Color separation in the Transverse Field Detector (TFD) is based on the possibility of collecting, at different surface electrodes, the carriers generated at different depths in a depleted region under the semiconductor surface. This kind of carrier collection can be performed by superposing, to the depletion electric field (perpendicular to the device surface), transverse components (parallel to the device surface) generated by a suitable biasing scheme of the surface electrodes. Due to the different absorption depth of different wavelengths, the proposed structure allows the color discrimination[8]. A four-channel version of this device is designed with four *N-type* reverse biased junctions on a low-doped *P-type* active layer[9] (Fig. 2a).

Each surface junction collects the carriers generated down to a different depth x_i. A suitable choice of the junctions biasing voltages determines the transverse electric field configuration that causes carriers generated by blue, green and red radiation to be collected mainly by electrodes #1, #2 and #3 respectively. Contact #4 collects from the remaining depth and thus the larger is the thickness of the Si active layer, the larger is its response in the NIR band.
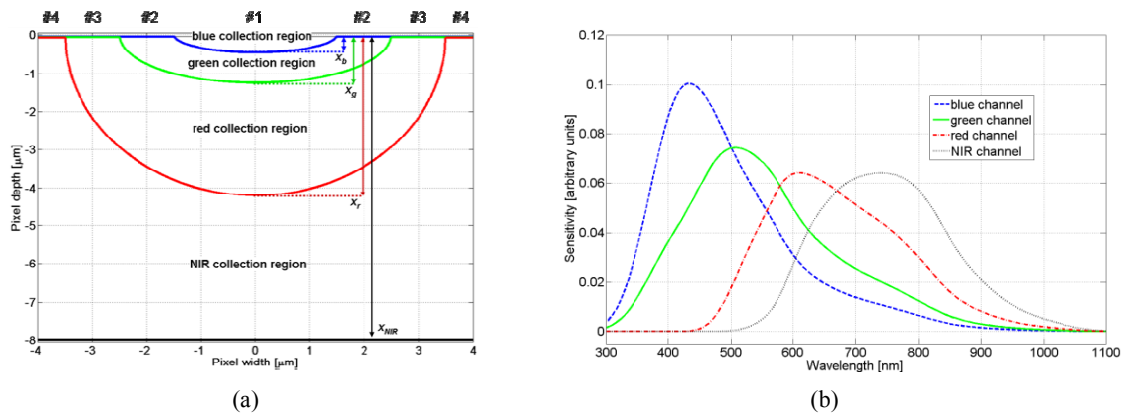

(a)   (b)

Figure 2. Example of TFD simplified geometry for *Matlab*® simulations (a). The result of each simulation is represented by four response spectra with peaks in the blue, green, red and near infrared bands (b).

On a set of simplified TFD geometries (an example is shown in Fig. 2a) we have performed radiation absorption simulations using *Matlab*®. These simplified simulations allow a fast analysis prior to detailed electron device simulations. The routine assumes that all the carriers generated in a certain sub-region of the device are collected by the corresponding contact, apart from carriers generated in a thin surface layer (few tens of nm), which are thought to be subject to recombination. The simulation sweeps the four parameters x_b, x_g, x_r and x_{nir} that represent the maximum collection depth of the blue, green and red contacts and the overall active layer thickness respectively. Parameters are swept over realistic values inferred by previous TFD implementations and experimental measurements. It is also assumed that a micro-lens can focus the radiation in the inner ($\pm$ 1 μm) device region, where color separation as a function of depth best happens. For every quartet of x_i the result of the simulation is a set of four spectral responses corresponding to the collected photo-signals from the different regions. An example is shown in Fig. 2b, with illustrative values of x_i. In the next section it will be examined how the signal processing algorithm works and which criteria are used to determine the optimum result for the joint capture of visible and NIR radiation.

4. VISUAL AND NIR INFORMATION PRE-PROCESSING

Each 4-channel TFD pixel outputs four values corresponding to the integral of the incoming radiation spectral density $s(\lambda)$ multiplied by the corresponding channel sensitivity:

$$R = \int s(\lambda) \cdot r(\lambda) d\lambda, \quad G = \int s(\lambda) \cdot g(\lambda) d\lambda, \quad B = \int s(\lambda) \cdot b(\lambda) d\lambda, \quad NIR = \int s(\lambda) \cdot nir(\lambda) d\lambda \quad (1)$$

Quantitatively, the integration starts to be significant roughly at a wavelength of 300 nm, below which the Silicon sensitivity is almost null due to surface charge recombination. As no hot mirror is used, the integral takes radiation up to 1100 nm, corresponding to the Si bandgap. The extraction of the VIS and the NIR information from the four signals collected by the detector requires:

1) a mapping from the TFD four dimensional color space *[r(λ), g(λ), b(λ), nir(λ)]* to a three dimensional color space *[r'(λ), g'(λ), b'(λ)]*. This operation corresponds to the subtraction of the unwanted infrared-generated signals inside the visible channels and it is obtained through a 3x4 matrix M_{NIR}. The derived spectral responses *R'*, *G'* and *B'* should at this point resemble *RGB* responses typically obtained when a hot mirror is used (i.e. cut-off around 700-750 nm);

2) the conversion from the three dimensional sensor color space *[r'(λ), g'(λ), b'(λ)]* into the standard CIE XYZ color space *[x'(λ), y'(λ), z'(λ)]* through a second matrix M_{CCM}. As usual, also in this work this matrix is obtained considering the matching of the values obtained from the 24 patches of the Macbeth Color Checker (MCC).

The two operations above can be clearly combined in a single matrix but it is preferred here to separately put in evidence the effects of the two steps:

$$
\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} =
\begin{bmatrix} c_{1,1} & c_{1,2} & c_{1,3} \\ c_{2,1} & c_{2,2} & c_{2,3} \\ c_{3,1} & c_{3,2} & c_{3,3} \end{bmatrix} \cdot
\left\{
\begin{bmatrix} d_{1,1} & d_{1,2} & d_{1,3} & d_{1,4} \\ d_{2,1} & d_{2,2} & d_{2,3} & d_{2,4} \\ d_{3,1} & d_{3,2} & d_{3,3} & d_{3,4} \end{bmatrix} \cdot
\begin{bmatrix} B \\ G \\ R \\ NIR \end{bmatrix}
\right\} =
\begin{bmatrix} c_{1,1} & c_{1,2} & c_{1,3} \\ c_{2,1} & c_{2,2} & c_{2,3} \\ c_{3,1} & c_{3,2} & c_{3,3} \end{bmatrix} \cdot
\begin{bmatrix} B' \\ G' \\ R' \end{bmatrix} \quad (2)
$$

The combined use of these conversions should determine:

a) a low propagation of noise, which corresponds to the presence of small negative coefficients in the two matrixes M_{NIR} and M_{CCM} (though signal is subtracted, noise is always added, thus lowering the SNR);

(b) a mean color reconstruction error (generally expressed in the *CIELAB* color space) that is acceptable for the target application (e.g. $\Delta E_{a*b*} \sim 3$ in photography).

Both these requirements, together with the requirement of a sufficient sensitivity in the NIR channel at the wavelength corresponding to the peak of the structured light, depend on the maximum collection depths of the different channels. The device optimization corresponds to an optimum choice of the four different depths x_i

4.1 Subtraction of NIR-generated signals from the visible channels

Consider the NIR portion of the spectrum (i.e. the wavelength range $\lambda > 780$ nm) in Fig. 2b. Due to the fact that at these wavelengths the inverse of the absorption coefficient $1/\alpha$ (that is the penetration depth) is much larger than x_i, one can express the exponential absorption law through its linear approximation. The efficiencies of the different channels thus become:

$$\eta_B^* \approx \alpha(\lambda) \cdot x_B, \quad \eta_G^* \approx \alpha(\lambda) \cdot x_G - \alpha(\lambda) \cdot x_B, \quad \eta_R^* \approx \alpha(\lambda) \cdot x_R - \alpha(\lambda) \cdot x_G, \quad \eta_{NIR}^* \approx \alpha(\lambda) \cdot x_{NIR} - \alpha(\lambda) \cdot x_R \quad (3)$$

These expressions, marked by an asterisk, are valid only in the near infrared range. Given the different depths x_i, it turns out that, in this range, the quantum efficiencies are linearly proportional, as evident from the right portion of the response spectra in Fig. 2b. On the contrary in the short wavelength range ($\lambda < 500$ nm), the NIR response is almost null. It thus turns out that infrared can be cancelled from visible channels simply by means of a linear subtraction. As a result, the matrix M_{NIR} has the quasi diagonal elements $d_{i,i}$ which are almost unitary, the elements of the fourth column $d_{i,4}$ which are negative (representing the subtraction of NIR) and the remaining elements that are almost null. In particular, the three negative coefficients are always very small as (from Eq. 3) they are strictly related to the sampling depth x_i as follows:

$$d_{1,4} \approx \frac{\eta_B^*}{\eta_{NIR}^*} = \frac{x_B}{x_{NIR}-x_R} \quad , \qquad d_{2,4} \approx \frac{\eta_G^*}{\eta_{NIR}^*} = \frac{x_G-x_B}{x_{NIR}-x_R} \quad , \qquad d_{3,4} \approx \frac{\eta_R^*}{\eta_{NIR}^*} = \frac{x_R-x_G}{x_{NIR}-x_R} \tag{4}$$

As the coefficients are small, the application of the conversion matrix M_{NIR} does not cause a significant propagation of noise across the colored image. The obtained R', G' and B' values represent the color coordinates in the sensor color space after the subtraction of unwanted infrared radiation.

4.2 Color conversion

In order to be interpreted by an external peripheral, the color coordinates must be converted into a standard color space, as for instance the CIEXYZ. This operation is done by considering a set of reference colors, the 24 patches of the MCC illuminated by the D65 standard source. The X, Y and Z triplets of these patches form a 3 x 24 matrix A.

Correspondingly, another 3 x 24 matrix B can be formed by considering the R', G' and B' values obtained by integrating the product of the TFD responses $r'(\lambda)$, $g'(\lambda)$ and $b'(\lambda)$ obtained above, again by the power spectral density of the 24 patches illuminated by the D65. The color conversion matrix can then be found as the product of the matrix A by the pseudo-inverse of the matrix B:

$$M_{CCM} = A \cdot [B^T \cdot (B \cdot B^T)^{-1}] \tag{5}$$

The modulus of the maximum negative coefficient of the resulting matrix depends on the shape of the spectral responses and on their relative amplitude and overlap. Also the goodness of the set of responses, expressed by the residual color conversion error[11] depends on these shapes. As a consequence an optimization of the spectral responses shape that takes into account these issues, together with the minimization of the crosstalk as described in section 4.1, must be done.

4.3 Optimum spectral responses

Considering realistic maximum absorption depths for the different channels in the TFD[8,9], the parameters were swept with $x_B \in [0.5\ \mu m - 1\ \mu m]$, $x_G \in [1.2\ \mu m - 1.8\ \mu m]$, $x_R \in [1.8\ \mu m - 3.6\ \mu m]$ and the maximum thickness $x_{NIR} \in [3\ \mu m - 8\ \mu m]$. For every simulation, the algorithm considers the minimum negative coefficient p_1 of the infrared subtraction matrix M_{NIR}, the minimum negative coefficient p_2 of the color conversion matrix M_{CCM} and a parameter p_3 representative of the color reconstruction error. The quality is then evaluated by minimizing the weighted sum:

$$q = \sqrt{(\alpha \cdot p_1)^2 + (\beta \cdot p_2)^2 + (\gamma \cdot p_3)^2} \tag{6}$$

In Eq. 6 the weighting coefficients α, β and γ are used to give more or less relative importance to the constraints discussed above. Giving more weight to α and β leads to a reduction of noise propagation; giving more weight to γ leads to a crosstalk reduction and to an improvement in the color reconstruction error of the *noiseless* sensor.

In case $\alpha = \beta = \gamma$ (that corresponds to the case of the best compromise) the values of the optimized depths turn out to be $x_B = 0.5\ \mu m$, $x_G = 1.4\ \mu m$, $x_R = 2.2\ \mu m$ and the active layer thickness is $x_{NIR} = 4.5\ \mu m$. The obtained matrixes are:

$$M_{NIR} = \begin{bmatrix} 1.015 & -0.022 & 0.298 & -0.231 \\ 0.079 & 0.885 & 0.3832 & -0.970 \\ 0.052 & -0.074 & 1.800 & -0.500 \end{bmatrix} , \quad M_{CCM} = \begin{bmatrix} 8.495 & -2.210 & 2.352 \\ 0.582 & 1.284 & -1.332 \\ -0.006 & -0.715 & 2.126 \end{bmatrix}$$

It can be observed that maximum negative coefficients are -0.97 and -2.21 respectively. This means that noise propagation is not significantly worsened by the crosstalk correction algorithm.

Fig. 3a shows the obtained result in terms of photoresponses, plotting the original response spectra $r(\lambda)$, $g(\lambda)$, $b(\lambda)$, $nir(\lambda)$ together with the reconstructed response spectra $r'(\lambda)$, $g'(\lambda)$, $b'(\lambda)$ on the whole wavelength range of interest in the optimal case. This case considers a slightly different geometrical width with respect to Fig. 2. Thanks to the subtraction algorithm, the crosstalk of the infrared signals inside the visible channels is reduced from 4.5%, 12.5% and 31.1% to 0.7 %, 1.2 % and 1.9% for the blue, green and red channels respectively. The corresponding geometry is reported in Fig. 3b, and a detail of the reconstructed responses in the visible range only is represented in Fig. 3c. Putting the latter photoresponses in the *ISET*® software [11], the mean color reconstruction error in the L*a*b* space, for the 24 patches

of the Macbeth Color Checker illuminated by a D65 standard source, turns out to be $\Delta E_{a*b*} \sim 3.8$, an acceptable value for fidelity in photographic color reproduction.
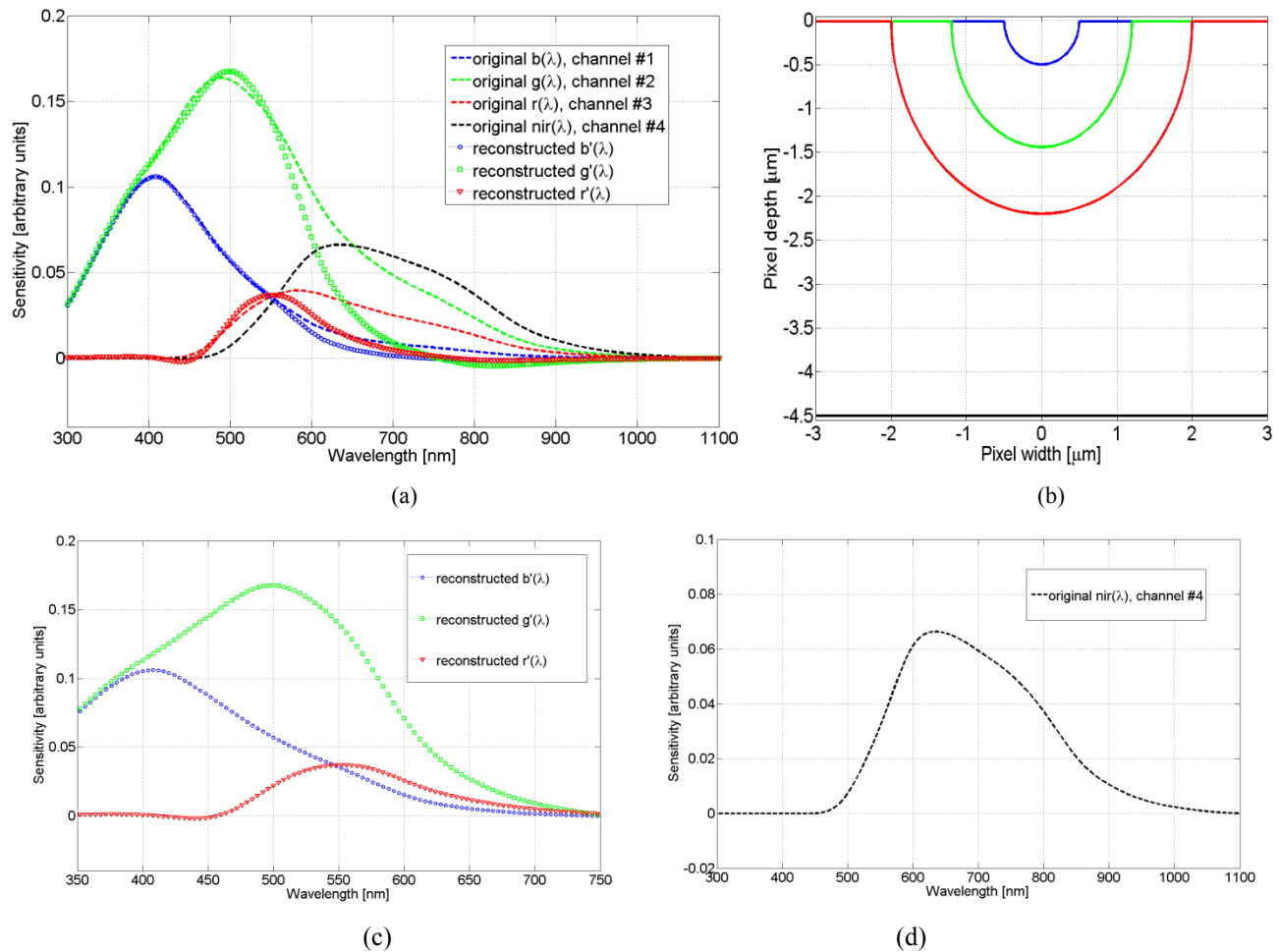

(a)

(b)

(c)

(d)

Figure 3. Optimum set of spectral responses obtained through *Matlab* simulations, sweeping over the different maximum collection depths x_i (a). Corresponding optimum geometry with $x_B = 0.5$ μm, $x_G = 1.4$ μm, $x_R = 2.2$ μm, and $x_{nir} = 4.5$ μm (b). Close up view of the responses $r'(\lambda), g'(\lambda), b'(\lambda)$, reported normalized and in the visible range only (c). Sensitivity of the infrared channel only: at the emission wavelength $\lambda = 850$ nm of the structured light source the sensitivity is still larger than the 30% of the peak value (d).

5. CMOS IMPLEMENTATION OF THE DEVICE

A CMOS implementation of the device is simulated using the *Dessis* software, considering the drift-diffusion model for the FEM solution of the electron device. In this simulation the design rules of a CMOS 130 nm are respected, and the doping profiles inside the semiconductor are realistically designed. In particular, the structure is formed using 7 *N-type* regions (two lateral *N* wells and five surface N^+ implants) with interjected *P-type* surface implants and Shallow Trench Isolation (STI) to provide better isolation from punch-through leakage current[25]. The overall pixel width is 6 μm but only half pixel (3 μm) is simulated, exploiting the symmetry of the structure.

The device doping configuration and the iso-potential lines are depicted in Fig. 4a and 4b. The contacts are biased at the following voltages: $V_1 = 0.5$ V, $V_2 = 2.2$ V, $V_3 = 2.7$ V and $V_4 = 4.0$ V. A detailed view of the electric field streamlines, representing carrier collection trajectories in the depleted region, is reported in Fig. 4c. It can be seen that the contacts collect carriers up to different depths. Contact n. 4 collects also from a larger depth by means of carrier diffusion.
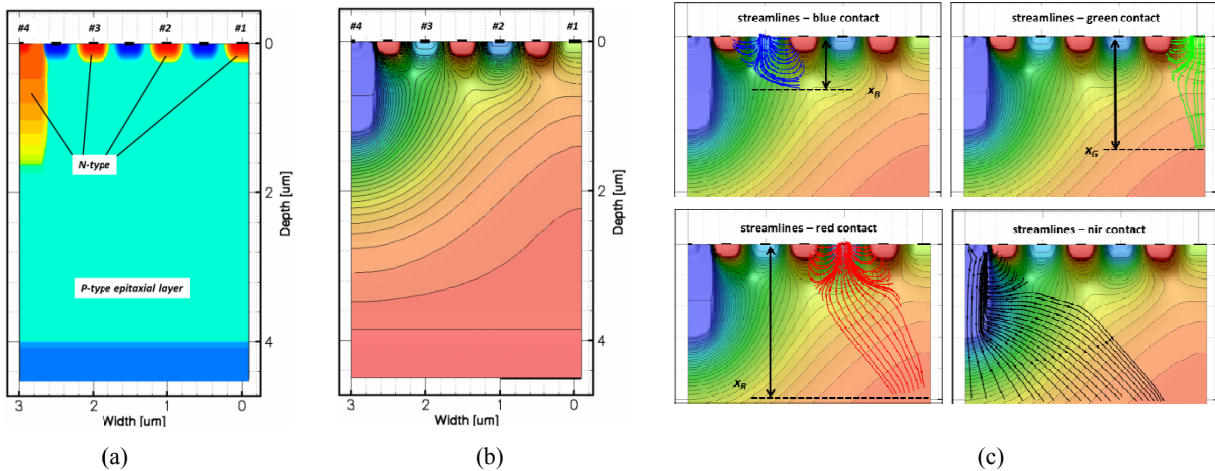
Figure 4. Doping configuration of the TFD using the electron device simulator *Dessis*. Only half pixel is simulated (a). Configuration for the iso-potential lines, at the biasing $V_1 = 0.5$ V, $V_2 = 2.2$ V, $V_3 = 2.7$ V and $V_4 = 4.0$ V (b). Close up view of the electric field streamlines corresponding to the carriers collection trajectories for the different contacts (c).

After solving the problem for the device without illumination, we then simulated a photogeneration impinging on the device at different wavelengths in the range of interest (300 nm – 1100 nm). As a result the spectral responses shown in Fig. 5 are obtained, that match with a good approximation the optimum result predicted by the simplified simulation shown above in Fig. 3.
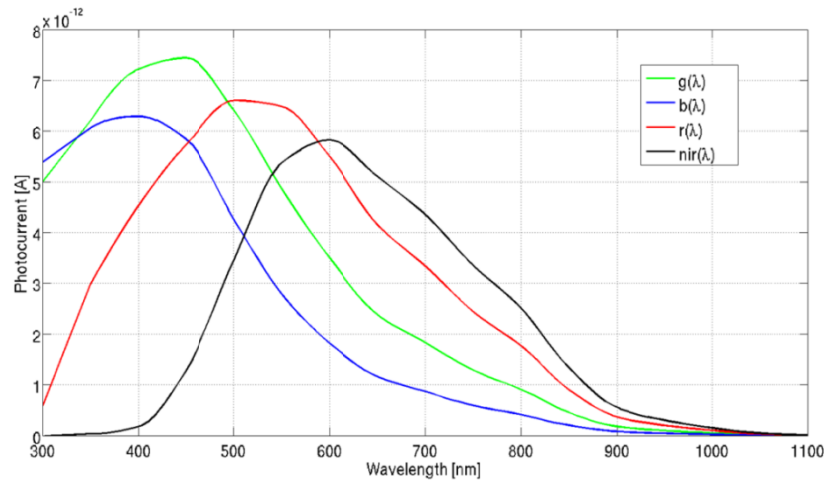

Figure 5. Spectral responses of the four contacts obtained using the *Dessis* simulator. Charge recombination at the interface between Silicon and Silicon Oxide was not included in the simulations, which explains the relatively large value of the photoresponses even at short wavelengths.

6. CONCLUSION

In this paper we have proposed the use of the Transverse Field Detector as a filter-less device for the joint capture of visible and near infrared radiation. The use of a monolithic detector makes easier the combination of the visible and infrared images in applications like 3D imaging, foreground/background segmentation and videoconferencing. With respect to other proposed device the TFD has advantages as it requires no demosaicking and the color and infrared channels has the same resolution. Through a suitable algorithm we have shown that potentially the TFD can consistently reduce the unavoidable crosstalk of infrared radiation inside the visible channels, while holding acceptable color reconstruction performance and sensitivity in the infrared channel. A feasibility study of a realistic TFD, built in a 130 nm CMOS technology has been done and the device is under design for realization and test.

REFERENCES

[1] X.P. Zhang et al., "Enhancing Photographs with Near Infrared Images", IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Anchorage, US (2008).

[2] L. Schaul et al., "Color Image De-hazing using the Near-Infrared", Proc. IEEE International Conference on Image Processing (2009).

[3] G. J. Iddan et al., "3D Imaging in the studio (and elsewhere…)", Proc. SPIE Vol. 4298(1), 48-55 (2001).

[4] M. Op de Beeck et al., "Towards an optimized 3D broadcast chain", Proc. SPIE Vol. 4864 (2002).

[5] Y. M. Lu et al., Designing color filter arrays for the joint capture of visible and near-infrared images, Proc. IEEE International Conference on Image Processing (2009).

[6] J. H. Lyu at el., "IR/Color Composite Image Sensor with VIPS (Vertically Integrated Photodiode Structure", 2007 International Image Sensor Workshop, June 7-10, 2007, Ogunquit, Maine USA.

[7] D. L. Gilblom et al., "Infrared and ultraviolet imaging with a CMOS sensor having layered photodiodes", Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications V, Proc. SPIE, Vol. 5301, (2004).

[8] A. Longoni et al., "The transverse field detector: a novel color sensitive CMOS device", IEEE Electron Device Lett., vol. 29, no. 12, pp. 1306–1308, Dec. 2008.

[9] G. Langfelder et al., "Further developments on a novel color sensitive CMOS detector", Proc. SPIE, Vol. 7356, 73562A (2009); doi:10.1117/12.822291.

[10] R. M. Turner et al., "Stacked pixels – An advanced 3-d technology gets more out of an image sensor", Photonics Spectra, November 2007

[11] J. E. Farrell et al., "A simulation tool for evaluating digital camera image quality", in Proc. SPIE IS, T Electron. Imaging Image Quality Syst. Perform., Y. Miyake and D. R. Rasmussen, Eds., 2004, vol. 5294, pp. 124–131.

[12] J. Salvi et al., "Pattern Codification Strategies in Structured Light Systems", in Pattern Recognition, Volume 37, Issue 4, pp. 827-849, April 2004.

[13] Lee, S.C. et al., "Signal Separation Coding for Robust Depth Imaging Based on Structured Light", in IEEE International Conference on Robotics and Automation, Vol. 4, pp. 4430-4436, 2005.

[14] Guan, C. et al., "Composite Structured Light Pattern for Three-Dimensional Video", in Optics Express, Vol. 11, Issue 5, pp. 406-417, 2003.

[15] Yang, R. et al., "Robust and Accurate Surface Measurement Using Structured Light", in IEEE Transactions on Instrumentation and Measurement, Vol. 57, No. 6, pp 1275-1280, June 2008.

[16] Wu, Q. et al., "Bi-Layer Video Segmentation with Foreground and Background Infrared Illumination", in MM '08, Proceedings of the 16th ACM International Conference on Multimedia. pp. 1025-1026, 2008.

[17] Wu, Q., "Robust Real-Time Bi-Layer Video Segmentation Using Infrared Video", Master Thesis, Dept. of Computing Science, University of Alberta, Fall 2008.

[18] Gunawardane, P. et al, "Invisible Light: Using Infrared For Videoconference Relighting", International Conference on Image Processing (ICIP 2010), Sept. 26-29, 2010.

[19] Woodham, R., "Photometric Method for Determining Surface Orientation from Multiple Images", Optical Engineering, Vol. 19, No. 1, pp. 139-144, 1980.

[20] Gokturk, S.B. et al, "A Time-Of-Flight Depth Sensor – System Description, Issues and Solutions", Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop, Vol. 3, p. 35, 2004.

[21] http://www.pmdtec.com/ Accessed 7/8/2010

[22] Crabb, R., "Real-time Foreground Segmentation via Range and Color Imaging", in CVPR Workshop on Time of Flight Camera Based Computer Vision, 2008.

[23] S. Koyama, Y. Inaba, M. Kasano and T. Murata, "A day and night vision MOS imager with robust photonic-crystal-based RGB-and-IR", IEEE Transactions on Electron Devices, 55(3), 754-759 (2008).

[24] S. Kawada, S. Sakai, N. Akahane, R. Kuroda and S. Sugawa, "A wide dynamic-range checkered-color CMOS image sensor with IR-cut RGB and Visible-to-Near-IR pixels", IEEE Sensors Conference, Christchurch, New Zealand (2009).

[25] G. Langfelder, "Isolation of Highly Doped Implants on Low-Doped Active Layers for CMOS Radiation Drift Detectors," IEEE Transactions On Electron Devices, vol. 56(8), 1767-1773 (2009).