Deepak V. Katkoria and Alberto C. Arjona

3-D Facial Recognition System
Based on the MPC5121e microprocessor

Introduction

Research in computer graphics has brought attention to 3-D modeling. With advanced progress in image recognition models, designers have opened a whole new field of applications, including 3-D human face recognition. This article proposes a design and algorithm for a 3-D facial recognition system using Freescale's versatile MPC5121e microprocessor. Its triple-core architecture features an e300 Power Architecture® processor core, 2-D/3-D graphics engine (MBX) and audio processor (AXE) core.

Face detection algorithms

To measure distance via triangulation, the calculation uses the baseline distance between a laser beam and a camera as well as their angles to a target point (see Figure 1).
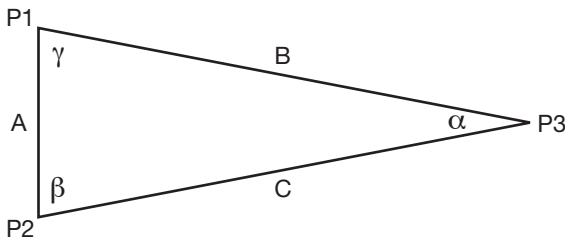
Triangulation Principle

Figure 1

P1 and P2 represent two reference points, the camera and laser beam, while P3 is a target point. The range B can be determined from the known values of the baseline separation A and the angles β and γ using the Law of Sines:

$$\frac{\sin \alpha}{A} = \frac{\sin \beta}{B} = \frac{\sin \gamma}{C} \tag{1}$$

In practice, this is difficult to achieve because the baseline separation and angles are hard to measure accurately. However, a demonstrated technique for obtaining range information via laser triangulation without knowing the values of A, β and γ has been developed. The algorithms of this technique could be successfully implemented in 32-bit processors, such as the MPC5121e multicore processor.

MPC5121e implementation

The MPC5121e integrated processor includes multiple cores and multiple buses that provide higher performance and allow for lower system cost and higher reliability.

The MPC5121e processor's rich set of integrated peripherals includes PCI, parallel advanced technology attachment (PATA), Ethernet, USB 2.0, twelve programmable serial controllers (PSC), display controller (DIU) and video-in unit (VIU), all of which fulfill the requirements for a facial recognition system.

The system can use the integrated display controller (DIU) to support an LCD display with a maximum resolution of 1280 x 720p and color depth of up to 24 bits per pixel. This will create an excellent image model for the user. Another advantage of DIU is its blending capability, which can be used to blend up to three different planes on the display. The system uses DIU for displaying the image model and for overlaying the data images to guide the user's decision making process.

The VIU plays a crucial role in the system's video interface. The VIU core accepts an ITU656-compatible video stream from the video camera, providing a wide selection of display modes, from QVGA to XVGA 8-bit/10-bit ITU656 video input.

The internal DMA engine transfers all incoming video data from FIFO to memory. This data is analyzed and computed using the triangulation algorithm mentioned above. Once the processor calculates the position of the vertex (point of interest) from the memory, then the matrix parameters are transferred to the MBX core. This is a 3-D accelerator that recreates a real-time rendering of the matrix (stored in memory) on the display.
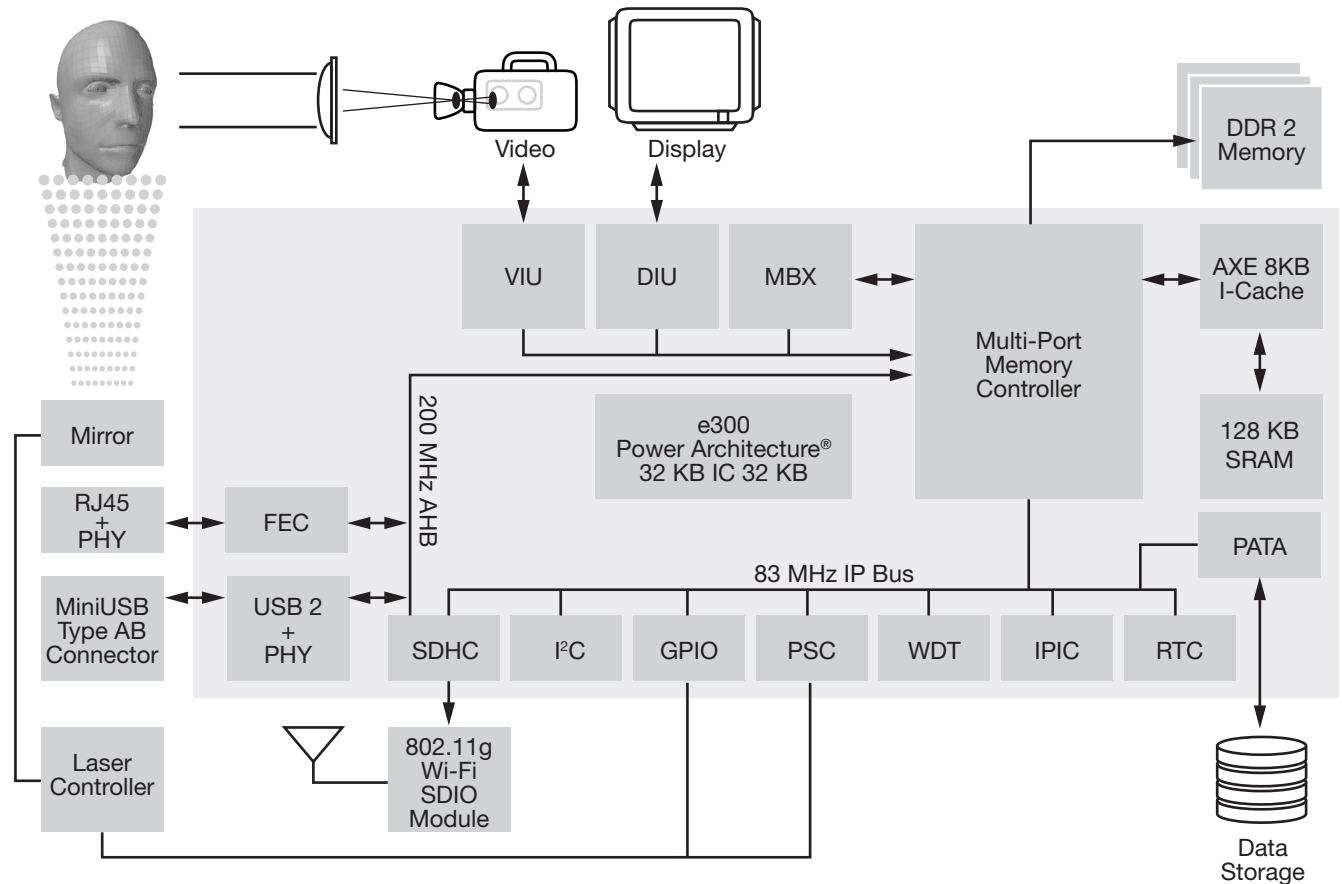
Figure 2

Image modeling is an important factor in a facial recognition system. Facial recognition, facial animation, video compression/coding, 3-D games, facial expression recognition, human action recognition for surveillance and object recognition are all image modeling outputs.

Users can store the display format and other useful information on a hard disk. The MPC5121e's PATA block offers two operation modes (PIO and DMA mode) that can be active at the same time.

A facial recognition system based on the MPC5121e processor can provide direct interface to display, video and storage using fewer components. The Power Architecture core and MBX core perform the post-processing for image reconstruction.

Designers can use MPC5121e's PSC and GPIO interfaces to control the laser's movement and intensity. The PSC can be used for serial communication between the laser unit and the MPC5121e processor while GPIO can work as on/off signals. Depending on the selected laser, a user can control its 180° circular movement with a timer, which turns the laser-positioning motor drive on and off to rotate the laser beam.

Real-time remote access and control of the facial recognition system from a central location is possible by using 100 Mbps Ethernet and a Wi-Fi interface.

In Figure 2, the VIU is a specific module where the video input is processed. It is used as a direct interface between the camera and the processor. The DIU is the module that controls the video output and LCD display. The laser control algorithms are processed by the MPC5121e. Some of the GPIOs and PSC are used as interfaces between the laser controller block and the e300 core.
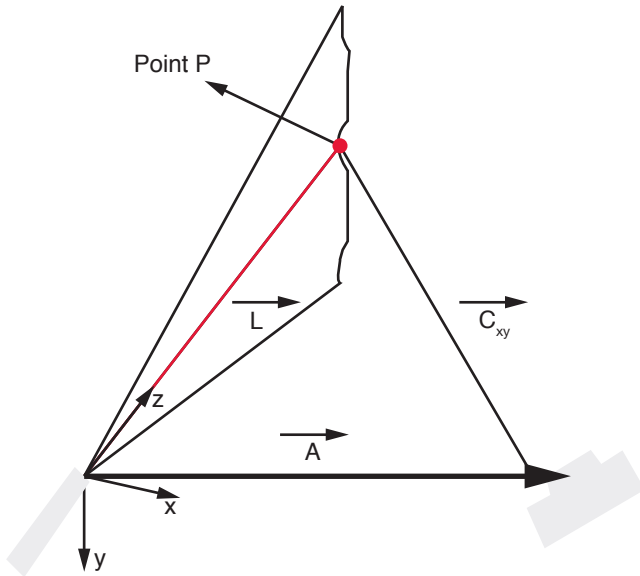
Figure 3

General triangulation equations

The elements of the data acquisition system along with the point where the object is illuminated by the laser form a triangle of vectors (Figure 3). The objective of laser triangulation is to find the spatial coordinates of the illumination point (Point P) in a defined frame of reference.

Where,
L is the vector between the origin of the laser beam and point P
A is the vector between the origin of the laser beam and the optical center of the camera
C is the vector between the optical center of the camera and point P

The geometrical relation is expressed in the following vectorial equation:

$$\vec{L} = \vec{A} + \vec{C}_{xy} \tag{2}$$

From this equation, vector A as well as the directional cosines from the vector Cxy are known (see Camera Calibration section below).

Equation (2) expressed in spherical coordinates, where φ is the angle between the vector L and X-Y plane and θ is the angle between the laser beam plane and X axis, is:

$$|L|\cos\varphi\cos\theta = A_x + Cxy_x \tag{3}$$

$$|L|\cos\varphi\sin\theta = A_y + Cxy_y \tag{4}$$

$$|L|\sin\varphi = A_z + Cxy_z \tag{5}$$

It is a non-linear system. Solving these equations is a problem of high computational cost. Nonetheless, by placing the reference system Z axis along the laser beam plane ($\varphi = 0$), the system is reduced. The new system of equations is given by:

$$L_x = |L|\cos\varphi = A_x + Cxy_x \tag{6}$$

$$0 = A_y + Cxy_y \tag{7}$$

$$L_z = |L|\sin\varphi = A_z + Cxy_z \tag{8}$$

The analytic solution of the set of equations (6) to (8) is given by:

$$L_x = |L|\cos\varphi = A_x + |Cxy|\alpha \tag{9}$$

$$|Cxy|\gamma = -\frac{A\gamma}{\beta} \tag{10}$$

$$L_z = |L|\sin\varphi = A_z + |Cxy|\gamma \tag{11}$$

…where α, β, γ are the directional cosines of the vector Cxy, in the laser's frame of reference.

These vector components are found using the camera's intrinsic and extrinsic parameters. Because of the importance of the camera location and its space orientation, it is necessary to find a transformation matrix. In general, the elements of a transformation matrix are the angles between each unitary vector from the original frame to the new frame. If both frames fit in one common axis, then the parameters are simplified in just one angle, as illustrated in Figure 4. This is the case for this application.
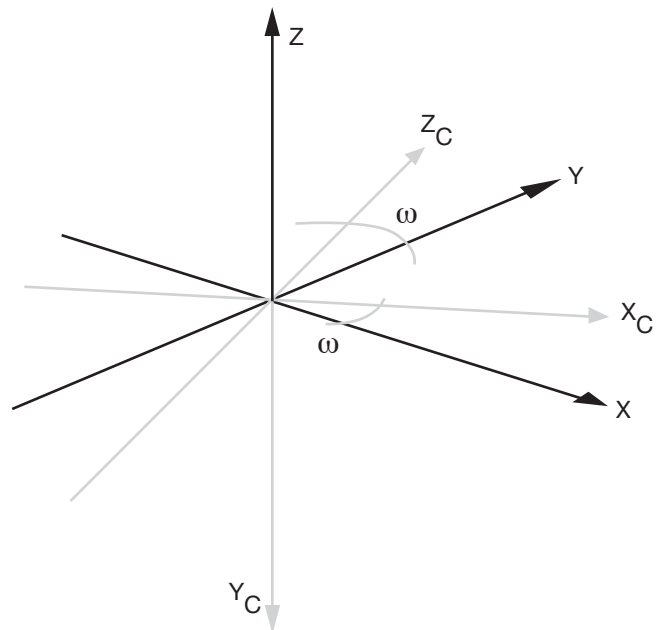
Figure 4

The parameter in Figure 4 is represented in the following transformation matrix:

$$\begin{bmatrix} \alpha\hat{x} \\ \beta\hat{y} \\ \gamma\hat{z} \end{bmatrix} = \begin{bmatrix} \cos\omega & 0 & \sin\omega \\ -\sin\omega & 0 & \cos\omega \\ 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} \alpha_c\hat{x}_c \\ \beta_c\hat{y}_c \\ \gamma_c\hat{z}_c \end{bmatrix} \tag{12}$$

Replacing (12) into (9) to (11), the system becomes:

$$L_x = |L|\cos\varphi = A_x + |Cxy|[\alpha_c\cos\omega + \gamma_c\sin\omega] \tag{13}$$

$$|Cxy_y| = -\frac{A_y}{\gamma_c\cos\omega - \alpha_c\sin\omega} \tag{14}$$

$$L_z = |L|\sin\varphi = A_z + |Cxy|[-\beta_c] \tag{15}$$

Because the laser beam path projects a line on the exploration surface, a rotating mirror is used to project the line along this surface (in this case a human face). This implies a mobile frame of reference in the laser beam, which then means the information needs to be referenced to a fixed system. In general, a mobile frame has translation and rotational components. Conventional transformations are inherently complex and difficult to drive. A very good alternative is to use a homogeneous transformation matrix. This is a block matrix that separates rotations and translations in two isolated partitions.

$$\begin{bmatrix} L_x \\ 0 \\ L_z \\ 1 \end{bmatrix} = T(\psi)_{4x4} \begin{bmatrix} r_u \\ r_v \\ r_w \\ 1 \end{bmatrix} \tag{16}$$

If the reference system is located in the laser, the origin is defined as OUVW and the fixed system of reference as OXYZ, and $(r_u, r_v, r_w)^T$ is the triangulated point $(0, L_y, L_z)^T$, using equations (13) to (15). Transformation matrix T describes, in general, the dynamics of the mobile mirror frame in a partition of rotation and a partition of translation.

$$T(\vartheta, \phi, \psi, x, y, z)_{4x4} = \left[\begin{array}{c|c} \gamma_{3x3}(\vartheta, \phi, \psi) & \mathbb{T}_{3x1}(x, y, z) \\ \hline 0_{1x3} & 1_{1x1} \end{array}\right] \tag{17}$$

Where,

γ_{3x3} is the rotation partition

$\mathbb{T}_{3x1}$ is the translation partition

ψ is the angle of rotation of the mobile mirror on the z component of the fixed frame

Camera calibration

The objective of calibrating the 3-D laser scanning is to find the parameters of the triangulation equations (13) to (15). Those parameters are the vector A, the directional cosines of the vector Cxy and the parameters of the homogeneous transformation matrix T.

Although there are various ways to model a camera, in computer vision the pinhole model is often preferred because of its simplicity.[1] Finding the parameters of the model is a problem of optimization, specifically a parameter estimation problem. However, it is possible to find several algorithms that solve the problem in an iterative way.[2] It is important to take into consideration that image deviations in cameras are founded in lateral extremes due to chromatic and radial distortion in the lenses. Many calibration techniques omit these deviations for simplicity.

Output

After digitizing the model, the output files can be manipulated using OpenGL-ES and displayed using an embedded controller, such as the MPC5121e processor. [See application note "3-D Graphics on the ADS512101 Board Using OpenGL ES" at **www.freescale.com** (search for AN3793)]

All the digitized points can be stored in a data base. For face recognition, these points are the comparison parameters that are different for every face scanned. Traditional techniques are based on extracting landmarks, or features, from a two-dimensional (2-D) image of the subject's face. These features are affected by changes in lighting, relative position and perspective variations. Some of these traditional techniques are based on the Eigenface method, neural networks or hidden Markov models. These techniques are very complex and expensive.

Application and Driver Tools and Software

ID	Description	Software and Tool Type	Vendor
CW-MOBILEGT	Integrated tool suite environment for mobileGT® applications. *Board support packages are offered free of charge, as-is.	IDE: debug, compile and build tools	Freescale
KINETIC ECG ALGORITHIM	The Kinetic ECG Algorithim provides signal processing and interpretation of the ECG waveform, thereby aiding health care professionals in assessing cardiac parameters	Protocol stacks and middleware	Monebo
ADS512101	MPC5121e, ADS5121, Mini-ITX, USB, audio, video, Ethernet, ATA, secure digital, automotive, consumer, industrial, portable	Evaluation/development boards and systems	Silicon Turnkey Express
MPLUS5121	MPC5121e, Mini-ITX, Media5200, LCD, Ethernet, PATA, SATA, CAN, DVI, mobileGT, telematics, automotive, ADS512101, graphics	Hardware components	Silicon Turnkey Express
MBX-SDK	Available from FAE and marketing team	Software development kit based on Linux® OS	Open Source

Table 1

On the other hand, the 3-D facial recognition system measures the positions of different points on a face and uses this information to create a 3-D surface image that contains the distinctive features of a specific face. The principal advantage of 3-D facial recognition is that it can identify a face from a number of viewing angles. In addition, it is not affected by changes in lighting as are the traditional techniques mentioned previously.

Tools and software

For developing applications and drivers, Freescale and its partners provide the tools and software listed in Table 1. Soon, additional drivers are expected to be available from other third-party vendors.

Conclusion

A new facial recognition technology based on Freescale's MPC5121e processor can provide a low-cost solution to help ensure public safety. The MPC5121e processor is a multi-featured solution that can help designers transform their ideas into reality, turning a big box facial recognition system into a compact, low-cost device for security applications.

To learn more about the MPC5121e processor, visit **www.freescale.com/mobilegt** or contact your local FAE. (Reference manual URL: **www.freescale.com/32bit**)

References

[1] "Multiple View Geometry in Computer Vision," Richard Hartley and Andrew Zisserman, Second Edition, Cambridge University Press, March 2004

[2] Applied Mathematics and Computation ISSN:0096-3003, Rudolf Scitovski, Marcel Meler

Deepak V. Katkoria is an IMT application engineer, formerly with Freescale Semiconductor. He has his degree in engineering from Pune University. While at Freescale, Deepak supported the MPC5200B and MPC5121e product families focusing on graphic support. Alberto Arjona Cabrera (Dmitri XIII) is a university professor and Freescale application engineer in Guadalajara, Mexico. Throughout his Freescale career he submitted several mathematical algorithms related to computer vision and signal processing applications. He received a master of science degree with a control theory specialty and has previously served as a captain in the Mexican Air Force.