ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

Effective Geometric Calibration and Facial Feature Extraction Using Multi Sensors

MOHD NORZALI Haji Mohd*, Masayuki KASHIMA**, Kiminori SATO**, Mutsumi WATANABE**
**Department of Information Science and Biomedical Engineering, Graduate School of Science and Engineering , Kagoshima University, Korimoto 1-21-40,Kagoshima,890-0065,Japan
*Department Of Computer Engineering, Faculty Of Electrical and Electronic Engineering, University Tun Hussein Onn Malaysia (UTHM),Parit Raja, Batu Pahat Johor,Malaysia.

Abstract*— This paper aims to present facial feature extraction by integrating 3 different sensors that might be used in the estimation of internal mental state. RGB-D camera is used at the pre and post monitoring phase while thermal infrared and visible camera is being used in the stimulus experiment. The measurement of three facial areas of sympathetic importance through purely imaging means that is periorbital, supraorbital and maxillary is done on the second stage. An Accurate and efficient thermal-infrared camera calibration is important for advancing computer vision research approach for geometrically calibrating individual and multiple cameras in both thermal and visible modalities. We also propose new printed Fever Cold Plaster (FCP) chessboard using a popular existing approach which is comparatively accurate and simple to execute. Based on the experiment conducted by comparing the degradation of image quality with the current approach, our proposed chessboard can be more clearly located than those on the applied standard chessboard by 39%.

***Index Terms*— Thermal-Infrared Geometric Camera Calibration, Thermal Face, Facial Feature Extraction.**

I. INTRODUCTION

Various methods for internal state measurement such as mental stress have been previously proposed which utilizes the changes of physiological quantities such as blood pressure, heart rate and salivary amylase. These quantities can be measured noninvasively by the use of inexpensive instruments. In this paper we introduce an integrated noninvasive measurement through purely imaging means. The propose booth setup is equipped with Thermal, Visual and RGB-D camera Sensors.

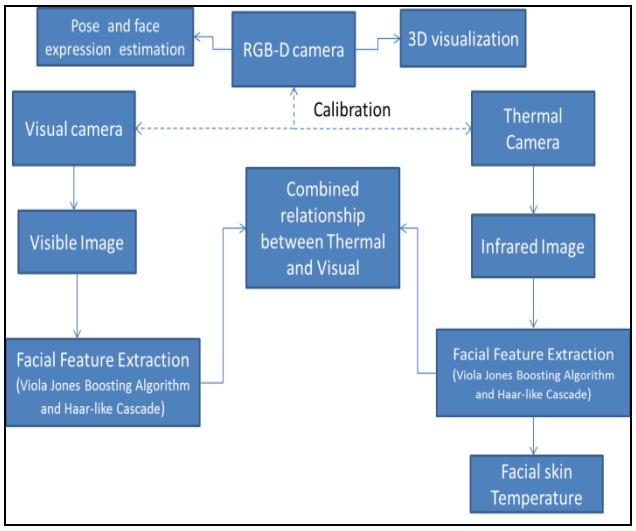


Fig.1. Overall system

A monitor and a speaker are also installed to provide mental and acoustic stimulus to monitor changes in facial sign and brain waves. These experiments will be done at a later stage. This paper focuses on the experimental Booth setup, Thermal-Visual Camera calibration and facial feature extraction in thermal, visual and RGB-D camera.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

Fig.2. Camera setup (Thermal Camera (Right): NEC TH7800, Visual Camera (Left): USB CMOS Imaging Source DFK22AUC03 and Windows RGB-D camera (Upper))

II. RELATED WORK

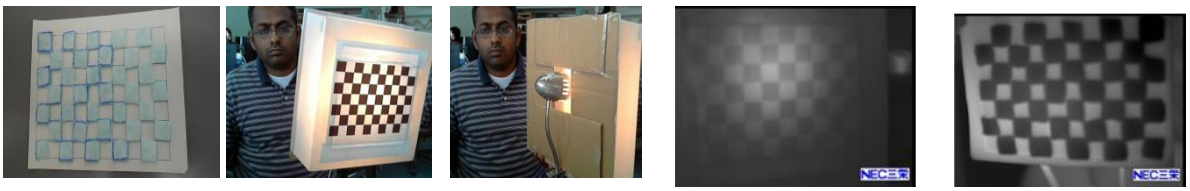
A proven approach for the accurate calibration of the intrinsic and extrinsic parameters of cameras involves the use of a planar calibration pattern or a 3-D calibration object. The known geometry of the pattern or object is exploited in order to determine how the camera views the world .A sequence of images of the calibration pattern or object from different angles is typically required by most calibration algorithms [3].
The most popular calibration pattern is a regular printed chessboard. This pattern has been demonstrated to be highly effective and convenient for the task of calibrating visible spectrum cameras.

III. PROPOSED SYSTEM

1. Registration of IR and Visible

A. Calibration Board

Accurate geometric camera calibration must be performed for both single and multiple devices imaging system as a precursor to many existing computer vision algorithms. Algorithms that greatly benefit from accurate geometric calculation include those for feature matching, localization and 3D reconstruction. Such algorithms form critical components in system applied in domains such as surveillance, medical assessment and industrial inspection. The different characteristics of thermal-infrared cameras make them superior or complementary to conventional visible cameras from many applications. For example, thermal-IR cameras are more robust to poor lighting or lighting changes and also affect the process of calibration and therefore the accuracy of result. The proposed system is an adaptation of Zhang`s method. The calibration system [3] techniques are flexible, robust and low cost. The techniques only require the camera to observe a planar pattern shown at a few different orientations.
The most popular calibration pattern is a regular printed chessboard [3]. This pattern has been demonstrated to be highly effective and convenient for the task of calibrating visible spectrum cameras. The calibration points are easily located in the image with high precision due to high image contrast.
With some small adjustment and preparation where cold fever plaster is attached to the back of the the calibration points on a chessboard, the calibration points can be reliably located in a thermal –Infrared image. The proposed method enables the use of the calibration chessboard for thermal camera calibration involves heating the pattern through exposure to a flood lamp [4]. This result in the appearance of the pattern in thermal modality resembling its appearance of the pattern in the visible modality. Fig.2. Shows the quality of the image that could be produced using the heated chessboard pattern attached to cold fever plaster.



(A) (B) (C) (D)

Fig. 2. (A) Chessboard attached with Fever Cold Plaster (back View), (B) proposed collaboration board attach with flood lamp, (C) Thermal image chessboard (light source 500 W,conventional chessboard) , (D) Thermal image chessboard (light source 500 W, proposed Chessboard)

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

B. Calibration of Thermal IR , Visible and RGB-D camera Camera

The output of the calibration method includes the relative rotation R and translation T (Fig. 4 (a)) of the cameras with respect to the left hand view and the internal parameters of each camera, focal length f_c, principal point c_c ,skew coefficients α_c and radial and tangential distortion k_c.Following stereo calibration,coordinates of a 3D point, $p_c = [x_c, y_c, z_c]$ as follows $P_c = \boldsymbol{T} + \mathbf{R}P$, where *R* and *T* are their respective relative rotation and translation with respect to the world coordinates. Here, for the surface reconstruction procedure , we make use of the image point projection of the scene normalized so as to follow the pinhole camera model [5], Let the normalized (pinhole) image projection, $p_n = [x, y]$ be given by:

$$p_n = \begin{bmatrix} x_c/z_c \\ y_c/z_c \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}. \qquad [1]$$

After including the lens distortions, the new normalized point coordinates $p_d = [x_d, y_d]$ is obtained

$$p_d = \begin{bmatrix} x_d \\ y_d \end{bmatrix} = (1 + kc(1)r^2 + kc(2)r^4 + kc(5)r^6)p_n + dx \qquad [2]$$

Where $r^2 = x^2 + y^2$, and dx is the tangential distortion vector

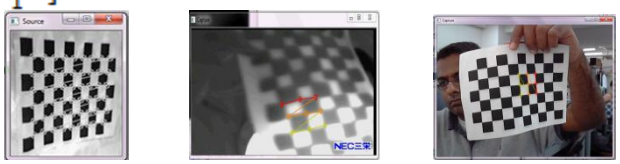
$$d_x = \begin{bmatrix} 2kc(3)\,x + kc(4)(r^2 + 2x^2) \\ kc(3)(r^2 + 2y^2) + 2kc(4)xy \end{bmatrix}$$

With these ingredients, we can relate the normalized coordinate vector, p_d ,and the pixel image coordinates, x_d and y_d as follows

$$\begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = K \begin{bmatrix} x_d \\ y_d \\ 1 \end{bmatrix} \qquad [3]$$

Where K is known as the camera parameter matrix,which can be expressed making use of the calibration output variables as

$$K = \begin{bmatrix} f_c(1) & \alpha_c * f_c(1) & cc(1) \\ 0 & f_c(2) & cc(2) \\ 0 & 0 & 1 \end{bmatrix}$$



(A) (B) (C)

Fig. 3. Collaboration based on Zhang using the proposed chessboard (A) RGB-D camera (Width:3 Height:3) for (B) Thermal and (C) Visual camera.

C. Stereo Vision

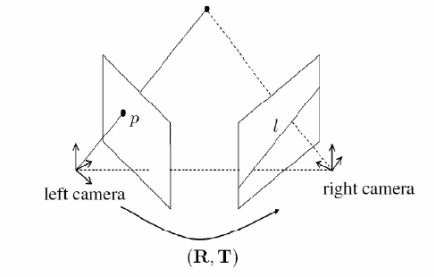
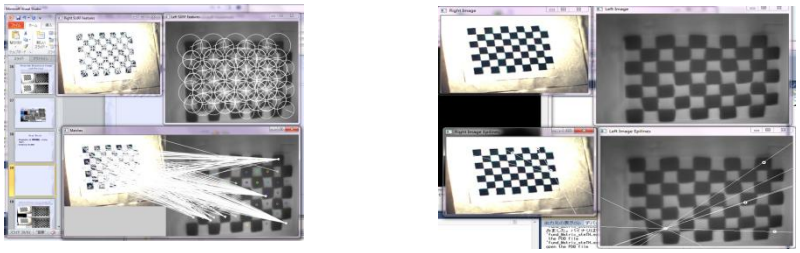


Fig. 4. (a) Triangulation and Epipolar Geometry

The principle of computational stereo [6] is well known in computer vision community due to its wide range applications. Stereo vision is the technique to calculate Depth or 3D coordinates of a number of points is a scene given at least two images. One of the common strategies to simplify the correspondence problem is to exploit epipolar geometry constraints, as illustrated in the figure 4 below. Given a pair of calibrated camera, for a point in the visible view is constrained to lie on the corresponding epipolar line. This reduces the search space for the whole image to a line [7].

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

The steps that require in computing the fundamental matrix between two calibrated stereo images are to detect SURF feature points and select 7 good matches. After that, Using openCv cv::findFundamentalMat functions to compute the fundamental matrix from this 7 matches (Fig. 5.)



(A) **(B)**

Fig. 5. (A) Detecting SURF Feature Points between Calibrated Stereo Images. (B) Computing Fundamental Matrix and Drawing Epipolar Lines from Thermal-Visual Stereo Calibrated Images

2. Facial Feature Extraction

A. Thermal and Visual Facial Feature Extraction

Thermal face image analysis has many applications such as sensation evaluation and face recognition. Facial feature extraction in the IR image is an essential step in these applications. Certain facial areas such as periocular,nasal,cheeks and neck region produce different thermal patterns for different activities or emotions[8].Skin temperature of facial features, such as the nose and forehead, could be an effective indicator in objectively evaluating human sensations such as stress and fatigue [9][10].Most existing approaches manually locate the facial feature in IR image or the subject are required to wear marker[11], as it is hard to automatically locate facial features, even for the obvious features such as the corners of the eyes and mouth. This problem is caused by poorer contrast between the features and the face in IR images. In our experiment, we have found out that using the propose face detection using Haarcascade in both Thermal and Visual can detect faces and eye region automatically. The flowchart on how to locate a face region and the temperature is as Fig.5.

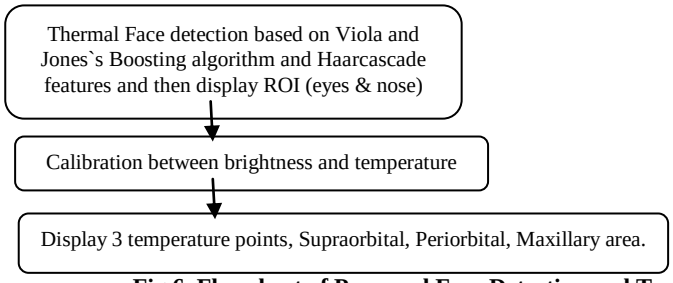


Fig 6. Flowchart of Proposed Face Detection and Temperature Acquisition in Thermography

In our proposed approach, at first, we detect the face region using Viola and Jones`s Boosting algorithm [12] and a set of Haar-like cascade features. Then the eye area is minimized by assuming the eyes estimated position to be at the upper part of the face. Next the centroid of the face is located as the nose based on ratio given in figure 6.

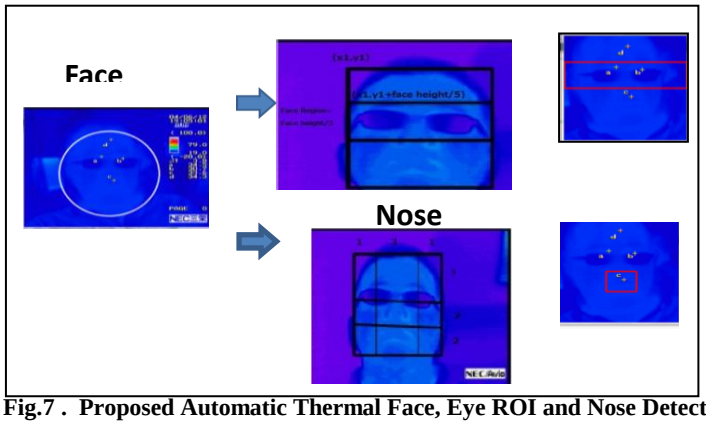


Fig.7 . Proposed Automatic Thermal Face, Eye ROI and Nose Detection

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

Having known the the eye ROI and nose, the location and temperature value of the appropriate region as figure 7 are then detected.When sympathetic activation occurs such as mental stress, blood pressure, heart and breathing rate elevated. The supraorbital area is correlated to blood supply to the corrugator muscle [13].The per orbital correlate to blood supply to orbital muscle [14],[15] whereas the maxillary area correlate to blood perfusion in the maxillary area [16]. Based on the past research ,we have found out that 3 ROI namely Supraorbital, per orbital and Maxillary region are the most affected during Mental Stress and focuses on the temperature in these areas (see Fig.7) .The Average brightness value located around these 3 main areas is then converted to the temperature value (figure 8).

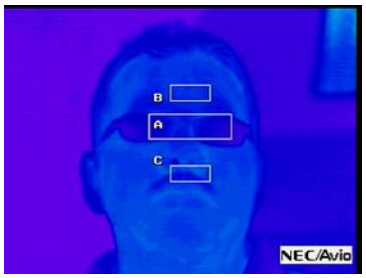


Fig. 8 Detection of Location and Temperature Value around
B: Supraorbital, A: Periorbital and C: Maxillary areas.

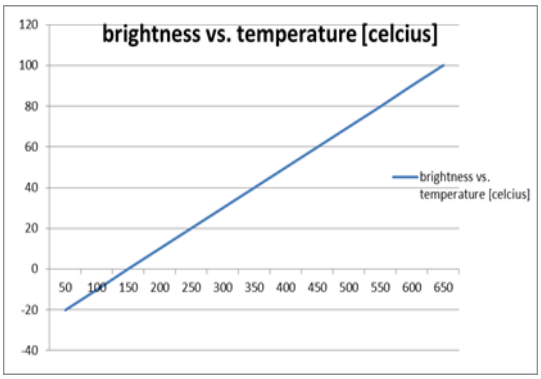


Fig. 9. Calibration between Brightness and Temperature (-20°C- 120° C)

The same approach for locating face image in Visual face was also done. At first we detect the face using Viola and Jones` Boosting Algorithm and a set of Haar- like cascade features. Next, Haar like Feature cascade was also used to detect Eyes ROI (see Fig. 9).

Fig. 10 Detection Result for Face and Eyes ROI

Viola and Jones`[12] face detection algorithm is based on the Haar-like features. The four types of rectangular box are used to extract the features. The shaded area of the rectangular regions is subtracted from the white area of the rectangle. Figure below shows the four types of Haar-like features used for feature detection. Upper two rectangles give edge features, below left gives line features and the right one gives rectangle features.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

C. RGB-D camera Face Feature Extraction

RGB-D camera camera can easily locate human head and estimate head pose. A depth AAM algorithm is developed and applied to to locate the detailed face structure. The depth AAM algorithm [17] takes four channels R, G, B and D which combines the color and depth of input images.

1. Head Pose Estimation

The depth imaging technology has advanced dramatically over the past few years, especially with the launch of RGB-D camera. Pixels in depth image indicate the calibrated depth in a scene, rather than a measure of intensity of color. Randomized decision trees and forests are effective multiclass classifiers for human pose recognized. A forest is an ensemble of T decision trees, each containing of split and leaf nodes. Each split nodes Consists of a feature f_θ and a τ threshold. At the leaf node reach in the tree t, a learned distribution $P_t(c|I,x)$ over body part labels c is stored. The distributions are averaged together for all the trees in the forest give the final classification: $P(c|I,x)=\frac{1}{T}\sum_{i=1}^{T} P_t(c|I,x)$

Because the facial information of depth information is not sufficient, we can`t locate the facial feature effectively. That`s why we need to apply AAM algorithm to locate facial feature which can use texture and depth information comprehensively. The four upper body parts converging the head are merged to localize the head. The head poses describe for $(X_{head,}, Y_{head,} Z_{head})$ YYZ coordinates.

2. Depth AAM Algorithm

The shape of an AAM algorithm is composed of the coordinates of the vertices that makes up the mesh: $S = (X_1, Y_1, X_2, Y_2, \ldots, X_n, Y_n)^T$ This means that the shape s_0 can be expressed as a base shape plus a linear combination of n shape vectors S_i :

$$s = s_o + \sum_{i=1}^{n} P_i S_i$$

In this expression the coefficients P_i are the shape parameters, and for the vectors S_i are just the orthornormal eigenvector obtained from the training shapes. We apply Principal Component Analysis(PCA) to the training mashes with the hand labeled training images. The base shape s_0 is the mean shape and the vectors s_i are the n eigenvectors corresponding to the largest eigenvalues.

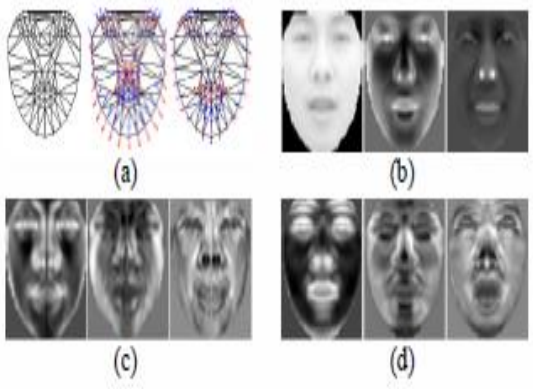


Fig.11. Multi-Band Aams (A) Mean Shape and The First Two Shape Bases Learned By PCA. (B) To (D) Are Mean Appearance and the First Two Appearance Bases of the Intensity Band, X-Direction Band and Y-Direction Band Respectively.

1. The depth AAM algorithm is applied using the following algorithm Get pairs of texture and depth images with RGB-D camera camera and train the appearance models with depth information
2. **while** get texture and depth images with RGB-D camera camera **do**
3. Recognize human pose using equations
4. **if** human pose recognized, then segment facial images and estimate head pose $(X_{head,}, Y_{head,} Z_{head})$
5. **if** head pose sufficient than fit the input data (texture and depth images)
6. **end if**
7. **end if**
8. **end while**

In our research, the face tracking and rendering with mask superimposed .The facial expression is then map into an egghead to visualize user expression (fig.12)

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

Fig.12 Visualizing User Expression Using RGB-D Camera

IV. EXPERIMENTS

To evaluate the alternative approaches and the effectiveness of the proposed calibration board, MRE (Multiple Regression *Equations) is used as the metric for comparison.* The MRE (Multiple Regression *Equations)* are used as the metric for comparison. In general .It is calculated using only frames included in the calibration process. We choose to calculate MRE using entire input that has not actually been selected for calibration. The units are in pixels.

$$\text{MRE} = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N}||p(m,n) - q(m,n)||}{MN}$$

M, is the total number of frames in the extended sequence, and N is the total number of calibration points per frame.p(m,n) is the pixel location of a point in the pattern, and q(m,n) is the projected location of that point using an estimate of the pose of the pattern in the particular frame.

An experiment was conducted in a normal office environment, with an ambient temperature of approximately 24°C.The calibration objects were placed at the minimum focal distances for which the pattern was fully visible, approximately 50cm from the lens. A workspace approximately 1m³ was required for the experiment.

The proposed calibration board is evaluated and compared with existing approaches for calibrating thermal and visual cameras. 100 calibration frames of each pattern were used to compare performance.

2 types of chessboard were printed on an A4 sheet paper , One type Which is the conventional chessboard and another which fever plaster which is in the cool gel form was attached at the back of the chessboard as in figure 2(A). A 500 W heat lamp was used for approximately 5s to heat the pattern as even as possible. Footage was then captured immediately for approximately 10s, at which point the image contrast had degenerated significantly. An identical software framework was used for both method.OpenCV function "findChessboardCorners ()"[18] was used.

Table 1. Pattern Effectiveness (MRE and Standard Deviation)

Configuration	MRE
Thermal (chessboard)	**0.814±0.015px**
Thermal (*FCP*)	**0.614±0.011px**

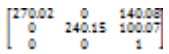
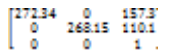
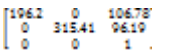
From the above table, it is clear that calibration points on the proposed chessboard *FCP* can be more clearly located than those on the standard chessboard.

Table 3. Intrinsic parameters of the three cameras

	Visual vs. Thermal	RGB-D camera vs. Thermal
Rotation Matrix	$\begin{bmatrix} 0.999938 & 0.018628 & 0.020553 \\ -0. & 315.41 & -0.023347 \\ -0.020975 & 0 & 0.99979 \end{bmatrix}$	$\begin{bmatrix} 0.99727 & -0.039956 \\ 0.054310 & 0.99979 \end{bmatrix}$
Translation Vector	$\begin{pmatrix} -79.829 \\ -0.92904 \\ -0.023286 \end{pmatrix}$	$\begin{pmatrix} 650.08 \\ -13.817 \\ 248.66 \end{pmatrix}$

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

Table 4. Extrinsic Parameters of the Three Cameras Relating to Its Position

Camera	Thermal Camera(right)	Visual camera(left)	RGB-D camera Camera(upper)
Matrix			
Optical Distortions coefficients	K1: -0.39235 K2: 0.25438 P1: 0.0011792 P2: 0.0051091	K1: -0.42242 K2: 1.1506 P1: -0.001509 P2: 0.0004468	K1: -0.82349 K2: -2.8381 P1: -0.0085913 P2: 0.00075683

V. DISCUSSION

The comparison between traditional heated chessboard and *FCP* is as table 1.The cost involved in manufacturing *FCP Chessboard is* considered cheap in comparing to other calibration board made from polished copper plate coated with high emissivity paint or calibration rig, Even though heating is required in both method, it is difficult to get an even coverage comparing to our method. The pattern also can last longer about 20 minutes

Table 5. Evaluation of Proposed Approach

	Chessboard	**FCP Chessboard**
Manufacturing	**Printable from a standard printer**	**Printable from a standard printer with cheap fever cold plaster**
Heating	**A powerful (500W+) flood lamp and an external power supply are required. Difficult to get even coverage.**	**A powerful (500W+) flood lamp and an external power supply are required. Easy to get even coverage.**
Footage	**The pattern is only effective for a few seconds after heating.**	**The pattern can be used easily and effectively for about 15-20 minutes.**
Searching	**Generally requires preprocessing (inversion/ thresholding) Many conventional algorithms will struggle to find the pattern automatically**	**The algorithms are very common. No preprocessing required.**
Accuracy	**Less (see experiment)**	**Higher (see experiment)**

VI. CONCLUSION

Facial Feature extraction for thermal, visual and Kinect was presented. A new simple and effective calibration board has been presented and compared with the existing board. The pattern is cheap to manufacture and result shows that it can obtain 39% more effective that the conventional heated chessboard to be used for thermal camera calibration. The calibration system calculating intrinsic and extrinsic parameter was also presented for thermal , visual and Kinect sensors. Future works include monitoring the mental state of a person by analyzing the facial feature.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 1, Issue 2, November 2012

ACKNOWLEDGMENT

This work was partially funded by the Japan Society for the promotion of Science (JSPS), Kagoshima University and UTHM. We would like also to give special thanks to the Laboratory member for their invaluable inputs and assistance.

REFERENCES

[1] L. Pavlidis, J. Levine and P.Baukol," Thermal imaging fro anxiety detection,"Proc. IEEE Workshop Computing Visual beyond Visible Spectrum: Methods Appl. 2000, pp, 104-109.

[2] L.B.Wolff, D.A. Socolinsky and C.K.Evel ,"Quantitative measurement of illumination invariance fro face recognition using Thermal Infrared Imagery "in Proc.SPIE Vol 4820, Inferred Technology and Applications XVIII,B.Andresen ,G.F.Fulop,M. Stronmjnik,Eds,Jan 2003,pp.140-151.

[3] Z.Zhang,"A Flexible New Technique for camera calibration", IEEE Trans. Pattern Anal. Mach. Intell., Vol.22, no. 11, pp. 1330-1334, Nov.2000.

[4] Prakash S.,.Lee P.Y.,Caeli T.,and Raupach T.," Robust Thermal camera calibration and 3D mapping object surface temperatures".Thermosense XXVIII, 6205:62050J1-62050J8,2006.

[5] Hartley R. and Zisserman A., "Multiple view geometry in computer vision". Cambridge University Press, 2000.

[6] Koschan A., "What is a New in Computational theory since 1989: A survey of current stereo paper."Technical Report, 93-22, 1993.

[7] Faugeras O., "Three Dimensional Computer Vision: A geometric Viewpoint". The MIT Press, 1993.

[8] InterSense Inc, InertiaCube [Online] www.intersense.com

[9] H.enno.A.Saijo, H.yoshida, R.Suzuki and M. Osumi,"Using Facial temperature to objectively evaluate sensation"Inter. Journ. Ind. Ergonomics, Vol. 19, no 2, pp161-171 Feb. 1977.

[10] H.enno.A.Saijo, H.yoshida, R.Suzuki and M. Osumi," Non- contact method for measuring facial skin temperature", Inter. Journ. Ind. Ergonomics, Vol. 19, no 2, pp147-159 Feb. 1977.

[11] Jian-Gang Wang," Facial Feature Extraction in an infrared Image by Proxy with a Visible Face Image"IEEE Transc. On Instrumentations and Measurement, Vol 56, No 5, pp 2057-2066, Oct. 2007.

[12] P.Viola and M.Jones,"Robust real time object detection", Inter.Journal of Computer Vision, Vol.57, No 2, 2002.

[13] C.Puri,L.Eberharndt and J.Levine," StressCam:Non- Contact measurement of users' emotional states through thermal imaging,"inProc. Of ACM Conf. Hum. Factors Computing System (CHI), 2-5, 2005, pp.1725-1728.

[14] J.Levine, I.Pavlidis, and M.Cooper," The face of Fear", Lancet, vol. 357, no 9270, p.1757, Jun 2001.

[15] I.Pavlidis, N.L. Eberhardt and J.Levine," Human Behaviour: Seeing through the face of deception", Nature, vol.415, no 6867, p.35, Jan.2002.

[16] Dvijesh et.al, "Imaging Facial Signs of Neurophysiological Responses"IEEE Transactions on Biomedical Engineering Vo.56, No.2, February 2009.

[17] Qui Jin et. Al."Facial Feature Extraction with a Depth AAM Algorithm", Inter. Conf. on Fuzzy System and Knowledge Discovery (FSKD 2012), 1792-1796.