

# 색도 변환과 퍼지 클러스터링을 이용한 입술영역 추출

김 정 엽<sup>†</sup>

## Extraction of Lip Region using Chromaticity Transformation and Fuzzy Clustering

J.Y. Kim<sup>†</sup>

### ABSTRACT

The extraction of lip region is essential to Lip Reading, which is a field of image processing to get some meaningful information by the analysis of lip movement from human face image. Many conventional methods to extract lip region are proposed. One is getting the position of lip by using geometric face structure. The other discriminates lip and skin regions by using color information only. The former is more complex than the latter, however it can analyze black and white image also. The latter is very simple compared to the former, however it is very difficult to discriminate lip and skin regions because of close similarity between these two regions. And also, the accuracy is relatively low compared to the former. Conventional analysis of color coordinate systems are mostly based on specific extraction scheme for lip regions rather than coordinate system itself. In this paper, the method for selection of effective color coordinate system and chromaticity transformation to discriminate these two lip and skin region are proposed.

**Key words:** Lip/Skin Color Analysis, Lip Region Extraction

### 1. 서 론

입의 영상으로부터 추출된 입술영역의 정보를 활용하는 대표적인 분야는 음성-시각 자동 언어인식(Audio-visual automatic speech recognition)이다 [1]. 해당분야에서는 언어인식에 있어 오디오 데이터 외에도 부가적으로 입술모양을 활용한 비디오 데이터를 결합하여 복합적인 해석을 시도하여 언어 인식을 높이고 있다. 이와 같은 시각적인 비디오 데이터를 자동으로 추출하기 위해서는 화자의 입술영역을 추출하는 것이 필수적이다. 과거 흑백영상의 경우, 입술영역은 1차적으로 선정된 입 주변 영역을 대상으로 템플릿[2], 능동모형[3] 등의 방법으로 추출

하였다. 템플릿의 경우, 에너지 함수를 최소화 시키는, 목표의 입술모양에 맞는 최적의 파라미터로 구성된 곡선을 구성하는 방법을 적용한다. 능동모형은 입술영역에 대한 확률모델을 기반으로, 미리 정해진 모델의 투영과 실제 영역의 차이를 이용하여 모델 파라미터를 갱신하는 방법이다. 칼라영상의 경우는 입술에 특정 색을 칠하거나[4], 적절한 환경에서 미리 촬영된 입술영상 데이터베이스를 이용하는[5] 방법 등이 있다. 본 연구에서는 입술영역을 추출하기 위한 기법보다 사용된 칼라좌표계 또는 칼라변환에 초점을 두어 분석하고자 한다. Michael Vogt[6]는 HSI 좌표계 중 단지, 색 정보만 사용하기 위하여 HS와 LUT(Look Up Table)를 결합하고 피부색과 입술 색

※ Corresponding Author : Jeong Yeop Kim, Address: (612-743) 99 Pilbong-gil Haeundae-gu, Busan, Korea, TEL : +82-51-540-7105, FAX : +82-51-540-7280, E-mail : neocopy@ysu.ac.kr

Receipt date : Apr. 21, 2014, Revision date : Jun. 3, 2014  
Approval date : Jun. 12, 2014

<sup>†</sup> School of Undeclared Majors, Youngsan University

※ This research was supported by the Research Grant funded by Youngsan University.

경계를 추출하여 입술영역을 구하였다. M. U. Ramos Sanchez 등[7]은 RGB 좌표계를 이용하여 미리 정해진 피부색 영역을 중심으로 기하학적인 위치정보를 이용한 입술영역 추출을 시도하였다. Demas Sanger 등[8]은 칼라변환인 Lrg 항목과 미리 정해진 피부색과 입술색의 분포정보를 이용하여 입술영역을 추출한다. 두 영역의 분포가 일정부분 겹치는 것을 볼 수 있으며, 이를 극복하기 위하여 기하학적인 위치정보를 활용하였다. T. Wark[9], Trent[10], Vladimir[11] 등은 (R/G, G/B) 형태의 RGB 좌표계의 각 요소 비(ratio)를 이용한 칼라변환을 중점적으로 활용하고, 이 값들이 미리 정해진 문턱치를 만족하면 입술영역으로 판별한다. M. Wimmer 등[12]은 RGB 좌표계의 세 요소를 각각 정규화한 형태인  $R/(R+G+B)$ 의 값들에 대하여 문턱치를 적용하여 입술영역을 구한다. 미리 정해진 문턱치는 실험을 통한 데이터베이스에 의존하며 성능을 결정한다.

Liew[13], Yi Wu[14]는 CIELab, CIELUV 등의 좌표계를 이용하였다. Liew 등은 두 개의 좌표계를 결합하여 벡터를 만들거나 색조(hue)를 계산한 후, 일정한 문턱치를 적용하는 방식을 사용하였고, Yi 등은 CIELUV 좌표계 중 U 성분에 대하여 일정한 문턱치를 적용하였다. Simon Lucey 등[15]은 RGB 좌표계를 활용하여 (R/G, 평균, 분산), rg 등의 칼라변환을 위주로 분석하였다. Erol 등[16]은 RGB,  $R/(R+G)$ , 색조(hue),  $(r,g)=(R/(R+G+B), G/(R+G+B))$  등의 4가지 칼라좌표계와 칼라변환을 복합적으로 사용하여 각각의 요소가 개별적으로 적용되는 것에 비하여 복합적인 구성이 더욱 효과적임을 보였다. Rohani 등[17]은  $R/(R+G)$ 의 가상색상을 이용하여 영역분할을 시도하였다. Wenchao 등[18]은 CIELUV 좌표계와 DHT (Discrete Hartley Transform)를 이용하여 입술영역을 추출하였다. 그 외에 후백정보만으로 입술영역을 추출하는 경우[19]도 있다.

이상의 기존 알고리즘에서 칼라좌표계들을 활용하는 경우, 기본 좌표계와 파생되는 칼라변환을 복합적인 형태로 구성하여 적용하는 것이 대부분이다. 입술영역과 피부영역의 색차를 최소화하는 좌표계에 대한 분석과 데이터 해상도(resolution)의 고찰보다는 최대한 여러 가지 형태의 단편적인 변환들을 결합하는 구조로 칼라정보를 활용하고 있다. 본 논문에서는 단편적인 결합의 구조보다는 입술과 피부 두 개의

칼라 영역을 구분하는데 가장 유리한 좌표계에 대한 고찰을 하고자 한다. 2장에서는 입술영역 추출에 사용되는 기존 좌표계의 활용 형태에 대해 기술하고, 3장에서는 기존의 좌표계를 중심으로 피부영역과 입술영역의 색차(color difference)를 계산하고, 기존 좌표계에 대하여 본 논문에서 제안하는 데이터 해상도를 기준으로 성능분석을 한다. 얼굴 데이터베이스에서 추출한 영상에서 피부색과 입술 색 영역의 2개 칼라 데이터 그룹에 대하여 데이터 해상도를 각각 계산하여 변별력이 우수한 좌표계를 선정할 수 있다. 4장에서는 효율적인 특징의 계산, 색도변환 방법의 제안 및 영역분할방법을 설명하고, 5장에서 실험결과를 보인다.

## 2. 입술영역 추출을 위한 칼라좌표계

입술영역 추출에 사용되는 칼라좌표계 또는 칼라변환은 주로, RGB 좌표계 및 이를 응용한 변환, IHS 좌표계 및 관련 변환, CIELab/LUV 좌표계 및 관련 변환의 세 가지 항목이 주로 이용된다. 부가적인 변환 등이 없는 간단한 데이터를 제공하므로 RGB 좌표계가 가장 많이 이용되며 주로 밝기 값(intensity value)을 기준으로 정규화한 r, g 또는 R/G와 같은 각 요소의 비(ratio)를 기준으로 한다. 특히 색상(hue)과 색 포화도(saturation) 정보를 이용하기 위하여 IHS 좌표계의 H와 S 요소를 이용하기도 한다. 사람이 판단하는 색상의 차(color difference)에 가장 근접한 것으로 평가되고 있는 국제조명위원회(CIE) 계열의 좌표계도 이용되고 있다. CIELab, CIELUV 등의 좌표계이며, 색차를 계산할 때 등간격인 점을 활용한다. 본 장에서는 주요한 칼라 좌표계의 활용 형태를 고찰한다.

M. U. Ramos Sanchez 등[7]은 RGB 좌표계를 이용하여 미리 정해진 피부색 영역을 중심으로 기하학적인 위치정보를 이용한 입술영역 추출을 하였다. 피부색 영역은 조명의 영향을 배제하기 위하여 정규화된 RGB 값을 이용한 일반적인 영역분할의 결과를 이용한다.

Erol 등[16]은 RGB,  $R/(R+G)$ , 색조(hue),  $(r,g)=(R/(R+G+B), G/(R+G+B))$  등의 4가지 칼라좌표계와 칼라변환을 복합적으로 사용하였다. 식(1),(2)와 기본적인 RGB 좌표계, 색조(hue)를 다양하게 조합하

여 실험하였으며 성능은 식(1), (2) 및 색조 세 가지를 결합한 경우가 입술면적 대비 85%로 가장 좋은 것으로 나타났다.

$$\frac{R}{R+G} \tag{1}$$

$$\left(\frac{R}{R+G+B}, \frac{G}{R+G+B}\right) \tag{2}$$

Simon Lucey[15]등은 Sanchez와 유사하게 식(2)~(5)의 칼라변환을 위주로 분석하였다.

$$\frac{R}{G} \tag{3}$$

$$SD(i,j) = \ln \sqrt{\frac{\sum_{p=i-3}^{i+3} \sum_{q=j-3}^{j+3} \left[\frac{R}{G}(i,j) - AI(p,q)\right]^2}{49}} \tag{4}$$

$$AI(i,j) = \frac{\sum_{p=i-3}^{i+3} \sum_{q=j-3}^{j+3} \frac{R}{G}(p,q)}{49} \tag{5}$$

여기서 (i,j)는 영상에서 i행 j열을 나타낸다. Lucey는 식(2), 식(3), 식(3)~(5)의 세가지 유형에 대하여 입술영역 추출에 대한 실험을 하였으며, 식(2)만 사용한 경우 87% 내외, 식(3)만 사용한 경우 89% 내외, 식(3)~(5)를 사용한 경우 90% 정도의 인식률을 달성하였다. 주로 식(3)에 의존적이며 하나의 영상에 대한 식(3) 계산 값의 7x7창 평균과 분산을 고려하였다.

T. Wark[9], Trent[10], Vladimir[11] 등은 (R/G, G/B) 형태의 RGB 좌표계의 각 요소 비(ratio)를 이용한 칼라변환을 중점적으로 활용하였으며, M. Wimmer등[12]은 RGB 좌표계의 세 요소를 각각 정규화한 형태인 R/(R+G+B) 형태의 값들에 대하여 문턱치를 적용하였다.

Demas Sanger등[8]은 칼라변환인 Lrg 항목과 미리 정해진 피부색과 입술색의 분포정보를 이용하여 입술영역을 추출하였다. 상기, 칼라변환의 경우[8, 10,12]들은 사전 정보의 획득 및 문턱치의 결정에 어려움이 있다.

Michael Vogt[6]는 HSI 좌표계 중 단지, 색 정보만 사용하기 위하여 HS와 LUT(Look Up Table)를 결합하고 피부색과 입술 색 경계를 추출하였다.

Liew[13], Yi Wu[14]는 CIELab, CIELUV등의 좌표계를 이용하였다. Liew등은 두 개의 좌표계를 결합하여 벡터를 만들거나 색조(hue)를 계산한 후, 일

정한 문턱치를 적용하는 방식을 사용하였고, Yi등은 CIELUV 좌표계 중 U성분에 대하여 일정한 문턱치를 적용하였다. 이러한 접근 방법에서도 적절한 문턱치의 결정은 어려운 문제이다.

본 논문에서는 이와 같이 여러 가지 좌표계의 원형 또는 칼라변환을 이용할 때, 입술 색 영역과 피부 색 영역의 분포를 분석하여 각 영역들의 특징을 추출하고, 두 개의 영역이 최대한 분리되어 있는 좌표계 또는 칼라변환을 제시하고자 한다. 두 개의 영역에 대한 분리도가 가장 우수한 좌표계와 관련 파라미터를 활용하면 입술영역의 추출이 유리해 진다.

### 3. 제안한 칼라좌표계 분석 방법

본 논문에서는 칼라좌표계 또는 칼라변환 중에서 사용 빈도가 높은 RGB, XYZ, rgY, HSV, YCrCb, R/G, CIELUV[20]의 7개를 대상으로 분석을 하고자 하며, 기준영상에서 추출된 입술 색과 피부색의 2개 그룹 데이터 집합에 대하여 각 그룹간의 거리를 중심으로 분석한다.

#### 3.1 입술과 피부 영역에 대한 평균과 표준편차

본 논문에서 사용한 얼굴영상은 MIT-CBCL-facerec-database[21]에서 추출한 5장의 영상을 이용하였고, 일정한 영역을 중심으로 피부영역과 입술영역에서 각각 40개씩의 샘플 화소를 추출하였다. 본 절에서는 5장의 영상에서 추출된 샘플화소를 기준으로, 각 좌표계로 변환한 다음 입술과 피부영역에 대한 평균벡터와 표준편차를 식(6)과 같이 계산하였다. 평균벡터를 중심으로 계산된 표준편차는 각 클래스터의 밀집된 정도를 표현한다.

$$C_{avg} = \frac{1}{N} \sum_{i=1}^N C_i \tag{6}$$

$$\sigma = \frac{1}{N} \sum_{i=1}^N (C_i - C_{avg})^2$$

위 식에서 C<sub>i</sub>는 임의의 샘플 화소에 대한 각 좌표계 별 칼라 벡터이고, C<sub>avg</sub>는 계산된 평균벡터이다.

#### 3.2 칼라좌표계의 성능 분석을 위한 새로운 기준

##### 3.2.1 입술과 피부색 영역의 클래스 분류 해석

두 개 칼라 영역의 분류는 전통적인 2 클래스 문제

Table 1. The matching rate for color coordinate systems

Color Coordinate Systems	Matching rates	
	Lip_hit [%]	Skin_hit [%]
RGB	85.0	90.3
XYZ	80.9	86.6
rgY	88.4	93.4
HSV	80.3	88.1
YCrCb	94.7	96.3
R/G	94.4	100.0
LUV	96.9	99.7

(2 class problem)로 해석되며, 해석 도구는 근접 이웃(Nearest neighbor) 분류기[22]를 사용하였다. 입술과 피부의 2개 영역 데이터 집합 80개의 모든 데이터를 3.1절에서 구한 두 개의 평균벡터를 이용하여 가까운 거리에 해당하는 클래스로 분류하였다. 표 1에 실행 결과를 나타내었으며, Lip\_hit은 원래의 입술 데이터 집합 소속의 데이터가 올바르게 입술 집합으로 분류된 백분율을 나타낸다. 예를 들면, 40개 중에서 36개 데이터가 올바르게 분류되면 90%가 되는 것이다. 표 1에 나타난 결과는 5개의 테스트 영상에 대하여 7개의 좌표계별로 각각의 백분율을 계산한 후, 평균값을 취하여 전체적인 성능을 평가하였다.

Lip\_hit을 기준으로 분석한 결과, CIELUV 좌표계는 가장 우수한 결과를 보였고, HSV좌표계는 분류 성능이 가장 낮았으며, 오히려 변환을 하지 않은 RGB좌표계보다 성능이 낮은 것이 특징이다. 또한, Skin\_hit을 기준으로 하면, YCrCb는 96.3%와 R/G는 100%로 나타났으며, CIELUV는 99.7%로 나타나 R/G와 같이 가장 우수함을 알 수 있다. 그러므로 종합적인 평가를 하면, CIELUV 좌표계가 분류성능이 가장 우수하게 나타났다.

### 3.2.2 정규화 거리(Normalized distance)

본 논문의 3.1절에서는 평균벡터만을 이용하여 기존의 칼라 좌표계 상에서의 직접적인 거리를 기준으로 각 좌표계의 변별력(discriminability)을 비교하였다. 그러나 각 좌표계의 척도(scale)가 모두 다르기 때문에 직접적인 변별력의 비교가 어렵다. 그림 1에서  $C_S$ 는 피부색을 가지는 화소들의 칼라 좌표 집합이며,  $C_L$ 은 입술 색 영역 화소들의 칼라 좌표 집합을

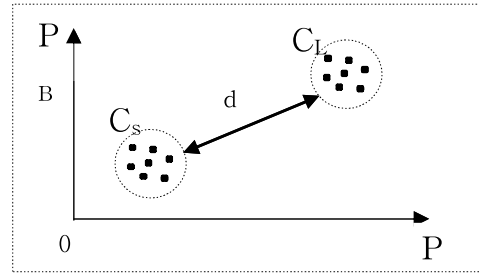


Fig. 1. Color distribution and difference between lips and skin region.

나타낸다. 3.1절에서는 단순하게  $C_S$ 와  $C_L$ 의 직접적인 거리인  $d$ 만을 고려하였다. 여기서  $P_A-P_B$ 는 HSV 좌표계의 (H,S) 칼라평면과 같은 변환 가능한 임의의 칼라 좌표계 요소 축을 의미한다.

여러 가지 칼라 좌표계들의 두 개 칼라좌표 집합의 거리에 따른 변별력을 비교하기 위하여 각 좌표계 별로, 발생 가능한 최대의 색차인  $D_{Max}$ 를 기준으로 하여 식(7)과 같이  $D_N$ 을 계산할 수 있다. 계산값  $D_N$ 은 각 좌표계에서 정규화된 해상도를 기준으로 하는 두 개 칼라좌표 집합의 상대적인 거리를 최대거리 대비 백분율[%]로 나타내고, 이는 두 개 칼라좌표 집합의 변별력으로 간주한다.

$$D_N = \frac{d}{d_{Max}} \tag{7}$$

### 3.2.3 칼라좌표 집합의 분리도(Separability for the set of color coordinates)

비록 각 좌표계 별로,  $D_N$ 의 값이 상당히 크더라도 각 클러스터 분포상의 모든 좌표 값들이 넓은 범위에 퍼져 있으면, 즉 표준편차가 크면 두 개 칼라좌표 집합의 변별력이 크게 저하되며, 경우에 따라서는 두 개 칼라 좌표 분포가 상당부분 겹쳐질 수도 있다. 이러한 경우는 두 개 칼라좌표집합의 구별이 어려워진다. 그러므로 각 집합의 분리도를 계산하면 변별력을 결정하는 하나의 척도로 사용할 수 있다. 분리도  $S$ 는 입술영역에 대한 데이터 분산과 피부영역 데이터 분산을 이용하여 식(8)과 같이 계산할 수 있다. 클러스터의 밀집도가 높고 클러스터간 거리가 멀면 분리도가 높게 나타난다.

$$S = \frac{D_N}{\sqrt{\frac{\sum_{i=1}^N (V_{Li} - V_{LA})^2}{N}} \times \sqrt{\frac{\sum_{i=1}^N (V_{Si} - V_{SA})^2}{N}}} \tag{8}$$

여기서  $V_{Li}$ 는 해당 칼라 좌표계의 입술영역 좌표 벡터를,  $V_{LA}$ 는 각 좌표집합의 입술영역 평균벡터를 의미한다.  $N$ 은 칼라집합의 샘플 수이며 본 논문에서는 40으로 하였다.

### 3.3 효율적인 칼라좌표계의 선정방법

본 논문에서 사용한 얼굴영상은 MIT의 CBCL-facerec 데이터베이스에서 추출한 5장의 영상을 이용하였고, 일정한 영역을 중심으로 피부영역과 입술영역에서 각각 40개씩의 샘플 화소를 수동으로 추출하였다. 제안한 기준을 이용한 칼라 좌표계의 성능분석을 위한 각 테스트 영상 IM1~IM5에 대한 정규화 거리  $D_N$ 은 표 2와 같이 비교된다.

rgY좌표계가 가장 정규화 거리가 크게 나타나 2개 클러스터간 거리가 상대적으로 우수하다. 반면, R/G의 경우는  $d_{Max}$ 가 매우 크므로 식(13)에 의하여 정규화 거리가 거의 0에 접근한다. 이는 잡음 등의 요인을 고려할 때 매우 비효율적인 좌표계임을 의미한다. 분리도  $S$ 는 표 3과 같이 표현된다.

표 2에서 정규화거리가 크게 나타날수록 입술영역과 피부색영역의 분리도가 좋음을 나타낸다. rgY 좌표계가 가장 정규화 거리가 크게 나타나 클러스터간 거리가 상대적으로 우수하나 식(8)에 의한 분리도는 낮은 편으로 나타났다. 이는 피부색과 입술색의 두 개 클러스터 분포밀도가 낮아서 결집도가 낮음을

의미하고, 분류오류가 증대됨을 나타낸다. 3.2절의 성능평가에서 CIELUV좌표계가 최상의 성능을 나타냈으며, 본 장에서의 결과도 분리도가 가장 우수하였다. 참고로, 여섯 번째인 R/G 좌표계의 경우는 실험결과 분리도는 CIELUV 좌표계보다 매우 우수하게 나타났지만 계산 불가의 항목이 존재한다. 즉,  $G$ 의 값이 0이 되는 경우 파라미터 계산이 불가하여 실제에 사용하기가 적절하지 않은 단점이 있다. 계산상의 안정성을 고려하여 CIELUV좌표계가 가장 우수하다고 판단하였다.

## 4. 입술영역 추출에 유용한 특징의 선정과 영역 분할 방법

### 4.1 입술과 피부영역의 칼라 특성 및 색도변환

실험에 사용된 영상의 집합은 캘리포니아공대(California Institute of Technology)의 얼굴 데이터베이스[23]를 사용하였다. 데이터베이스에는 17명의 남자와 11명의 여자에 대한 450장의 896x592 크기의 영상을 포함하며, 개인 당 20여 가지의 다양한 환경에서 촬영한 것이다. 실험을 위하여 개인당 2장씩 총 54장의 영상을 선택하였고 그림 2에 간략하게 나타내었다.

입술과 피부영역의 특성을 파악하기 위하여 각 실험영상 마다 입술과 피부 각각 60개씩의 샘플화소를

Table 2. The matching rate for color coordinate systems

Test Images	Color coordinate systems						
	RGB	XYZ	rgY	HSV	YCrCb	R/G	LUV
IM1	10.104	8.912	11.460	9.556	6.698	0.059	3.043
IM2	9.104	8.300	10.151	8.680	5.775	0.071	3.082
IM3	10.261	9.793	10.772	9.709	6.176	0.066	3.369
IM4	8.690	7.189	10.018	7.748	6.346	0.075	3.473
IM5	9.831	9.504	10.006	9.194	5.878	0.077	3.638

Table 3. The separability for color coordinate systems

Test Images	Color coordinate systems						
	RGB	XYZ	rgY	HSV	YCrCb	R/G	LUV
IM1	5652	5482	7312	5961	12062	0	26882
IM2	4196	4099	5194	4227	9624	0	35699
IM3	4944	4590	6027	5109	10013	0	22758
IM4	5028	4483	6331	4813	12452	0	41020
IM5	11161	10557	15009	12017	21740	0	42924

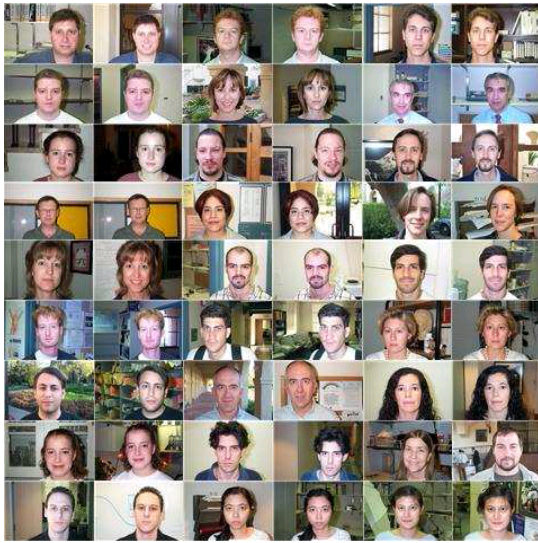


Fig. 2. Color face images for test.

추출하여 칼라 특징 값을 분석하였다. 여러 가지 방법으로 특징을 분석한 결과, 가장 구분이 잘 될 수 있는 데이터는 UV평면으로 나타났으며 그림 3과 4에 일부 나타내었다.

54장의 입력영상에 대한 각각의 UV값의 평균값

을 모아서 나타낸 것이 그림 5이며, U와 V의 차이 값은 입술과 피부영역을 구별할 수 있는 좋은 특징 값으로 분석되었다.

그림 5에 나타난 평균값을 기준으로 하면 |U-V|의 값은 안정적으로 나타나지만 실제에 있어서 해당영상 화소의 칼라 값들은 일정한 데이터 분포를 가지므로 그림 5(b)와 같이 명확하게 구분되지는 않은 경향이 있다. 본 논문에서는 U와 V의 계산과정을 수정하여 입술과 피부 영역에 대한 색차가 커지도록 하는 색도 변환(chromaticity transformation)을 제안한다. 기존의 CIELUV 좌표계에서의 U와 V는 식(9)와 같이 표현되며 U-V 색도평면 데이터는 그림6과 같이 나타난다. 여기서 ( $u_{pn}, v_{pn}$ )는 백색점(white point)을 의미한다.

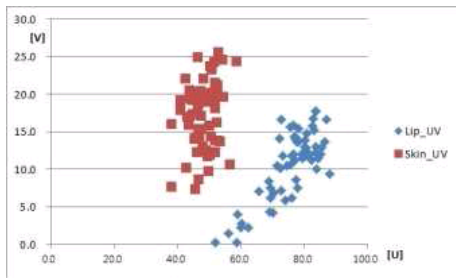
$$L = 116 \left( \frac{Y}{Y_W} \right)^{\frac{1}{3}} - 16 \tag{9}$$

$$u = 13L(u_p - u_{pn})$$

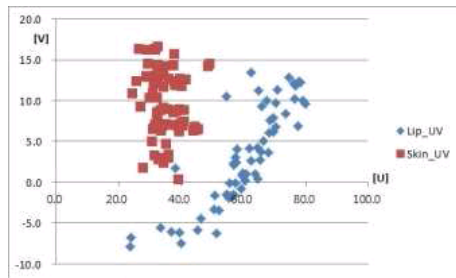
$$v = 13L(v_p - v_{pn})$$

여기서,  $u_p = \frac{4X}{X+15Y+3Z}, v_p = \frac{9Y}{X+15Y+3Z}$  이고,

$u_{pn} = \frac{4X_W}{X_W+15Y_W+3Z_W}, v_{pn} = \frac{9Y_W}{X_W+15Y_W+3Z_W}$  이다.

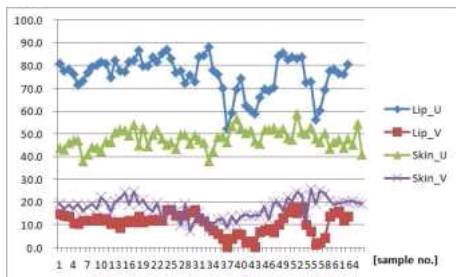


(a)

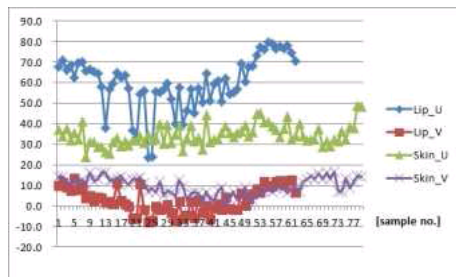


(b)

Fig. 3. Distribution of lips and skin samples in UV plane. (a) image #1 (b) image #2.



(a)



(b)

Fig. 4. U-V graph for the lips and skin samples. (a) image #1 (b) image #2.

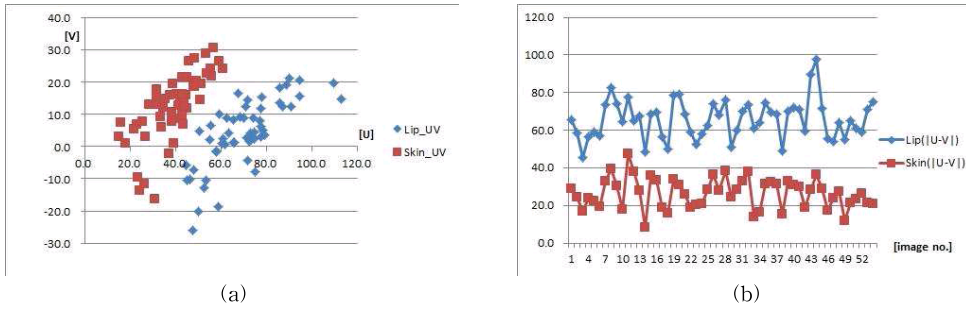


Fig. 5. U-V distribution and graph for the lips and skin samples, (a) UV distribution (b) |U-V| values.

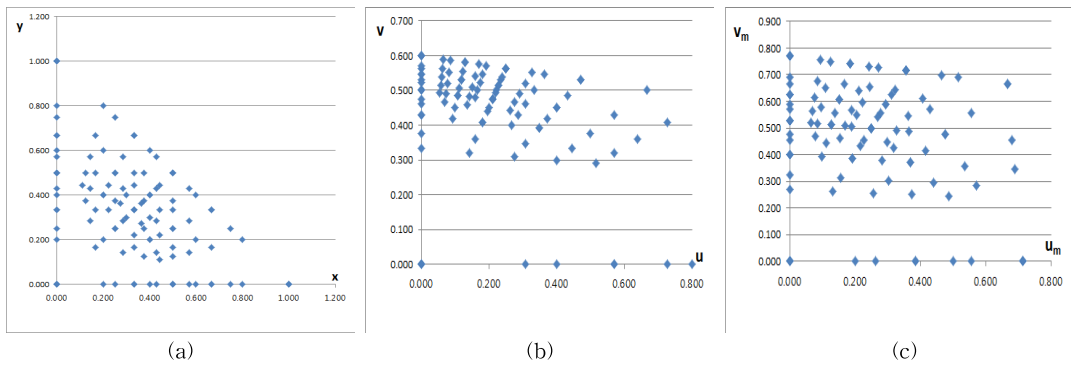


Fig. 6. Data distribution of CIE x-y and U-V planes. (a) x-y distribution (b) U-V distribution (c) Um-Vm distribution.

그림 6의 (b)에서와 같이 실질적인 U와 V값은 x-y 평면에 비하면 선형화되었으나, 입술과 피부에 대한 색도좌표는 매우 가깝게 나타나므로 U-V 분포를 더욱 넓게 표현하면 두 영역에 대한 분리 성능이 개선될 것이다. 이와 같은 가정 하에 상기 식(9)를 변형하여 식(10)을 제안하였다. 식(10)의  $U_m$ - $V_m$  색도좌표 분포는 그림 6의 (c)와 같이 개선됨을 볼 수 있고, 변환된 색도를 입술과 피부영역에 대한 영역분할에 적용한다.

$$L = 116 \left( \frac{Y}{Y_W} \right)^{\frac{1}{3}} - 16 \tag{10}$$

$$u_m = 13L(u_{pm} - u_{pmn})$$

$$v_m = 13L(v_{pm} - v_{pmn})$$

여기서,  $u_{pm} = \frac{5X}{X+13Y+6Z}$ ,  $v_{pm} = \frac{10Y}{X+13Y+6Z}$  이고,

$$u_{pmn} = \frac{5X_W}{X_W+13Y_W+6Z_W}, v_{pmn} = \frac{10Y_W}{X_W+13Y_W+6Z_W}$$

이다.

#### 4.2 입술-피부 영역분할 방법

기존의 입술-피부 영역분할 방법은 Liew[13] 등이 제안한 공간 퍼지 c-means 알고리즘과 Rohani [17] 등이 동일한 영역분할 방법에 새로운 특징 값을 제안하는 것이 대표적이다.

Liew[13] 등은 CIELab, CIELUV와 이 두 칼라 좌표계를 각각 그리고 조합하여 3차원 또는 5차원의 특징벡터를 구성하여 사용하였다. 또한 색상 정보를 계산하여 문턱치를 적용하였고, 칼라 영상으로부터 입술영역을 추출하기 위하여 공간 퍼지 c-means 알고리즘을 사용하였다. 알고리즘의 구성은 그림 7과 같다.

입력된 영상에 대한 밝기 값의 보정을 실시한 다음, 공간 퍼지 c-means(SFCM) 방법[13]을 사용하여 입술영역에 대한 멤버십을 계산한다. 입술 후보영역에 대한 모폴로지 필터를 적용하여 추출될 영역을 부드럽게 만들고, 입술의 기하학적인 구조에 대한 대칭처리를 한다. 밝기 값을 조정하고 입술모양에 대한 템플릿을 이용하여 최종 보정한다. 본 논문에서는 전체의 과정 중에서 입술 영역에 대한 멤버십 계산 부

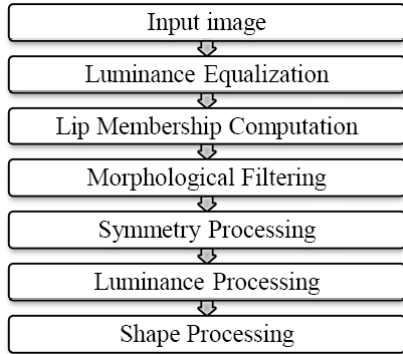


Fig. 7. Flowchart of Liew et. al.'s algorithm.

분만을 비교대상으로 삼아 특징치의 우수성을 검증하고자 한다.

Rohani[17] 등은 가상의 색상이라는 개념의 특징치를 도입하여 상기 멤버십 계산을 동일한 방법으로 수행하였다. 가상의 색상은  $R/(R+G)$ 의 형태로 화소의 빨강, 녹색의 비율을 이용하여 계산이 가능하며, 입술의 경우 피부색 보다 약간 더 붉게 보인다는 특성을 활용한 것이며 처리방법은 그림 8과 같다.

입력영상에 대하여 얼굴부분을 검출하고, 아래 쪽 1/3의 영역을 입술 후보영역으로 지정한다. 입술 후보영역에서 크와 같은 얼굴의 다른 부분이 배제 될수록 검출성능이 향상된다. 가상의 색상을 미리 계산한 다음 공간 퍼지 c-means 알고리즘을 이용하여 입술 영역에 대한 멤버십을 계산한다. 이후, 가우시안 마스크를 이용하여 영역을 매끄럽게 수정하고, 입술모양의 템플릿을 이용하여 모양의 보정을 최종적으로 한다. Liew 등의 방법보다 특징치가 간단하고 입술 후보영역의 선정이 적절하면 약간 더 좋은 성능을 보인다. 본 논문에서는 위의 경우와 마찬가지로 입술 후보영역에 대한 멤버십 계산부분을 제안한 특징치를 이용한 경우와 중점적으로 비교하였다.

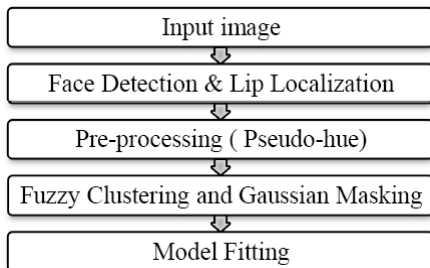


Fig. 8. Flowchart of Rohani's algorithm.

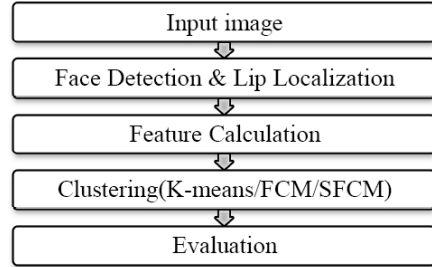


Fig. 9. Flowchart of Proposed extraction algorithm.

제안한 특징치를 이용한 입술영역 검출방법의 처리과정은 그림 9에 나타내었다.

제안한 방법을 샘플영상을 기준으로 표현하면 그림 10과 같다.

위 그림을 기준으로,

단계1: 입력영상  $In(i,j)$ 로부터 Nilsson[24]의 방법을 이용하여 얼굴영상  $F(i,j)$ 을 추출함

단계2: 얼굴영상  $F(i,j)$ 로부터 미리 정해진 기준에 따라 후보영역을 추출한다. Rohani와 같이 하단 1/3 영역을 기준으로 하였고, 폭은  $F(i,j)$ 의 절반으로 선택함

단계3: 입술 후보영역에 대하여 K-means(KM), Fuzzy c-means(FCM), 공간 Fuzzy c-means(SFCM)의 세 가지 분류기에  $|U_m - V_m|$ 의 특징치를 적용하여 영역분할을 함

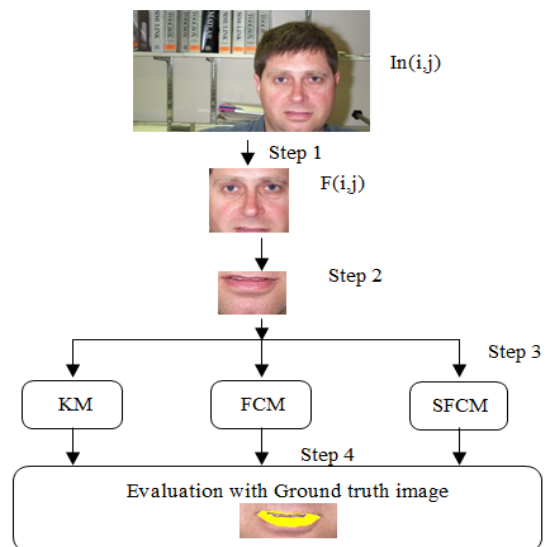


Fig. 10. Flowchart of proposed processing for the sample images.



단계4: 수작업으로 미리 구해 놓은 기준 입술영역과 추출된 입술영역의 겹치는 정도를 기준으로 성능을 평가함

본 논문에서는 우수한 특징치를 기존의 영역분할 방법(FCM/SFCM)에 적용하여 성능을 평가하므로, Liew와 Rohani 등이 사용한 복잡한 모델은 적용하지 않고 간단하게 처리하였다.

### 5. 실험 결과 및 고찰

본 논문에서 제안한 특징 치에 대한 성능평가는, 추출된 입술영역과 미리 준비된 기준 입술영역의 겹치는 정도를 이용하여 표현된다. 그림 11에서 기준 입술영역은 중간밝기의 녹색이고, 추출된 입술영역은 가장 어두운 밝기인 빨강으로 표시할 때, 겹쳐지는 가장 밝은 밝기인 노란색이 올바르게 추출된 유효한 화소들을 의미한다.

성능평가를 위하여 두 가지의 파라미터를 정의하여 사용한다.

- Lip\_hit : 기준영역(입술) - 추출영역(입술)
- Lip\_miss : 기준영역(입술) - 추출영역(피부)

샘플영상에 대한 실험결과를 그림 12에 나타내었다. 그림에서 (a)는 입력영상, (b),(c),(d)는 Liew등의 방법으로 CIELabUV의 5차원 특징 벡터를 사용하여



Fig. 11. Extracted lips region and ground truth lips region.

각각 KM, FCM, SFCM으로 처리한 결과를 나타낸다. (e),(f),(g)는 Rohani등의 가상 색상을 기준으로 처리한 결과이며 (h),(i),(j)는 제안한 특징 치인  $|U_m - V_m|$ 를 기준으로 처리한 결과이다. 전체 54장의 영상에 대한 각 방법의 처리결과를 표를 이용하여 나타내면 다음과 같다. 표 4의 제1,2 열은 K-means 방법의 Lip\_hit, Lip\_miss를 각각 나타내고 3,4열은 FCM, 5,6열은 SFCM의 결과이다.

전체적인 성능을 분석해 보면 제안한 색도 변환 특징치를 이용한 경우, 가장 좋은 결과를 얻을 수 있었다.

### 6. 결 론

본 논문에서는 단편적인 결합의 구조보다는 입술과 피부 두 개의 칼라 영역을 구분하는데 가장 유리한 좌표계에 대한 고찰을 하였으며, CIELUV좌표계가 분리성능에서 가장 우수함을 확인하였다. 기존의 논문들에서 볼 수 있는 특정 알고리즘에 준한 좌표계의 성능비교가 아닌, 관찰 데이터의 분포상태를 직접적으로 비교 및 분석할 수 있는 방법을 제안 하였다. 추가적으로 입술영역과 피부색영역의 각각 밀집도에서는 R/G좌표계가 가장 우수하나, 3.2절에서의 오 분류 비율이 10%에 근접하고, 계산 불가 항목 때문에 최상의 좌표계로 확인되지 않았다. 추후, R/G좌표계의 밀집도와 CIELUV좌표계의 우수한 분류성을 결합하여 분리성능이 우수한 데이터 변환 구조의 연구가 필요하다. 제안한 입술영역 추출을 위한 색도변환 특징치인  $|U_m - V_m|$ 는 평균인식률( $S_{hit}$ ) 96%로 다른 방법들에서 사용된 특징 치인 CIELabUV (84.6%), R/(R+G)(90.1%)에 비하여 우수한 성능을 보였고, 제안한 색도변환의 적용으로 성능이 개선되었다.

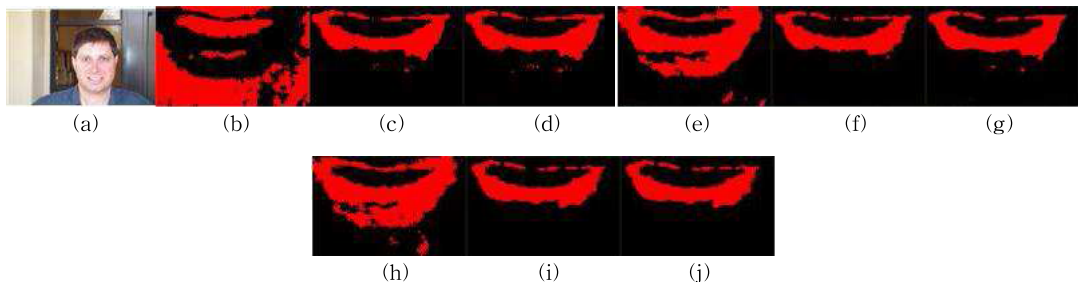


Fig. 12. Extracted lips region for the sample image.

Table 4. The matching ratio for Liew's CIELabUV case

Test Images	K_hit	K_miss	F_hit	F_miss	S_hit	S_miss
#1	94.1	172.1	88.2	41.6	88.2	45.1
#2	88.2	236.5	77.9	63.2	77.9	67.0
#3	99.8	312.1	95.6	129.6	95.7	136.8
#4	96.1	109.3	84.6	56.4	84.7	58.1
...						
#50	98.8	27.5	91.0	20.1	91.0	20.1
#51	1.9	180.0	91.7	18.7	91.7	18.7
#52	97.5	95.9	95.1	12.9	95.1	12.9
#53	99.5	37.2	94.7	10.0	94.7	10.7
#54	99.7	25.9	99.0	17.1	99.0	17.5
Avg	<b>70.1</b>	<b>130.1</b>	<b>86.4</b>	<b>64.6</b>	<b>84.6</b>	<b>69.8</b>
$\sigma$	42.7	114.8	24.6	91.3	27.4	101.3

Table 5. The matching ratio for Rohani's R/(R+G) case

Test Images	K_hit	K_miss	F_hit	F_miss	S_hit	S_miss
Avg	<b>91.3</b>	<b>78.8</b>	<b>89.4</b>	<b>39.3</b>	<b>90.1</b>	<b>43.2</b>
$\sigma$	18.9	73.9	14.1	36.7	14.0	40.2

Table 6. The matching ratio for the proposed |Um-Vm| case

Test Images	K_hit	K_miss	F_hit	F_miss	S_hit	S_miss
Avg	<b>96.6</b>	<b>71.3</b>	<b>96.0</b>	<b>29.8</b>	<b>96.0</b>	<b>30.0</b>
$\sigma$	13.3	74.5	3.3	21.6	3.3	21.6

## REFERENCES

- [ 1 ] G. Potamianos, C. Neti, J. Luettin, and I. Matthews, *Audio-Visual Automatic Speech Recognition: An Overview*, In: G. Bailly, E. Vatikiotis-Bateson and P. Perrier (Ed.), *Issues in Audio-Visual Speech Processing*, MIT Press, 2004.
- [ 2 ] A.L. Yuille, P.W. Hallinan, and D.S. Cohen, "Feature Extraction from Faces using Deformable Templates," *International Journal of Computer Vision*, Vol. 8, No. 2, pp. 99-111, 1992.
- [ 3 ] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham, "Active Shape Models: Their Training and Application," *Computer Vision and Image Understanding*, Vol. 61, No. 2, pp. 38-59, 1995.
- [ 4 ] C. Bregler and Y. Konig, "Eigenlips for Robust Speech Recognition," *Proceeding of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 669-672, 1994.
- [ 5 ] G.I. Chiou and J.N. Hwang, "Lipreading from Color Video," *IEEE Transactions on Image Processing*, Vol. 6, No. 8, pp. 1192-1195, 1997.
- [ 6 ] M. Vogt, "Fast Matching of a Dynamic Lip Model to Color Video Sequences under Regular Illumination Conditions," *NATO ASI Series F*, Vol. 150, D. G. Stork, M. E. Hennecke (Eds.), pp. 399-407, 1996.
- [ 7 ] M.U.R. Sanchez, J. Matas, and J. Kittler, "Statistical Chromaticity-based Lip Tracking with B-Splines," *Proceeding of the IEEE*

- International Conference on Acoustics, Speech and Signal Processing*, pp. 2973-2976, 1997.
- [ 8 ] D. Sanger, Y. Miyake, H. Haneishi, and N. Tsumura "Algorithm for Face Extraction Based on Lip Detection," *Journal of Imaging and Science and Technology*, Vol. 41, No. 1, pp. 71-80, 1997.
- [ 9 ] T. Wark, S. Sridharan, and V. Chandran, "An Approach to Statistical Lip Modelling for Speaker Identification via Chromatic Feature Extraction," *Proceeding of the 14th International Conference on Pattern Recognition*, Vol. 1, pp. 123-125, 1998.
- [10] T.W. Lewis and D.M.W. Powers, "Lip Feature Extraction using Red Exclusion," *Proceeding of Pan-Sydney Workshop on Visual Information Processing*, Vol. 2, pp. 61-67, 2001.
- [11] V. Vezhnevets, S. Soldatov, A. Degtiareva, and I.K. Park, "Automatic Extraction of Frontal Facial Features," *Proceeding of 6th Asian Conference on Computer Vision*, Vol. 2, pp. 1020-1025, 2004.
- [12] M. Wimmer and B. Radig, "Adaptive Skin Color Classifier," *Proceeding of the first ICGST International Conference on Graphics, Vision and Image Processing*, Vol. 1, pp. 324-327, 2005.
- [13] A.W. Liew, S.H. Leung, and W.H. Lau, "Segmentation of Color Lip Images by Spatial Fuzzy Clustering," *IEEE Transactions on Fuzzy Systems*, Vol. 11, No. 4, pp. 542-549, 2003.
- [14] Y. Wu, R. Ma, W. Hu, T. Wang, Y. Zhang, J. Cheng, and H. Lu, "Robust Lip Localization On Multi-view Faces In Video," *Proceeding of IEEE International Conference on Image Processing*, pp. 481-484, 2007.
- [15] S. Sridharan S. Lucey, and V. Chandran, "Adaptive Mouth Segmentation using Chromatic Features," *Pattern Recognition Letters*, Vol. 23, No. 11, pp. 1293-1302, 2002.
- [16] O. Erol, Y. Berkay, K. Harun, E. Hakan, and U. Mustafa, "Lip Segmentation using Adaptive Color Space Training," *Proceeding of the International Conference on Auditory-Visual Speech Processing*, pp. 219-222, 2008.
- [17] R. Rohani, F. Sobhanmanesh, S. Alizadeh, and R. Boostani, "Lip Processing and Modeling based on Spatial Fuzzy Clustering in Color Images," *International Journal of Fuzzy Systems*, Vol. 13, No. 2, pp. 65-73, 2011.
- [18] W. Wenchao, W. Shilin, K.E. Engin, M. Xiaoli, L. Shenghong, L. Jianhua, and L.M. Ni, "Optimization of Lip Contour Estimation," *Chinese Journal of Electronics*, Vol. 23, No. 2, pp. 341-347, 2014.
- [19] K. S. Jang and I. Lee, "Lip Shape Representation and Lip Boundary Detection Using Mixture Model of Shape," *Journal of Korea Multimedia Society*, Vol. 7, No. 11, pp. 1531-1539, 2004.
- [20] G. Wyszecki and W.S. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, John Wiley&Sons, New York, 1982.
- [21] MIT Facerec Database, <http://cbcl.mit.edu/software/datasets/heisele/facerecognition-database.html> (accessed April, 15, 2014).
- [22] R.O. Duda, P.E. Hart, and D.G. Stork, *Pattern Classification, 2nd edition*, Wiley-Interscience, Hoboken, NJ, 2000.
- [23] California Institute of Technology Face Database, [http://www.vision.caltech.edu/Image\\_Datasets/faces/faces.tar](http://www.vision.caltech.edu/Image_Datasets/faces/faces.tar) (accessed April, 15, 2014).
- [24] M. Nilsson, "Face Detection using Local SMQT Features and Split up Snow Classifier, [www.mathworks.fr](http://www.mathworks.fr) (accessed April, 15, 2014).



김 정 업

1986년 3월~1990년 2월 경북대  
학교 전자공학과 학사

1990년 3월~1992년 2월 경북대  
학교 대학원 전자공학과  
석사

1992년 3월~2001년 2월 경북대  
학교 대학원 전자공학과  
박사

1994년 6월~2001년 2월 삼성종합기술원 전문연구원

2001년 3월~2013년 2월 영산대학교 게임콘텐츠학과 부  
교수

2013년 3월~현재 영산대학교 자유전공학부 부교수

관심분야: 칼라영상처리, 컴퓨터 비전.