A $0.18\mu m$ CMOS On-Chip Skin Detection Scheme based on NPNP-Triple-Junction Structure

Xiaojin Zhao[1], Shoushun Chen[1], Farid Boussaid[2], Amine Bermak[1]
[1]Department of ECE, the Hong Kong University of Science & Technology, Kowloon, Hong Kong
[2]School of EECE, the University of Western Australia, WA 6009, Australia
Email: eexjzhao@ust.hk, dazui@ust.hk, boussaid@ee.uwa.edu.au, eebermak@ust.hk

***Abstract*— In this paper, a $0.18\mu m$ CMOS color processing scheme based on NPNP-triple-junction structure is proposed to achieve real-time skin detection. The need for building complex demosaicing circuitry is completely eliminated, which dramatically reduces the fabrication cost and simplifies the pixel space sampling algorithm. In addition, real-time detection is enabled through the use of logarithmic readout mode and Gilbert normalization cell. The fabricated chip and related measurement results have verified the concept of a compact, low cost, CMOS fully compatible micro-sensor suitable for skin detection.**

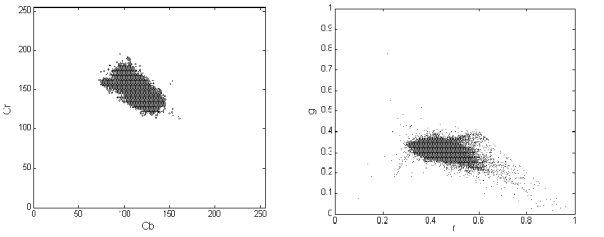
I. Introduction

The past few years have witnessed an increased research activity in the area of face detection especially due to the rapid emergence of security applications. Skin can play a major role in face detection algorithms. The purpose of skin detection is to locate the skin objects like faces or hands through a quick and coarse search of the scene. Most approaches to skin detection are based on color information, which is an often used feature in the detection with inexpensive computation; therefore, it is well suited for real-time VLSI implementations, which would make the integration of image-recognition capabilities on a single chip possible. Recent advances in the CMOS IC industry have already provided this unique opportunity to integrate sensing and processing functions together on a single chip. A fully integrated vision micro-sensor can thus feature built-in processing circuitry to emulate early human vision functions such as object detection or recognition [1][2], for which the way has been paved by several conventional contributions, including the on-chip color segmentation architecture and the Hue-Saturation-Intensity (HSI)-based pixel segmentation architecture proposed by Perez *et al* [3] and Etienne-Cummings *et al* [4], respectively. In both approaches, HSI color space was preferred to RGB (Red-Green-Blue) color space due to its effectiveness for segmenting scenes under various illumination environments. However, the main drawback of the proposed approaches, is related to a significant increase in silicon area and algorithmic complexity. Take the second approach for example, the proposed imager achieves skin or pattern recognition through template matching in the HSI color space, but it requires on-chip storage elements for the templates, resulting in a non-negligible increase in the imager core area. In addition, real-time on-chip skin detection is not achieved as the RGB components are sampled-and-held using a color filter wheel. In this paper, an RGB-space skin detection scheme based on NPNP-triple-junction structure is proposed and characterized through electrical measurements on a test chip fabricated using CMOS $0.18\mu m$ technology. The proposed circuitry enables the concept of a compact, low cost, CMOS fully compatible micro-sensor with integrated real-time skin detection function. The paper is organized as follows. In Section II, a detailed description of the proposed skin detection algorithm is provided. In Section III, the VLSI implementation of the proposed algorithm is presented. Section IV describes the chip characteristics and experimental results supporting the proposed architecture. Finally, a conclusion is presented in Section V.

II. Skin Detection Algorithm

Human skin detection can be realized by locating the skin color pixels from an input scene and a number of algorithms have been proposed to achieve this. In this paper, we propose to modify the algorithm by Chai *et al* [5] to make it suitable for VLSI implementation. In Chai's technique, skin color is modeled in YCbCr color space and the conversion from gamma-corrected RGB data to YCbCr data is given by

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}$$



(A) (B)

Fig. 1. (A) Skin color map in CbCr space; (B) Skin color map in rg chrominance space.

For primary RGB inputs in [0, 255], chrominance components Cb and Cr range from 0 to 255. Fig. 1 (A) shows the skin color map in CbCr space, obtained from hand-segmented images containing people with various skin tones (dark, white, yellow and red). According to Chai's technique, pixels are classified into skin and non-skin by using four threshold

values, which form a rectangular region in the CbCr space. A pixel is classified to have skin tone if its components (Cb, Cr) fall within the ranges, i.e, $Cb_{min} \leq Cb \leq Cb_{max}$ and $Cr_{min} \leq Cr \leq Cr_{max}$. The threshold values proposed in [5] are

$$\begin{aligned} [Crmin, Crmax] &= [133, 173] \\ [Cbmin, Cbmax] &= [77, 127] \end{aligned} \quad (1)$$

To make this algorithm hardware friendly, two modifications are proposed. Firstly, the RGB components are normalized to make colors less dependent on the change in light intensities.

$$\begin{aligned} r &= \frac{R}{R+G+B} \\ g &= \frac{G}{R+G+B} \\ b &= \frac{B}{R+G+B} \end{aligned} \quad (2)$$

Because r+g+b=1, color information can be represented by any two of the three color components rgb. The skin color map in rg color space is shown in Fig. 1 (B). This map is obtained from the same pixels used to generate the skin color map in Fig. 1 (A). Secondly, the multivariate Gaussian probability density function is chosen to model skin color:

$$f(\vec{x}) = \frac{1}{2\pi\sqrt{|\sum|}} \exp\left[-\frac{1}{2}(\vec{x}-\vec{\mu})^T \sum^{-1} (\vec{x}-\vec{\mu})\right] \quad (3)$$

where $\vec{\mu}$ and $\sum$ are the mean vector and covariance matrix, respectively. A particular input pixel of chrominance vector $\vec{x}$ is classified as skin if the probability density function f($\vec{x}$)$\geq\tau$, where τ is a threshold. Taking the logarithm of Equation (3) leads to the following relation for skin detection:

$$(\vec{x}-\vec{\mu})^T \sum^{-1} (\vec{x}-\vec{\mu}) \leq F \quad (4)$$

where F is a new threshold related to τ. A given pixel of normalized components (r,g) will thus be classified as skin-pixel if it satisfies relation (5), that is:

$$\begin{aligned} &s_{11}(r-\mu_r)^2 + s_{22}(g-\mu_g)^2 \\ &+2s_{12}(r-\mu_r)(g-\mu_g) \leq F \end{aligned} \quad (5)$$

$$with \sum = \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}$$

To estimate the model parameters ($\vec{\mu}$, $\sum$), a training set containing over 10 million skin pixels was prepared by manually segmenting 200 images. The training set includes skin colors of people from different races and under different lighting conditions.

III. Sensor Architecture and Implementation

The proposed on-chip skin detection scheme consists of two basic blocks: RGB image capture block and image processing block. The sensor's VLSI architecture is depicted in Fig. 2 and Fig. 3: Fig. 2 shows the pixel circuitry using current readout structure and Fig. 3 shows the skin detection processing structure.

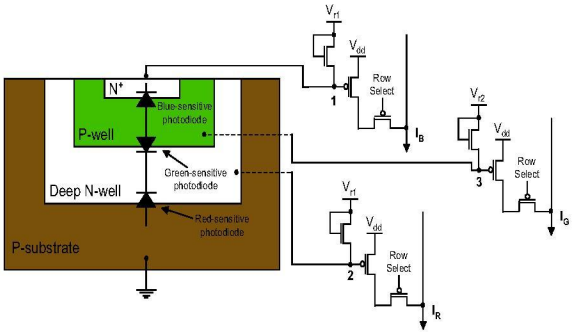


Fig. 2. Pixel circuitry.

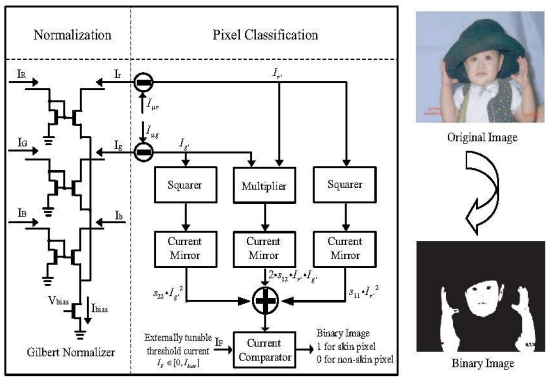


Fig. 3. Skin detection processing structure.

A triple-junction (NPNP) CMOS process is chosen to fabricate three vertically integrated photodiodes acting as the RGB color detector for each pixel. The device operation is based on the strong wavelength dependence of the photon penetration depth. For an incident monochromatic light penetrating in the silicon bulk with a given photon flux, the penetration depth is a function of the wavelength of the incident light. Therefore, the photons with different wavelengths can be captured at different penetration depths. As a result, all three RGB photocurrents can be simultaneously generated at different junction depths and sent to the inputs of the adjacent image processing block. Logarithmic readout circuitry is selected as the front stage of image processing block. All three RGB photocurrents I_R, I_G and I_B can be readout via 3 output buses with transistors operating in the subthreshold region. Real-time sensing can be realized since in the current-mode readout, the photodiode will be automatically restored to the initial state once there is no input light. Subsequently, the three readout currents are passed to the last processing stage, where they are normalized using a Gilbert cell with outputs described by

$$\begin{aligned} I_r &= \frac{I_R}{I_R+I_G+I_B} \cdot I_{bias} \\ I_g &= \frac{I_G}{I_R+I_G+I_B} \cdot I_{bias} \\ I_b &= \frac{I_B}{I_R+I_G+I_B} \cdot I_{bias} \end{aligned} \quad (6)$$

This keeps color components within a narrow range over which the processing circuitry can be set to operate correctly. Pixels of an input RGB image are then classified as skin or non-skin pixels by computing the multivariate Gaussian probability and evaluating the following expression:

$$s_{11}\left(I_r - I_{\mu r}\right)^2 + s_{22}\left(I_g - I_{\mu g}\right)^2 + 2s_{12}\left(I_r - I_{\mu r}\right)\left(I_g - I_{\mu g}\right) \leq I_F \quad (7)$$

The threshold current I_F along with parameters $I_{\mu r}$ and $I_{\mu g}$ are set externally as a fraction of the Gilbert normalizer bias current I_{bias}. Covariance matrix coefficients s_{ij} are implemented by means of a tunable active input current mirror. A four quadrant multiplier/divider is used to achieve multiplication and squaring [6]. A single current comparator is required here for skin pixel classification. This allows compensation for mismatch induced errors and for varying lighting conditions. The output is a binary image in which skin pixels are displayed in white and non-skin pixels in black.

With the CMOS triple-junction (NPNP) process, the proposed compact and low cost design is fully compatible with CMOS standard process and completely removes the need for complex on-chip demosaicing dedicated circuitry. In addition, by combining the benefits of focal plane processing and computation on readout, the proposed VLSI architecture enables real-time skin detection processing. Finally, the use of a statistical Gaussian skin color model offers an acceptable trade-off between skin detection performance and implementation complexity.

IV. Chip Design, Measurement and Results

In order to validate that the proposed compact, low cost, CMOS fully compatible scheme is suitable for skin detection, a test chip was designed and fabricated. The chip design, test setup and related experimental results are presented in the following subsections.

A. Chip Design and Test Setup

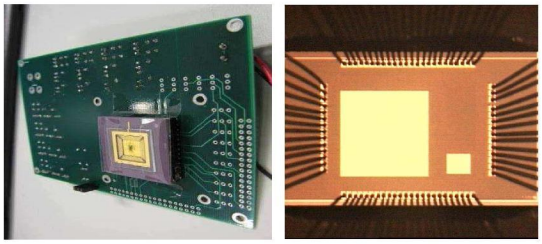
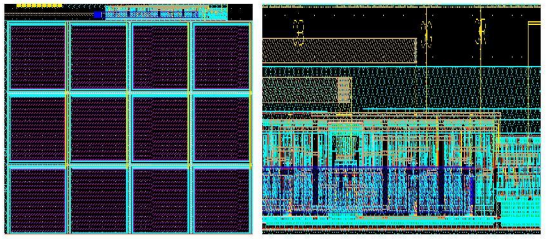


Fig. 4. Fabricated chip: packaged chip (left) and microphotograph (right, all details are obscured by the light shield layer).

The fabricated test chip comprises a set of test structures designed for the modular test and characterization of individual circuits as well as of the entire skin detection processing chain. The test chip and its layout views are shown in Fig. 4 and Fig. 5 respectively. The electrical measurements were carried out using the test setup depicted in Fig. 6, which includes the device under test (DUT), the HP4156A analyzer and the optical test platform consisting of a white light source,



(A) (B)

Fig. 5. Layout views: (A) Pixels; (B) Image processing blocks.

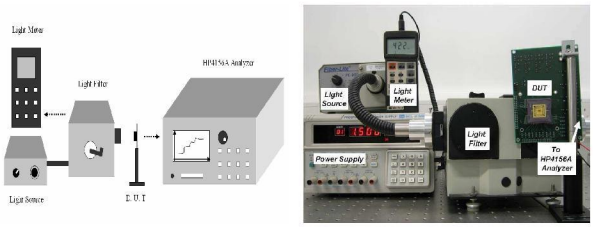
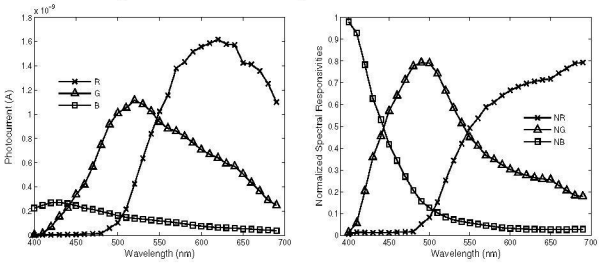


Fig. 6. Diagram (left) and photograph (right) of the test setup.

a monochromatic light filter, a light meter and a DC power supply.

B. Experimental Results

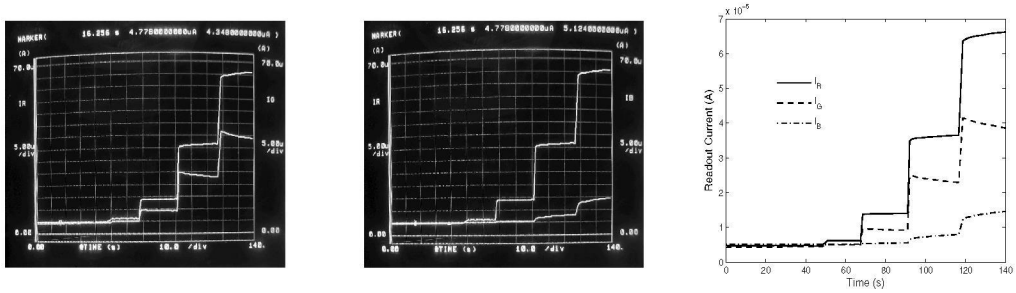
In this subsection, the experiments performed on the test chip (including RGB color detector, pixel readout circuitry, normalization and pixel classification circuits) are described and the acquired results are reported. The RGB color detector



(A) (B)

Fig. 7. (A) Spectral responsivities of the RGB photodetectors; (B) Normalized spectral responsivities of the RGB photodetectors.

was characterized using a monochromatic beam with the wavelength varied from 400nm to 700nm. Measured spectral responsivities and normalized spectral responsivities, shown in Fig. 7, validate the use of a triple-junction RGB photodetectors to discriminate between the three fundamental components of visible light. The logarithmic readout circuitry was measured by exposing the chip to an input white light. Fig. 8 shows the pixel output currents I_R, I_G and I_B with the incident white light intensity varied as a staircase function. The readout currents of the three channels versus the input light intensity are displayed in Fig. 9. In order to validate the processing chain including photodetectors, readout circuitry and Gilbert normalization circuit, the intensity of the blue color component was utilized as the input function, which was increased from 0 to 300lux with a 50lux step, and then subsequently decreased.



(A) (B) (C)

Fig. 8. Measured pixel output currents with a staircase input light intensity: (A) I_R and I_G (screen capture); (B) I_R and I_B (screen capture); (C) I_R, I_G and I_B replotted together because the experimental setup allows for the simultaneous display of at most 2 curves.

Note that normalized currents I_r and I_g vary as expected with the sum of all three normalized currents remaining constant. This is seen in Fig. 10 with the corresponding normalized

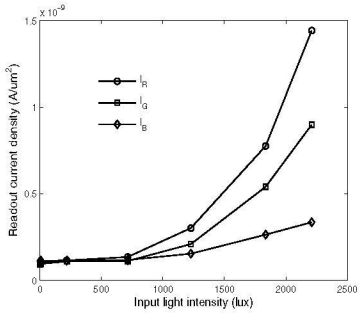


Fig. 9. Measured pixel output currents I_R, I_G and I_B versus input light intensity.

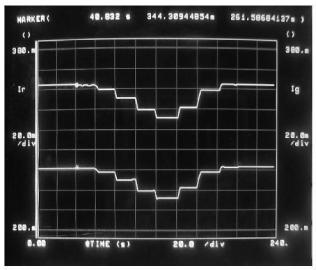


Fig. 10. Gilbert normalizer (screen capture): red I_r (top curve) and green I_g (bottom curve) outputs as a function of the blue color component. (At most 2 curves can be displayed simultaneously.)

output currents I_r and I_g as a function of the blue fundamental component. I_r and I_g are also plotted versus the blue color intensity in Fig. 11. Finally, the multiplier/divider and active input current mirror circuits used in the classification stage were also characterized. Implemented current mirrors exhibited less than 3% variation over a 5 decades photocurrent range. Implemented four quadrant multiplier/divider circuits can operate at a power supply of 1.8V, with a nonlinearity error of at most 2%. The prototype was fabricated using UMC 0.18μm 1-poly, 6-metal, 1.8V CMOS technology and the pixel fill-factor is 25%.

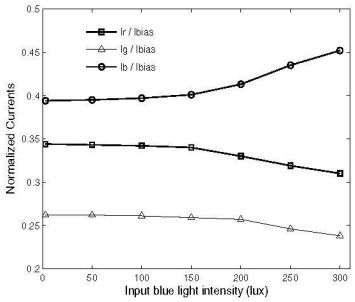


Fig. 11. Gilbert normalizer: red I_r, green I_g and blue I_b outputs as a function of the blue component intensity.

V. Conclusion

The prototype demonstrates a real-time CMOS on-chip skin detection scheme based on triple-junction structure, which is relatively simple, highly compact and enables the fabrication of an integrated skin detection micro-sensor in a standard CMOS process with significant cost savings. Reported experimental results validate the proposed system which can find applications in smart consumer imaging products.

Acknowledgment

This work was supported in part under Australian Research Council's Discovery Project DP0453371 and an Emerging High Impact Area (HIA05/06.EG03) from HKUST.

References

[1] A. El Gamal and H. Eltoukhy, "CMOS Image Sensors," IEEE Circuits and Devices Magazine, vol. 21, Issue 3, May-June 2005.

[2] E. Culurciello, R. Etienne-Cummings, K.A. Boahen, "A Biomorphic Digital Image Sensor," IEEE Journal of Solid-State Circuits, vol. 38, no. 2, pp. 281-294, 2003.

[3] F. Perez, C. Koch, "Towards Color Image Segmentation in Analog VLSI: Algorithms and Hardware", International Journal of Computer Vision, vol. 12, no. 1, pp. 17-42, 1994.

[4] R. Etienne-Cummings, P. Pouliquen and A. Lewis, "Color segmentation, histogramming and pattern matching chip," Proc. ISCAS, pp. 320-323, 2002.

[5] D. Chai and K. N. Ngan, "Face segmentation using skin color map in videophone applications", IEEE Transactions on Circuits and Systems for Video Technology, vol. 9, no. 4, pp. 551-564, 1999.

[6] F. Ismail and T. Fiez, Analog VLSI Signal and Information Processing, McGraw Hill, 1994.