Adaptive Circular Enclosure Colour Distribution Geometrical Model Utilizing Point In-Polygon for Segregation Between Lips and Skin Pixels

Lee Hao Wei, Seng Kah Phooi, Chia Wai Chong, and Ang Li-Minn

Abstract---**This paper is inspired from various boundary determination techniques which are used for segregating colours between background, skin and lips. Basic concept for this technique is based on colour segmentation with CIELAB colour-space utilized for justifiable reasons. Using LAB colour-space, lips colours were compiled into a colour-map and processed accordingly to our proposed algorithm of adaptive circular enclosure. Algorithm output was determined as a series of coordinates symbolizing boundary values surrounding colour-map. Separation of colours is based on these boundaries by creating a freeform polygon that defines if colour-value either belongs within colour-boundary polygon or not. This technique is famously known as the point in-polygon technique. Proposed technique evaluation uses XM2VTS database based on false-positive and false-negative to compute segmentation error. Simulation shows proposed algorithm yields segmented error of 5.55% with accuracy of 94.45%.**

Index Terms- Boundary Points, Circular Expansion, Colour Space, Geometrical Distribution Model, Lip Segmentation, Point-In-Polygon.

I. INTRODUCTION

Visual speech recognition researches recently gain popularity after audio based speech recognition research started to saturate, leading researchers to infiltrate into other possible solutions to unsolved problems. Audio-Visual speech recognition concept was introduced based on observation regarding human capability of bimodal interpretation [1]. At normal environment, audio interpretation supersedes visual, however visual interpretation priority increases upon attempts to understand spoken word under high noise environment [2].

In visual speech recognition, various features such as movement of lips, facial motion and discrete body languages can be used to assist in speech interpretation [2]. Out of all these speech features an estimated of 80% usable features for classification originates from lips movement [3]. Due to this fact research on attempting for perfect lips segmentation from mouth region is desirable as a post-processing procedure prior to classification for accurate feature extraction.

Colour-based lips segmentation from skin if often difficult due to reasons that colour values for skin and lips are almost similar especially for lower lip region and when people from different ethnic with different skin tones are involved. The existence of variety of colour-maps such as commonly universal Red-Blue-Green (RGB), Chroma Values (YCbCr) and other derived colour-space from RGB category such as

LAB-Colour-Space reflects on the necessity for researchers to utilized proper colour-space corresponding to intended usage [4].

In this paper, we intend to propose a new colour-space boundary enclosure method inspired by ellipse based geometric boundary and quadratic based colour-space boundary with tighter colour-boundary for better colour-based lip and skin separation utilizing components L* and B* from the CIELAB colour-space.

Proposed technique can be summarised as multiple circular boundaries coordinates derived from centroid of colour-map and extends to colour-map edge followed by point in-polygon algorithms for colour classification and thresholding for segregate between skin and lips pixels.

II. METHODOLOGY

Main methodology required to understand this work can summarized as CIELAB Colour-Space, Colour Distribution Geometrical Model, Circular Coordinates and Point In-Polygon.

A. CIELAB Colour-Space

The CIELAB Colour Space transformation can be given be the following two equations [5].

$$\begin{matrix} X \\ Y = \\ Z \end{matrix} \begin{pmatrix} 0.4125 & 0.3576 & 0.1805 \\ 0.2126 & 0.7152 & 0.0722 \\ 0.0193 & 0.1192 & 0.9505 \end{pmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

Where R, G, and B are red, green and blue colour values, X, Y, and Z are CIEXYZ Colour Space Values which will be used for CIELAB Computation.

$$L* = 116 \times f(Y/Y_n) - 16$$
$$a* = 500 \times [f(X/X_n) - f(Y/Y_n)]$$
$$b* = 200 \times [f(Y/Y_n) - f(Z/Z_n)] \quad (2)$$

$$f(t) = \begin{cases} t^{1/3} & if\ t > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 t + \frac{4}{29} & otherwise \end{cases} \quad (3)$$

Where X_n, Y_n and Z_n are CIEXYZ reference white point values.

B. Colour Distribution Geometrical Model

Examples of colour distribution geometrical model can be seen from the paper [6] proposed skin-segmentation colour space of red and green chromatic colour space. The paper uses two quadratic polynomials to create a boundary that separates skin colours from other colours which is used for classification. Another paper would be [7] which uses an ellipse equation to create a boundary surrounding skin pixels from background and other object colours. Illustration of both methods can be seen in Fig. 1.
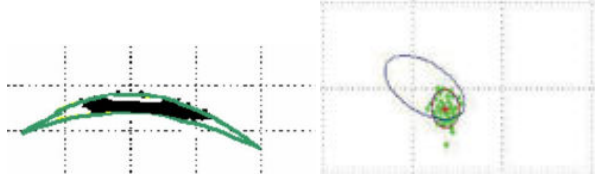

Fig. 1: Image on Left shows two quadratic line drawn surrounding colour-map and the right image show an ellipse boundary drawn surrounding desired colour-map

We believed that segmentation of lips utilizing only colour space values is possible as there have been several works that relates to this. An example would be from paper [8] that uses four linear equation to create a boundary.

For computational efficiency we pursue pure colour separation method for segregating skin pixels with lip pixels using multiple circle boundaries in order to avoid complicated classification computation which could cause system slowdown due to numerous calculations required during assumed real-time application.

C. Circular Coordinates

Pixels or coordinates which defined a circle can be said as individual pixels which surround a single center point while maintaining constant distance from it. Coordinates for circle can be given by equation below [9].

$$(x - a)^2 + (y - b)^2 = r^2 \qquad (4)$$

Where (a, b) = circle origin coordinates, r = circle radius, (x, y) = point coordinates.

D. Point In-Polygon

For classification, a simple method of point in-polygon algorithm is used. Point-in-polygon two main base concepts are angle-based and edge-based. Our research does not emphasize in this aspect but merely utilizes it for classification purposes; therefore paper will utilize the simpler version of point-in-polygon which is edge-based. Basic conditions for point-in-polygon for edge can be given below [10].

Function $f(e)$ = position function of edge as to ray, e=edge of polygon from start point to final point.

$$F = \sum_{i=1}^{n} f(e_i) \qquad (5)$$

$$X_c = X_i + \frac{(Y_0 - Y_i)(X_{i+1} - X_i)}{Y_{i+1} - Y_i} \qquad (6)$$

Where n = number of polygon edge, Edges of polygon = P_i and P_{i+1} (i=1…n, $P_{n+1}=P_1$), Coordinates of P_i and P_{i+1} = (X_i,Y_i) and (X_{i+1},Y_{i+1}), P = Test Point with coordinates (X_0,Y_0).

1) $Y_i>Y_0$ & $Y_{i+1}>Y_0$, F=F
2) $Y_i<Y_0$ & $Y_{i+1}<Y_0$, F=F
3) $Y_i=Y_{i+1}$ (Two Conditions)
   a. $Y_0 \neq Y_i$, F=F
   b. $Y_0 \neq Y_{i+1}$ & $(X_i<X_0<X_{i+0})//(X_i>X_0>X_{i+0})$, Output Result F = 1;
4) $Y_i<Y_0<Y_{i+1}//Y_i>Y_0>Y_{i+1}$ (Two Conditions)
   a. $X_0>X_c$, F = F
   b. $X_0<X_c$, F=F-1
   c. $X_0=X_c$, Output Result F = 1;
5) $Y_i<Y_{i+1}$ & $(Y_0=Y_i // Y_0=Y_{i+1})$, F = 0.5
6) $Y_i>Y_{i+1}$ & $(Y_0=Y_i // Y_0=Y_{i+1})$, F= -0.5

Initial F is set at 0. If F = 1, point is inside the polygon and if F = 0, therefore point is outside of polygon.

III. PROPOSED ALGORITHM

We proposed a new novel colour segmentation method inspired from convention methods of geometrical models and boundary equations and segmentation classification utilizing thresholding, artificial intelligence and various other modes of segregation. The Proposed algorithm will consists of three main parts which are Colour Space Transformation, Boundary Extraction and Pixel Classification.

A. Colour Space Transformation

Decision making had lead us to selecting CIELAB as the research base colour-space based on [4] which states that the colour space provides a decent separation of skin and lips pixels besides the more popular colour transformation and chromes of red and green.
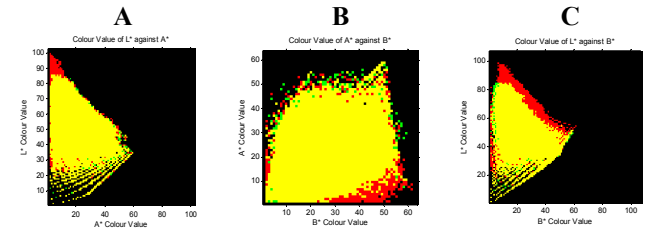

Fig. 2: From Left to Right Shows Colour-Space Plotting using L*-A*, A*-B* and L*-B* from CIELAB Space

Best colour-value data selection is based on segregation level of both skin pixels from lips pixels. Red section refers to skin colour value, yellow represent overlapping of skin and lips pixels and green represents purely lips colour context. Simple tabulation of L* against A*, A* against B* and L* against B* shows that L* V.S. B* provides best separation of lips and skin among the three different combination.

B. Boundary Extraction

We proposed a new method inspired from geometrical colour models of quadratic, linear and ellipse equations for boundary determination. Prior to extracting boundary coordinates, colour map is threshold with respect to density of each colour value in CIELAB. Best threshold value is found to be at 820 followed by some basic morphological techniques of fill and majority.
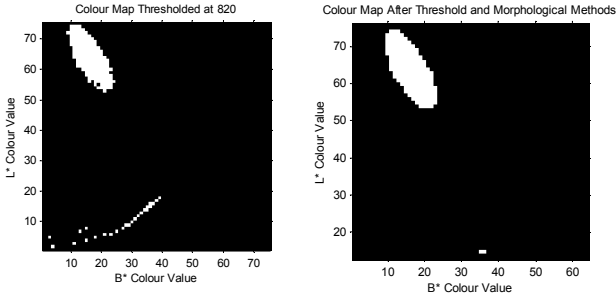
Fig. 3: Left Image Shows Thresholded Colour Map While Right Image Shows Similar Colour Map with Morphological Methods Applied

Pseudo code for adaptive circular boundary determination can be observed below.

1) Let Centroids = Thresholded Colour Space Center X-Y Coordinates and Trigger = 1.
2) While (Trigger = 1)
3) For every Coordinate pairs in Centroids
   a. Let circle radius, R = 1.
   b. Draw circle for radius, R using each pair of coordinate as circle center.
   c. Check drawn circle coordinates for pixel value = 1.
   d. If every pixel value = 1, R=R+1 else R=R-1.
   e. Radius, R values for respective Centroids coordinates are collected into variable Centroid Radius.
   f. Eliminate duplicate coordinates in all circle coordinates and compile into Boundary Coordinates.
   g. If there are any points in Centroid Radius Radius that correspond to R > 1, Trigger remain unchanged.
      i. Remove respective Centroid coordinates from Boundary Coordinates.
      ii. Centroids become Filtered Boundary Coordinates.
   h. If there are no coordinates that correspond to R>1, Trigger = 0.
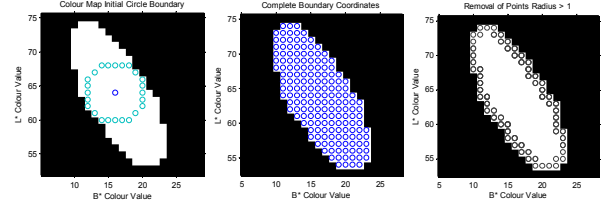4) Output Centroids as Boundary Coordinates.

Fig. 4: From Left to Right, Initial Single Point Circle Enclosure, Complete Circle Coordinates Selection, Filtered Boundary Coordinates based on Individual Radius Value.

Boundary Coordinates obtained at the end of this process is inserted into point in-polygon algorithm to map out colour boundary and determine if pixel is within pre-determined tolerable values.

IV. EXPERIMENTAL RESULTS

Machine specification for experimentation includes a Quad-core i5-460M with 4GB DDR3 RAM and HD5650 Graphics Card. Software used for simulation would be the popular simulation software MATLAB. Database utilized here would be from the popular XM2VTS Database.

Assumptions here include mouth regions have been segmented prior to separation of mouth pixels from skin pixels. Colour Map created based on manually segmented lip images from individual mouth images.
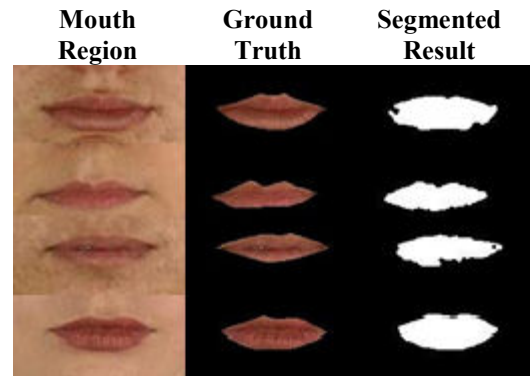
Fig. 5: Shows Raw Data, Ground Truth and Result for four different individuals found in XM2VTS Database

Classified into two different category which is False-Positive error where non-lip pixels are classified incorrectly and the vice versa False-Negative Error. Segmentation Error (SE) is known as the sum of both false positive and false negative pixels divided by total number of pixels. Equation can be deduced as following [11].

$$SE = \frac{F_N + F_P}{TL} \tag{7}$$

Where FN = Number of False Negative, FP = Number of False Positive, TL = Total Number of Pixels.

Using this evaluation method test was conducted on 144 XM2VTS segmented mouth region frames and results are tabulated in TABLE 1.

TABLE 1: RESULTS OBTAINED FROM VARIOUS TECHNIQUES FOR LIP SEGMENTATION PURPOSES

Method	Segmentation Error	Detection Rate
Proposed Method	5.55	94.45
Local Kernel Colour Histogram & Colour Information Fusion [11]	5.67	94.33
Local Kernel & Colour Histogram [12]	6.28	93.72
Gaussian Mixture Model [13]	6.54	93.46
K-Means and Fusion [12]	6.55	93.45
ISODATA Clustering [13]	11.84	88.16

Based on TABLE 1, our proposed method yields slightly better results than the more complex and higher computational methods of utilizing information fusion techniques as well as Gaussian and clustering methods.

V. CONCLUSIONS

Adaptive Circular Enclosure method compiles a list of coordinates which represents boundary surrounding colour map of lips also known as the geometrical shape of lips colour distribution model. Coordinates are placed into point-in-polygon algorithm to draw out a shape from coordinate list which is used as a replacement for quadratic/linear condition equation for separating lip pixels from skin pixels. Colour map was determined to be CIELAB colour space and from combination of three different subspace, L* and B* subspace seems to provide better colour separation between lips and skin. Experimental results yields segmentation accuracy of 94.45% which is the highest among listed of methods.

REFERENCES

[1] G, Stork D and M.E. Hennecke, *Speechreading by Humans and Machines*. Berlin, Germany: Springer, 1996.

[2] Lars A. Ross, Dave Saint-Amour, Victoria M. Leavitt, Daniel C. Javitt, and John J. Foxe, "Do You See What I Am Saying? Exploring Visual Enhancement of Speech Comprehension in Noisy Environment," *Oxford Journals Life Sciences & Medicine Cerebral Cortex*, vol. 17, no. 5, pp. 1147-1153, 2006.

[3] Katnak Qi Zhi, "Analysis of Lip Geometric Features for Audio-Visual Speech Recognition Systems," *IEEE Transactions on Man and Cybernetics*, vol. 34, no. 4, pp. 564-570, 2004.

[4] Alan Wee-Chung Liew and Shilin Wang, *Visual Speech Recognition : Lip Segmentation and Mapping*. Hershey, New York: Medical Information Science Reference, 2009.

[5] Yao WenJuan, Liang YaLing, and Du MingHui, "A Real-Time Lip Localization and Tracking for Lip Reading," in *IEEE 3rd International Conference on Advanced Computer Theory and Engineering*, 2010, pp. 363-366.

[6] Cheng-Chin Chiang, Wen-Kai Tai, Mau-Tsuen Yang, Yi-Ting Huang, and Chi-Juang Huang, "A Novel Method for Detecting Lips, Eyes and Face in Real Time," in *International Conference on Real-Time Imagging*, 2003, pp. 277-287.

[7] Wenjun Tan, Gaoyang Dai, Han Su, and Ziyi Feng, "Gesture Segmentation based on YCbCr Colour Space Ellipse Fitting Skin Color Modelling," in *24th Chinese Control and Decision Conference*, 2012, pp. 1905-1908.

[8] M. Shemshaki and R. Amjadifard, "Lip Segmentation Using Geometrical Model of Color Distribution," in *7th Iranian Machine Vision and Image Processing*, 2011, pp. 1-5.

[9] Rod Pierce and et.al. (2011) Maths Is Fun ? [Online]. http://www.mathsisfun.com/algebra/circle-equations.html

[10] Wu Jian and Cai Zongyan, "A Method for the Decision of a Point Whether In or Not in Polygon and Self-Intersected Polygon," in *IEEE 8th International Conference on Fuzzy Systems and Knowledge Discovery*, 2011, pp. 16-18.

[11] F. Salimi and Mohammad T Sadeghi, "Decision Level Fusion of Colour Histogram Based Classifiers for Clustering of Mouth Area Images," in *International Conference on Digital Image Processing*, 2009, pp. 416-420.

[12] F. Salimi and M. Sadeghi, "Segmentation of Lip Pixels Using Local Kernel Histogram in Colour Images," in *Proceedings in 16th Iranian Conference on Electrical Engineering*, 2008.

[13] M. Sadeghi, H. Kittler, and K. Messer, "Modelling and Segmentation of Lip Area in Face Images," *IEEE Proceedings in Conference Video, Imaging and Signal Processing*, vol. 149, no. 3, pp. 179-184, 2002.
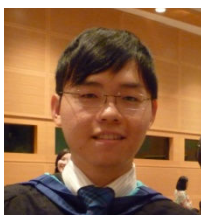
Hao Wei Lee received the Masters in Engineering from the Faculty of Engineering, The University of Nottingham in 2011. His research interest includes the fields of intelligent visual, audio processing, robotics and automation.

Wai Chong Chia received the B.Eng (1st class) and MSc (Distinction) from University of Nottingham Trent and University of Nottingham in 2006 and 2008 respectively. He is currently working as a lecturer at Sunway University Malaysia. His research interests include the fields of visual processing, embedded system, and wireless visual sensor networks.

Kah Phooi Seng received the B.Eng (1st class) and PhD from University of Tasmania, Australia, in 1997 and 2001 respectively. She is currently a Professor at Sunway University Malaysia. Prior to joining Sunway University, she was Associate Professor at School of Electrical and Electronic Engineering, Nottingham University. Her research interests include the fields of visual processing, multi-biometrics, artificial intelligence and wireless visual sensor networks.

Li-Minn Ang received the B.Eng (1st class) and PhD from Edith Cowan University, Australia, in1996 and 2001, respectively. He is currently a research staff at School of Engineering, Edith Cowan University. Prior to joining Edith Cowan University, he was Associate Professor at School of Electrical and Electronic Engineering, Nottingham University. His research interests include the fields of video compression, visual processing, wireless visual sensor networks and reconfigurable computing.