# An Intelligent Reader Based on Nios II

Wei-Liang Xu[1], Duo-Hua Xu[2], and Wen-Yan Chen[3], and Zhi-Yong Pang*

*School of Physics and Engineering, Sun Yet-Sen University*

*No. 135, Xingang Xi Road, Guangzhou, 510275, P. R. China*

** Instructor*

[1]932394919@qq.com

[2]898862479@qq.com

[3]1356368761@qq.com

*Abstract*— **With the rapid development of Human-Computer interaction techniques, the hand is no longer the only input means of human physique structure. This paper describes an intelligent reader system, which based on head movements for the man-machine interface, realizes the operations like zooming, page-turning and row-scrolling on the reading interface, through the rough face detection and localization of facial feature points to achieve the tack and recognition of head movements**.

***Keywords*—face detection; facial complexion model; facial feature points positioning; head motion identification; FPGA; SOPC; Nios II**

## I. Introduction of System

With the development of electronic technology, electronic reading has increasingly become a way of daily reading. Computer, laptop, tablet and mobile phone correspondingly have the function of web browsing and electronic text reading. Under this background, a kind of comfortable, convenient way of electronic reading there is the need of development.

### A. Introduction of system function

Today most of the consumer electronics (laptop, mobile phone, tablet, etc.) come with a front-facing camera. Taking advantage of this, we plan to design a kind of non-contact e-reader system by monitoring the reader head movement to realize such intelligent operations as rolling line, flip the current reading text and etc. The system is suitable for the vast majority of electronic consumers and e-books reading lovers. Users in the read only by the head movement can control text reading progress, in addition to provide a comfortable way of reading, but also a solution to some special cases of reading such as the reading of hand invalid, browsing mobile phone when standing on subway, copying the information from the screen.

When a user is faced with the screen, the front camera will capture video to the processing module which can detect the user's facial features and recognize the user's head movements by a certain test algorithm to control the e-reader.

Specific functions as follows:

*1)* The left-leaning and right-leaning of operator's head respectively control zoom-in and zoom-out of the reading interface;

*2)* The left-turn and right-turn of operator's head control the current reading interface flip back and forth;

*3)* The operator's looking up and down respectively controls the current reading interface of rolling line up and down.

Steps:

*1)* Face segmentation: sample of the image and segment the face region and the image, mark rectangular surrounding the face.

*2)* Facial feature location: get the coordinates of lips and eyes in the mark face rectangular to build initialization feature triangle of eyes and mouth.

*3)* Head motion detection: record characteristic parameters of real-time triangle, and compare them with the initialization feature triangle to recognize head movements and get operating instructions.

### B. Introduction of Hardware Platform

This system is mainly realized on the Altera DE2 FPGA multimedia development platform, which is equipped with rich logic elements and large memory components, and provides a wide variety of product features of low power consumption and high integration. It is the development tool of choice for complex digital system for its superior performance in video, audio, image and other data processing. To satisfy the demand of large amount of image data processing and high real-time, the FPGA can accelerates the data processing by using hardware.

The system hardware parameters are as follows:

*1)* Development platform: the DE2-115

*2)* User action space range: about 40-70 cm in front of the display screen

*3)* System interface: resolution of 720 p, color depth of 24

*4)* Camera parameters: D5M of Youjingkeji, the input resolution of 640 * 480 @ 60 hz

*5)* CPU frequency: 100 MHZ

## II. Introduction of Algorithm

### *A. A Half-adaptive Facial Complexion Model*

Facial complexion is dominant in the pixel of human face. While by facial complexion is not full proof to identify a face, color judgment as the speed of face detection preprocessing has an obvious advantage, in line with the requirements of system for real-time compared with other methods of face detection. At the same time, around the head movement will seriously affect the accuracy of most face detection algorithm, such as the Adaboost algorithm unable to recognize the positive face and side face at the same time, which can create a great difficulty to realize the detection of head movements. And the use of color detection can effectively avoid that.

According to principle of color clustering, there are three kinds of facial complexion models most commonly used as HSV, Fussy HSCC and YCbCr. However, none of them recognize can meet the requirements of dynamic accuracy after simulation under openCV. This is due to the parameter values of the models are fixed got from experiments with high requirements for background or every time before use need to take a corresponding background images without a face.

For that, we have designed a half-adaptive method -- before the start of the system, the interface will ask the user to enter his or her front face by camera to initialize the system, establishing a histogram of face in YCbCr color space and obtaining user facial complexion of Cb, Cr value component as the template of face detection (as the initialized complexion model), thus greatly avoiding the influence of different backgrounds. At the same time, the YCrCb color space (Y: brightness, Cr: value of red component, Cb: value of blue component) as the facial complexion distribution statistical mapping space has separated the luminance and chrominance, realizing the luminance and chrominance segmentation. CrCb the two dimensional independent distribution, also better limits the complexion area, making the system can be used within a certain range to the light intensity change, more precise than ordinary color recognition. The flow chart is shown in figure 2.1 a, the simulation rendering as shown in figures below.

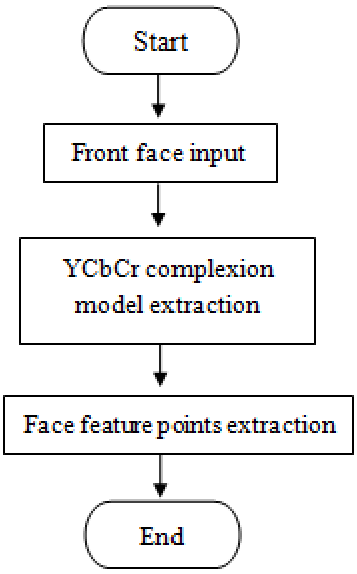


Fig.1 The initialization of face complexion model

Fig.2 Face of initialization

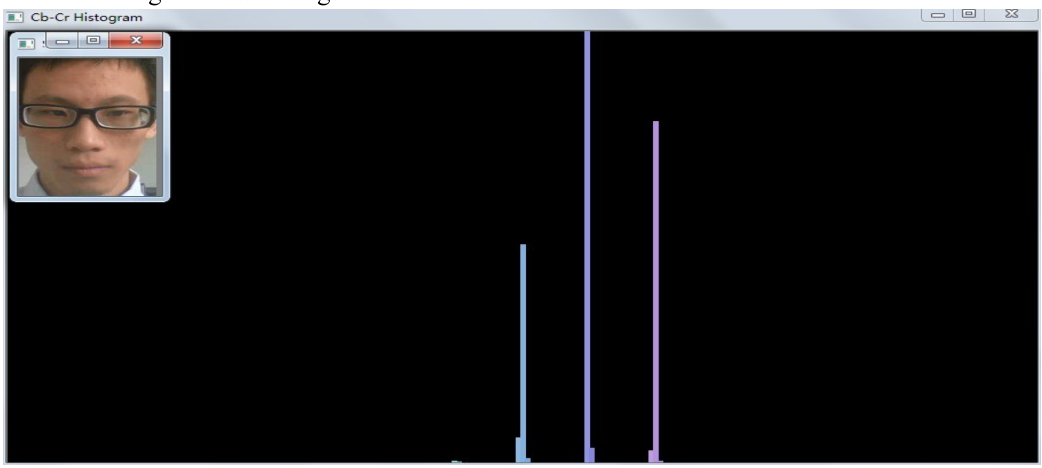


Fig.3 Histogram of initialization

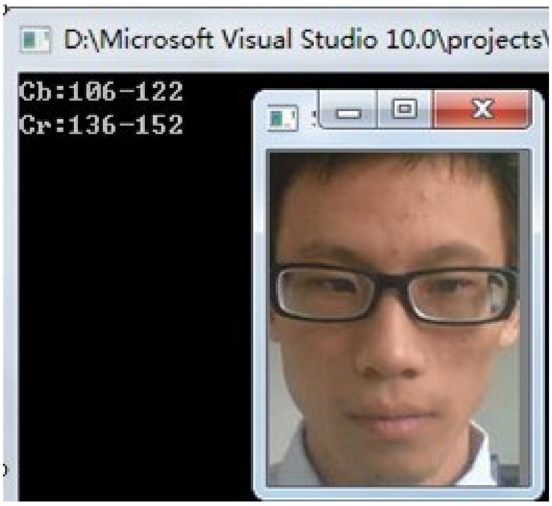


Fig.4 Initialized model of complexion

### B. *Bounding Box Of The Face*

Through the complexion segmentation designed according to the Initialized complexion model, we can separate the face region from the image—the pixels within the limits of the model are valued to be 1, and the other to be 0, finally creating a Binary image.

Set after the skin color segmentation binary image size for M * N, pixel values for each point f (x, y), so the minimum enclosing rectangle (bounding box) can be calculated as follows:

*1)* To calculate the horizontal projection (i.e., the sum of image pixels per line) Py [i], and the maximum of Py--MaxPy, MaxPy as the face width;

*2)* To calculate the vertical projection (i.e., the sum of each column of pixels image) Px [i];

*3)* To calculate the integral of Px in the (i, i + MaxPy - 1) interval noted as Sx [i], in which i is valued from 0 to M – MaxPy, and find out the coordinate of the maximum value point of Sx[i], as the bounding box border left, as well as left + MaxPy- 1 as the right (border) of the rectangle;

*4)* Considering the effect of noise points, MaxPy as face external rectangle width is often larger, therefore, we have to make a revision about boundaries: set a bound as a threshold, research from left to right the function Px (i), hit the first i meet the formula:

(Px (i) > bound) && (Px (i + 1) > bound) && (Px (i + 2) > bound) && (Px (i + 3) > bound)

Record left boundary value of the i ; as for right boundary, the i from right to left to investigate function Px (i), when first i value satisfies formula:

(Px (I) > bound) && (Px (I - 1) > bound) && (Px (I - 2) > bound) && (Px (I - 3) > bound)

Set right boundary value of the i ; Usually, the noise points are separated, not concentrated in a column, therefore such revision method can improve the accuracy of around the left and right boundary location;

*5)* To calculate horizontal projection Pylr [i] between left and right;

*6)* According to the proportion of facial morphology, often the face length-width ratio is less than 1.4:1, take Height = 1.4 * (right - left) as the external rectangular Height of coarse estimation;

*7)* To calculate the integral of Pylr in the (i, i + Height - 1) interval noted as function Sy [i], in which i is valued from 0 to N - Sy , and find out the coordinate of the maximum value point of Sy[i] as face external rectangular border on top, top + Height - 1 as bottom of the rectangular boundary.

The roughly estimated rectangular Height may be larger, the upper and lower boundaries are supposed for a revision as well.

At this point, we have got the bounding box of face. And the correct box is supposed to occupy the image area of more than 20%, and the proportion of complexion pixels in the box must be more than a certain value, which can be two standards to determine whether the region is face further (figure 5 ~ 8 is detected under different head pose).

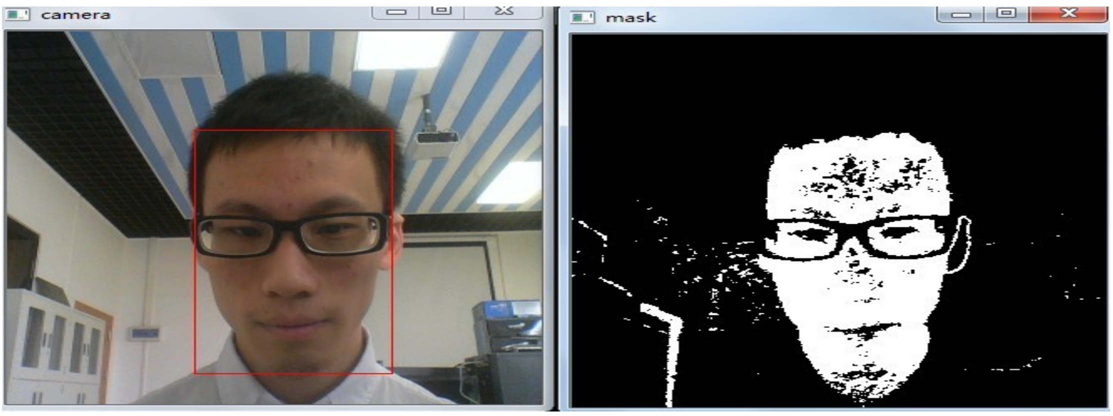


Fig.5 Front face

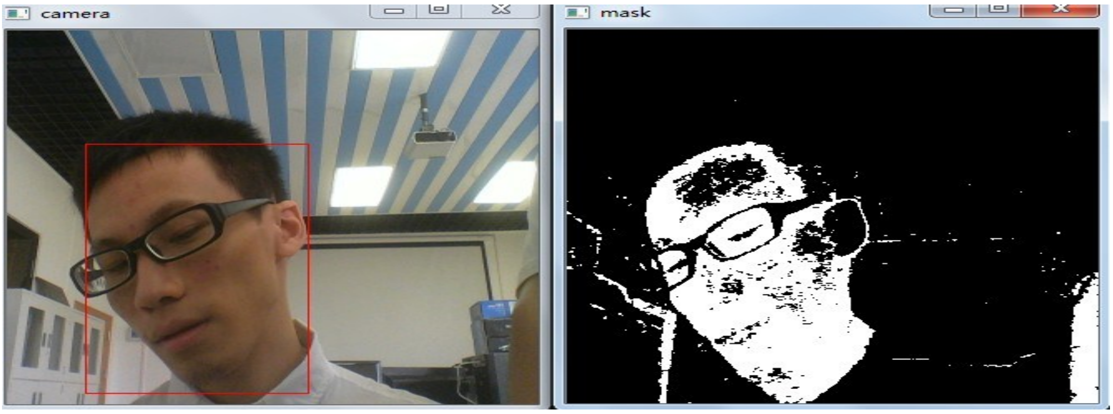


Fig.6 Left-leaning head

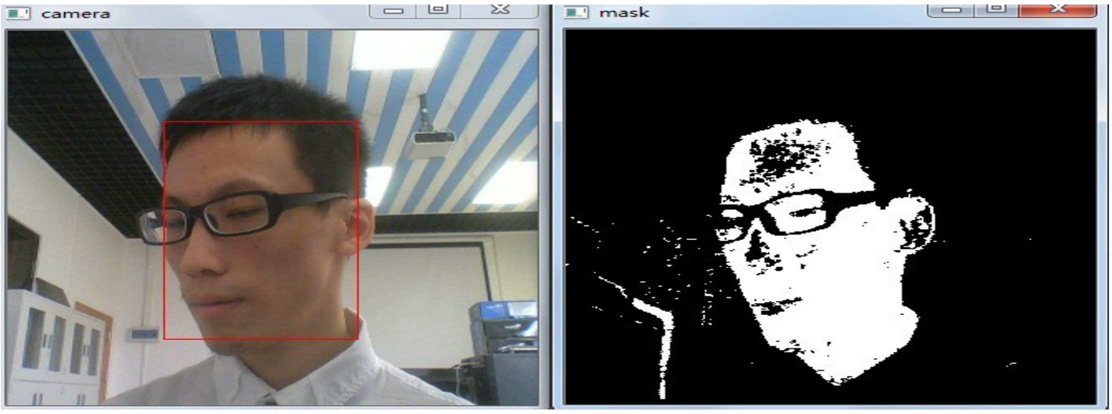


Fig.7 Left-turning head

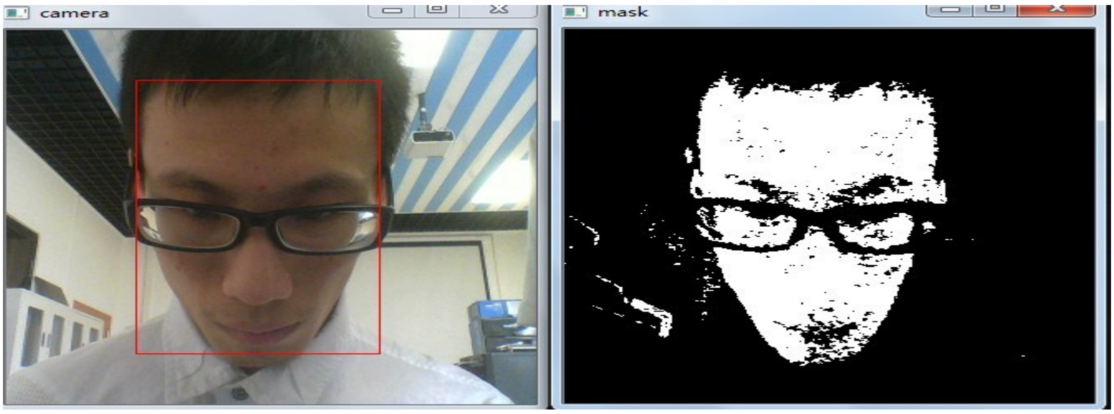


Fig.8 Looking-down head

### *C. Facial Feature Points*

Our system adopts the triangle of facial feature points (the eyes and lip) as for determining the head position.

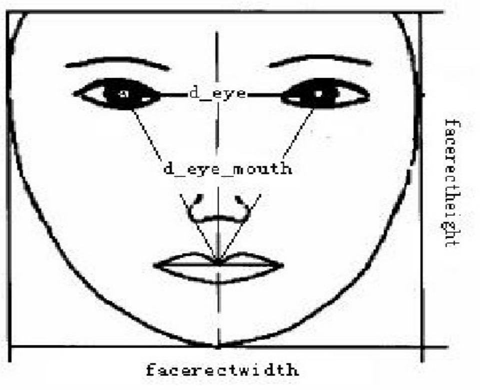


Fig.9 The triangle of face

First, we can figure out the regions of eyes and mouth by the classifier of Adaboost algorithm from our bounding box of face. And then calculate the coordinates of the centers of eyes and mouth by means of geometric gray projection.

Geometric gray projection is a classic in the field of facial feature location algorithm, which is based on that the facial features gray value is lower than those of the skin. It first find out the specific points of change by the grey value of horizontal or vertical direction projection, and then according to the prior knowledge, via the combination of the points in different direction, get the position of the facial features, with the advantage of low amount of calculation, fast solution.

### *D. Head motion identification*

The status and parameters of the characteristic triangle of face can be changed via user's head motion, including left-turn and right-turn in horizontal direction, look-up and look-down in vertical direction as well as left-leaning and right-leaning in a plane. Our system takes the characteristic triangle of initialized face as a benchmark, compared with the real-time monitoring triangle, to identify the head movements, and give the corresponding instructions.

The parameters we need is the same either initializing or real-time monitoring, including the coordinates of the eyes, the y-coordinate of mouth and the ratio distance of eyes to the distance between eyes and the mouth.

The process of identification:

*1)* When the real-time face around eyes y coordinates the absolute value of difference exceeds a certain threshold, it is judged that the head is left-leaning if the y-coordinate of right eye is greater, and the opposite decides to head right-leaning;

*2)* When the real-time face, the ratio distance of eyes to the distance between eyes and the mouth is less than the ratio of initialization and the difference above a certain threshold, it is judged that the head is left-turn if real-time eye center x coordinate is less than the initial model and right-turn conversely;

*3)* When the ratio of real-time face is greater than the ratio of initialization and the difference more than a certain threshold, it is judged that the head is look-down if real-time mouth center y coordinate is less than initial one and look-up conversely.

And we stipulate that (1) is prior to (2) (3), algorithm process as shown in figure 2.4. Among them, Left_y, Right_y respectively for real-time eye y coordinates and initialized ones; Act_rate, Model_rate respectively for real-time ratio and initialized one; Act_x, Model_x respectively for real-time eye center x-coordinate and initialized one; Act_y, Model_y respectively for real-time mouth y-coordinate and, initialized one;Th1, Th2 respectively for system setting threshold.

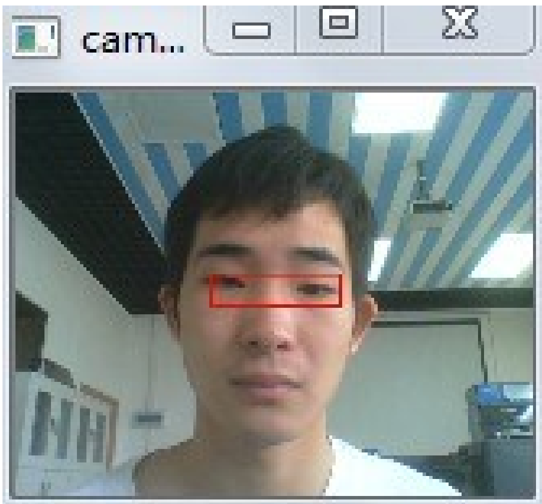


Fig.10 Detection of eyes

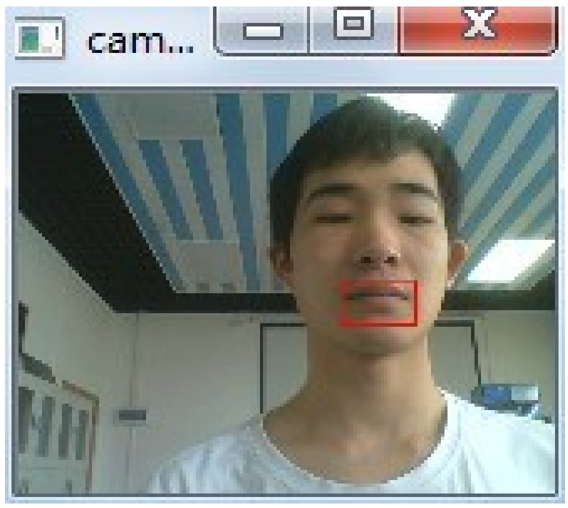


Fig.11 Detection of mouth

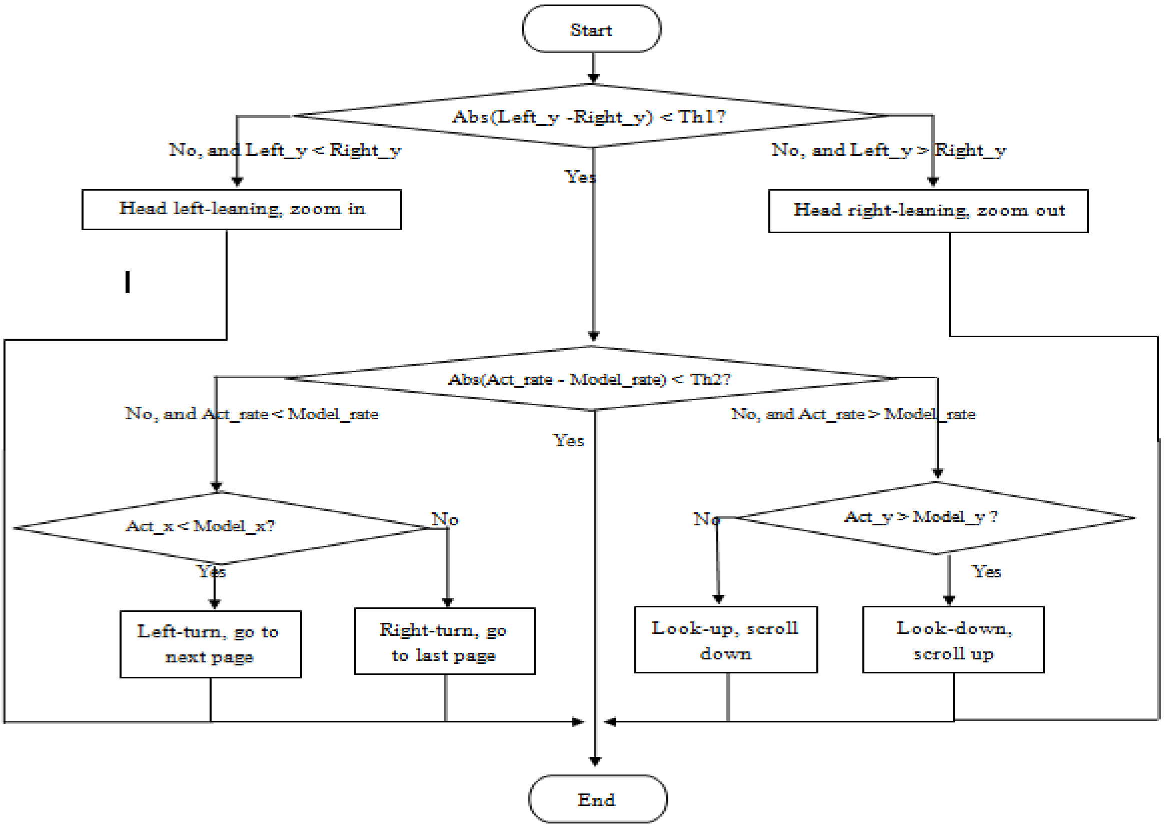


Fig.12 Head Motion Identification Algorithm Flow Chart

## III. SYSTEM STRUCTURE

### *A. Hardware*

*1)* System chart

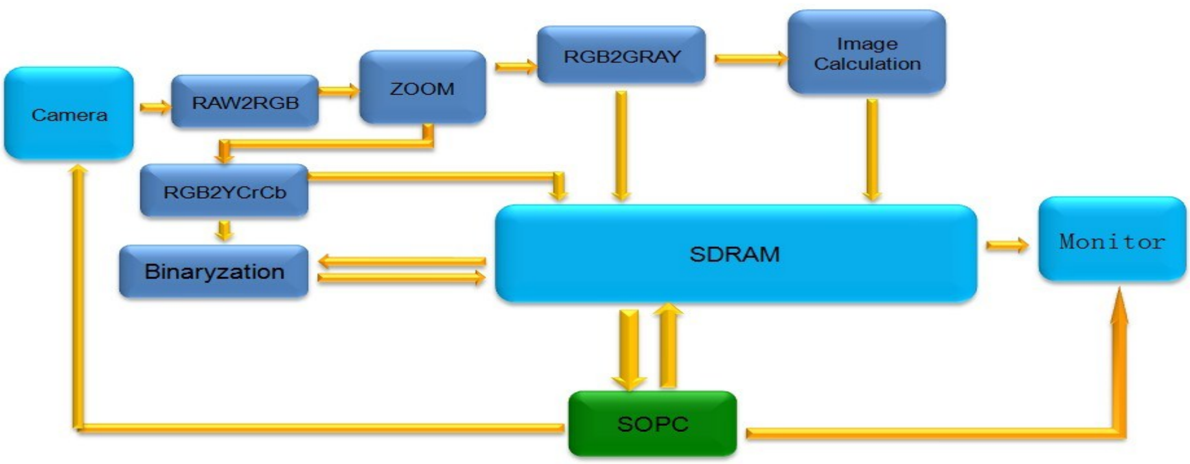


Fig.13

Dark blue modules

a. The dark blue modules are for FPGA hardware acceleration: RAW2RGB module will simulate the camera image data to interlaced and convert to RGB format; RGB2YCrCb module converts RGB image after scaling to YCrCb format and store in SDRAM; RGB2GRAY module converts the RGB image after scaling into grayscale, storing in SDRAM;

b. Zoom module resizes the images into a certain proportion of the scaling by using the image interpolation algorithm, to reduce subsequent image calculation;

c. Binarization module will get the complexion model initialized by SOPC from the SDRAM, and convert the subsequent input YCbCr image into binary image; Image pre-calculation module is a module that can do image horizontal and vertical integration and etc.

This system adopts analog camera, the capture of video refers to the DE2_70 routines, and convert the image data into RGB, YCbCr, Gray and use interpolation algorithm for image zooming, storing in SDRAM finally; At the same time, in order to ensure the system real-time performance, we make use of the FPGA hardware to accelerate ,for example realizing the calculation of integral image by hardware.

Light blue modules stand for the system external equipment including cameras, SDRAM and monitors, etc.

SOPC module shown as follows：

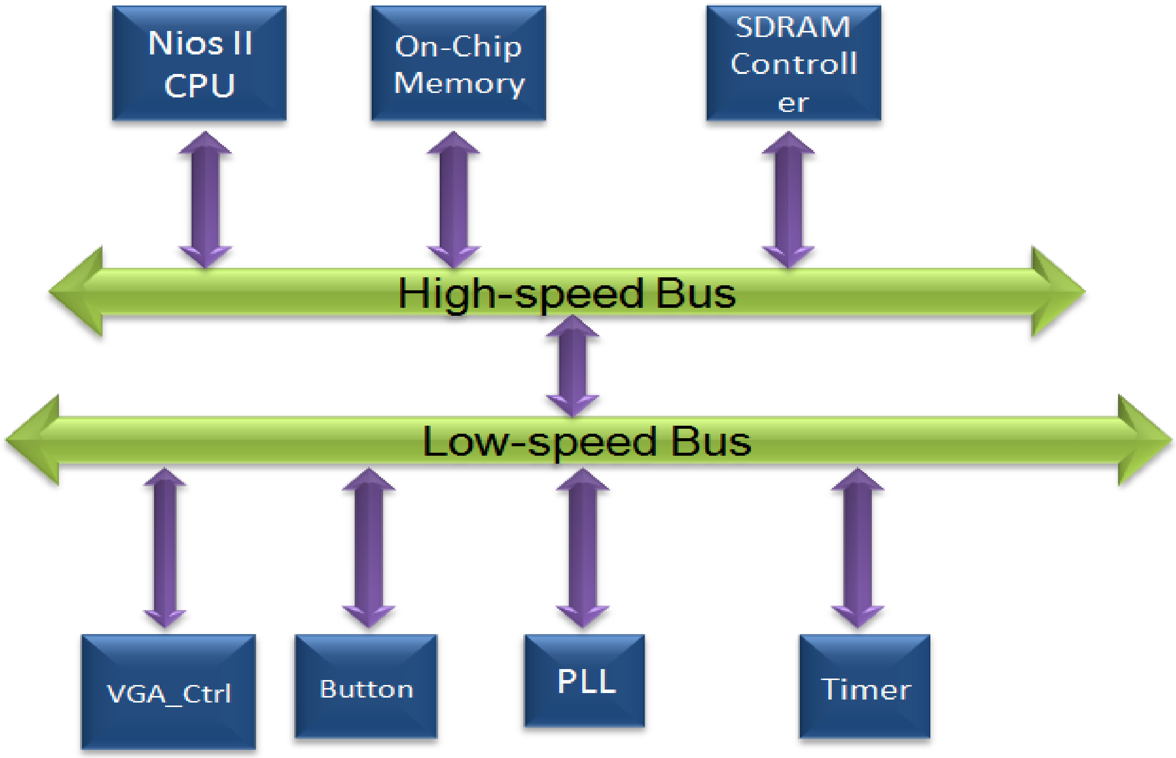


Fig.14

## B. *Software*

The software flow chart shown in figure 15.

## IV. SYSTEM FEATURES

### A. *Half-Adaptive Complexion Model*

The traditional fixed complexion model is not suitable for the intelligent reader system that is geared to individual user. Therefore, we propose the half-adaptive complexion model to segment face region, which proves to be more efficient. It needs to initialize the face complexion model whenever the user restart the system, which significantly improves the accuracy of face detection. Because the changes of background or environment while reading will result in the decrease of accuracy of face detection, which has already been taken consideration into the half-adaptive complexion model.

In addition, the face image used as input while initializing can also be used as a initial model for facial feature points. Hence, we needn’t search the specific face image from all kinds of images.

*B. Large Range of Movement*

In general, it is more difficult to identify dynamic objects than static objects and the range of movement should be limited. However, our system segments the rectangular area of face region first and turns it into a certain size by scaling, then determines the head posture by using the ratio of horizontal and vertical direction characteristic parameters. The above procedures both ensure the head to move around within a certain range.

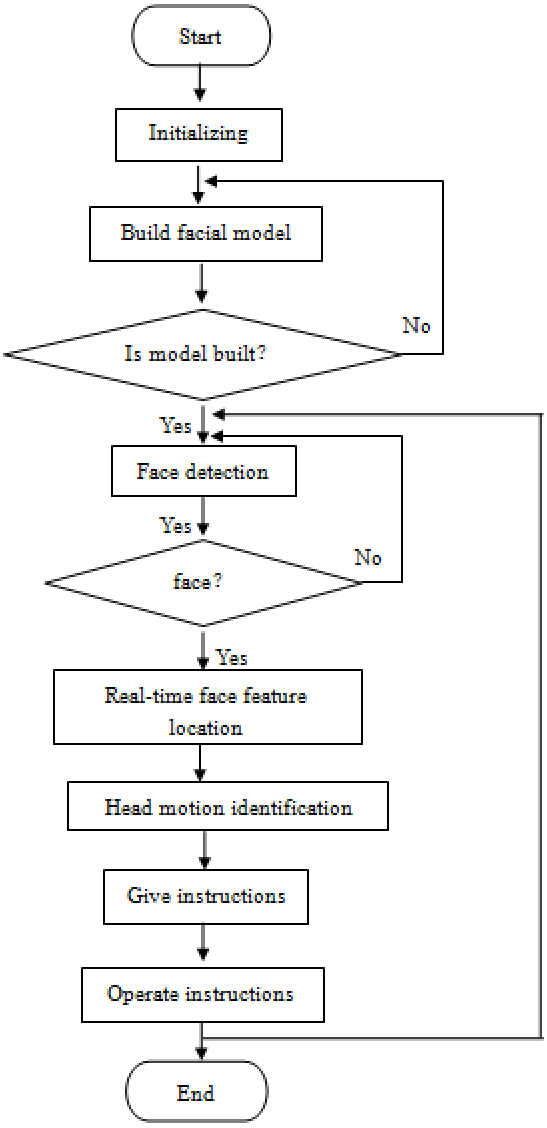


Fig.15

*C. Hardware-Software Co-design*

We realize the algorithm that with heavy computation burden in hardware, which can take full advantage of the characteristics of hardware in parallel and high speed. It makes sure that the system works with high efficiency and high real-time property.

## V. CONCLUSIONS

The idea of intelligent reader comes from the Samsung Galaxy S4 mobile phone which has developed smart scroll based on eye gaze tracking. However, after a preliminary trial, we find that using eyes to control the reader is not feasible in operation. On the one hand, the accuracy of eye gaze tracker is too low to control the reader correctly . On the other hand, there is no such a algorithm that can track the eye gaze with single camera while the head keeps moving. Therefore, users must keep their head static while using the intelligent reader, which is not suitable for long time reading. In order to solve the above problems, we propose a method to control the reader by using head postures. It demands the detection of eye feature and lip feature, used to determine the head postures then control the reader. We build the hardware architecture, at the same time, simulate and improve the algorithms. In a word, there is no doubt that as the technology is getting mature, the intelligent reader system will bring more exciting reading experiences for users in the near future.

## ACKNOWLEDGMENT

This work was supported in part by the Strategic emerging industry key technology special project of Guangdong Province (2011168014 and 2011912004).

## REFERENCES

[1] Eun Yi Kim, Sin Kuk Kang, Keechul Jung, Hang Joon Kim, Eye mouse: mouse implementation using eye tracking, ICCE on Consumer Electronics, 2005:207~208.

[2] Wen-Hsiang Lai, Chang-Tsun Li, Skin Colour-based Face Detection in Colour Images, IEEE International Conference on Video and Signal Based.

[3] Yan-Min Xu, Xing-Ming Zhang, Xiao-Hong Liu, Shu Su, Study of a Skin Segmentation Method Based on Color Clustering, Guangzhou Navigation Journal of Colleges and Universities ,2004,12(1):24~27.

[4] Wei-Xun Lin, Gang Pan, Zhao-Hui Wu, Yun-He Pan, Facial Features Localization Method Rivie, Chinese Journal of Image and Graphics, 2003,8(8):849~859.

[5] Xue-Wei Li, Xin-Rong Zhang, Fast Face Positioning and Feature Extraction Based on Gray level Difference Projection, Microprocessor, 2003,2:23~25.

[6] Zhi-Geng Pan, Peng-Cheng Zou, Rong-Hua Liang, Face Detection Based on Characteristics of Human Face and Statistical Skin Color, Journal of System Simulation, 2002, 16(6).

[7] Xin Tian, Skin Color Model Based on Different Color Space, Journal of Xi 'an Institute of Technology, 2001,21(4):369~372.