

컬러공간에 따른 영상내 사람 손 영역의 검출 성능연구

A Study on the Performance of Human Hand Region Detection in Images According to Color Spaces

김준엽* · 도용태**
(JunYup Kim · Yongtae Do)

Abstract - Hand region detection in images is an important process in many computer vision applications. It is a process that usually starts at a pixel-level, and that involves a pre-process of color space transformation followed by a classification process. A color space transformation is assumed to increase separability between skin classes for hands and non-skin classes for other parts, to increase similarity among different skin tones, and to bring a robust performance under varying illumination conditions, without any sound reasonings. In this work, we examine if the color space transformation does bring those benefits to the problem of hand region detection on a dataset of images with different hand postures, backgrounds, people, and illuminations. Results indicate that best of the color space is the normalized RGB.

Key Words : hand detection, color spaces, gesture

1. 서론

제스처는 오랫동안 인간의 의사소통수단으로 사용되어 왔던 통신 수단의 한 방법이다. 정보화 사회로 발전해감에 따라 사람과 컴퓨터와의 상호의사소통을 구현하기 위한 방법으로 보다 자연스러우며 단순하면서도 다양한 의미의 전달이 가능한 제스처를 이용한 사람과 컴퓨터의 인터페이스에 대한 연구가 활발히 진행되고 있다[1].

인간이 물체를 구별할 수 있는 가장 큰 특징 중의 하나가 컬러이며, 손 영역 검출에 관한 기존의 연구들도 대부분 컬러 정보를 사용한다. 컬러변환을 사용하는 이유는 첫째, 피부와 비피부 영역을 분리할 수 있는 능력을 증가시키기 위해서 이다. 둘째, 컬러가 조명 등 주위환경에 따라 민감하게 변하므로 이에 의한 영향을 줄이기 위해서 사용된다.

본 논문에서는 컬러 변환이 사람의 피부 영역 검출에 어떠한 차이를 보이는지에 대한 실험적인 결과를 보여줌으로써 손 영역 검출의 성능을 향상시킬 수 있는 근거를 제공한다. 카메라를 통하여 RGB영상을 획득하고, 이를 XYZ, YIQ, HSV, rgb(normalized RGB)로 변환한 후 각 컬러공간에 대하여 손 영역을 학습하고 다른 조명, 다른 사람, 다른 배경, 다른 손 모양에 적용하여 검출 성능에 어떠한 차이를 보이는지 살펴본다.

본 논문의 구성은 우선 2장에서 컬러변환을 다루고, 3장에

서 실험 및 결과를 제시하며, 마지막으로 4장에서 결론을 내린다.

2. 컬러 변환

카메라로부터 출력되는 영상은 대부분 RGB값이다. 우리는 이번 실험에서 모두 4가지 컬러공간 XYZ, YIQ, HSV, rgb(normalized RGB)를 사용하고, 각 컬러공간별로 XYZ, I, IQ, YIQ, H, HS, SV, HSV, rg, rb, gb 11개의 공간에서 비교한다. 각 컬러 공간은 다음 식에 의해서 변환한다[2].

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.619 & 0.177 & 0.204 \\ 0.299 & 0.586 & 0.115 \\ 0.000 & 0.056 & 0.944 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1.a)$$

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.586 & 0.115 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1.b)$$

$$H = \cos^{-1} \left(\frac{\frac{1}{2} [(R-G) + (R-B)]}{\sqrt{[(R-G)^2 + (R-B)(G-B)]^{1/2}}} \right) \quad (1.c)$$

$$S = 1 - \frac{3}{(R+G+B)} [\min(R, G, B)]$$

$$V = \frac{1}{3} (R+G+B)$$

저자 소개

*大邱大學校 大學院 情報通信工學科 碩士課程

**大邱大學校 電子情報工學部 正教授 · 工博

$$r = \frac{R}{R+G+B} \quad (1.d)$$

$$g = \frac{G}{R+G+B}$$

$$b = \frac{B}{R+G+B}$$

3. 결과

3.1 데이터 획득

우리는 세 대의 다른 특성을 가진 카메라로부터 총 372개의 영상을 획득하였다. 그림 1의 이미지는 다른 카메라, 다른 조명, 다른 배경, 다른 사람, 다른 손 모양의 조건아래 획득한 영상의 예이다.

첫 번째 카메라는 선광전자의 SK-2146AIN이며, 이 카메라는 1/3" Sony Color CCD의 이미지 센서로 30만화소, AGC(자동조정기능)내장, BLC(역광보정기능)탑재, 최저 조도 3Lux에서 촬영이 가능하다. 두 번째 카메라는 SECURA의 CCN-261 IA 이며, 1/3" IT CCD의 이미지 센서로 41만 화소, AGC, BLC, 최저 조도 0.8Lux에서 촬영이 가능하다. 세 번째 카메라는 ADS TECHNOLOGIES의 PYRO 1394 Web Cam으로 30만화소, AGC, WBS(White Balance Setting), 최저 조도 2.7Lux에서 촬영이 가능하다.

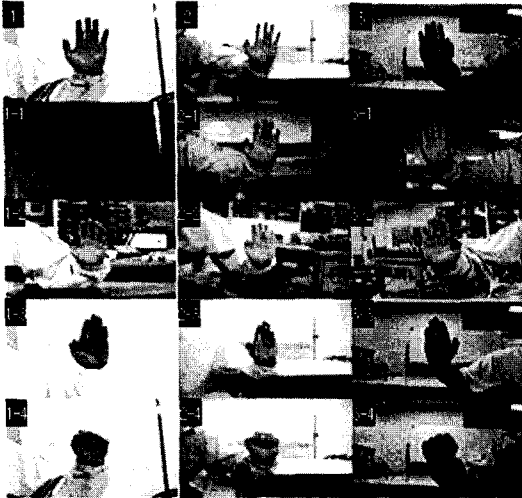


그림 1. 실험 영상의 예 : 1은 SK-2146AIN에서 획득한 영상이며, 2는 CCN-261 IA에서 획득한 영상이고, 3은 1394 Web Cam에서 획득한 영상이다. 1, 2, 3은 학습영상이고, 1-1, 2-1, 3-1은 다른 조명, 1-2, 2-2, 3-2는 다른 배경, 1-3, 2-3, 3-3은 다른 사람, 1-4, 2-4, 3-4는 다른 손 모양이다.

3.2 피부 화소 판정에 의한 손 영역 검출

획득한 영상에서 손 영역을 검출하기 위해 우선 학습용 영상 내에서 피부 영역인 그림 2의 사각 영역(a)내의 화소들을 각 컬러공간별로 평균(μ)과 표준편차(σ)를 구한다. 다음 식(2)

를 시험용 영상 전체에 적용하여 손 영역과 기타 영역(그림 2의 경우 영역(c))에 대하여 시험한다. 단, (b)영역은 손과 기타영역이 혼재하므로 실험에서 제외한다.

$$\mu - 2\sigma \leq I(x, y) \leq \mu + 2\sigma \quad (2)$$

특정 시험 영상에 대해 컬러 공간별로 적용한 후 결과는 그림 3에 나타내었다. 그림 4는 그림 3에서 획득한 피부 영역에 모폴로지 적용 후의 결과이다. 모폴로지 적용 후 XYZ, H공간에서는 더 악화 되었지만, 그 외의 공간에서는 더 향상 되었다.



그림 2. 영상내 손 영역과 기타 영역



그림 3. 컬러 공간별 피부 영역 획득 결과 : 1(RGB), 2(XYZ), 3(I), 4(IQ), 5(YIQ), 6(H), 7(HS), 8(SV), 9(HSV), 10(rg), 11(rb), 12(gb)



그림 4. 모폴로지 적용 후 결과

3.3 컬러 변환과 손 영역 검출 성능

결론부터 말한다면 컬러변환이 피부 검출에 도움을 준다. 그리고 그 성능은 카메라별로 다르다. SK-2146AIN카메라는 rgb가 거의 모든 상황에서 좋은 성능을 보이며, CCN-261 IA 카메라는 YIQ를 사용한 것이 가장 좋은 성능을 보이나, rgb를 사용한 것과 큰 차이는 없었다. 1394 Web Cam에서는 rgb가 가장 좋은 성능을 보였다. 3대의 카메라 모두 조명이 어두울수록 검출 성능이 낮았고, Web Cam은 배경의 변화에 따른 검출 성능의 영향이 상대적으로 컸다. 표 1은 각 카메라별로 피부 영역과 비피부 영역의 검출 성공률을 나타내고, 그림 5는 각 카메라 별로 검출 성능의 평균과 최악의 경우를 나타낸 것이다.

표 1. 피부 영역과 비피부 영역 검출 성공률

SK-2146AIN									
학습영상	배경1	배경2	조명1	조명2	손 모양1	손 모양2	다른 사람1	다른 사람2	
RGB	96.08	55.84	46.79	48.93	48.815	74.16	65.015	88.235	97.15
XYZ	95.185	53.275	45.255	47.815	46.215	73.25	63.725	90.005	97.08
YIQ	92.815	51.495	49.84	50	50	70.97	68.315	92.19	97.47
HSV	92.19	55.375	47.975	49.64	49.845	70.76	67.68	85.115	91.47
S.rg	95.515	68.56	92.145	97.925	68.03	87.095	88.095	94.465	93.33
S.rb	95.685	81.025	87.65	97.255	65.19	83.735	92.535	91.615	92.81
S.gb	96.36	80.185	88.425	97.82	67.655	83.59	95.83	92.4	91.26
CCN-261 IA									
RGB	95.425	67.04	68.98	71.7	74.95	74.95	78.595	95.5	89.585
XYZ	93.93	61.86	66.535	67.5	73.115	71.88	72.73	94.43	81.265
YIQ	96.13	71.136	71.845	74.745	53.245	73.31	80.865	97.68	73.85
HSV	95.325	75.13	76.115	77.37	51.47	74.405	75.75	97.86	81.815
S.rg	92.47	90.845	82.545	91.2	50.075	82.065	93.59	94.31	86.355
S.rb	87.54	88.545	80.735	82.395	49.16	78.105	90.98	93.07	85.14
S.gb	87.4	83.335	87.685	83.81	49.545	81.01	89.8	94.68	92.81
1394 Web Cam									
RGB	91.15	47.795	45.77	46.655	43.58	89.73	72.345	86.155	92.425
XYZ	85.49	43.74	42.315	44.785	35.99	82.895	65.71	74.695	85.845
YIQ	93.155	51.095	49.95	50.095	50.06	71.125	72.91	70.355	92.28
HSV	85.095	49.805	47.805	48.595	48.765	89.98	71.935	76.415	86.65
S.rg	93.89	91.07	95.935	95.355	80.88	77.655	81.89	76.765	89.33
S.rb	90.045	88.745	90.825	88.25	78.33	71.965	72.465	64.47	83.58
S.gb	89.995	86.31	91.01	88.16	76.955	73.21	71.68	62.895	81.105

3.4 조명에 따른 손 영역 검출 성능

표 1과 그림 3에서 조명이 어두울수록 모든 상황에서 검출 성능이 떨어지는 것을 볼 수 있다. 전체적으로 피부 검출을 위해 컬러의 구성 요소 중 밝기와 관련된 YIQ의 Y, HSV의 V값을 사용하지 않는 것이 피부와 비피부 검출 성능을 향상시킬 수 있었다.

4 결론

많은 컴퓨터 비전 알고리즘은 영상 내에서 손 영역을 검출하기 위해 컬러 변환을 사용한다. 본 논문은 컬러 변환이 실제로 검출 능력에 영향을 미치는지, 또 어떤 컬러 공간의 사용이 가장 좋은 결과를 가져오는지 실험적으로 조사하여 보았다. 그 결과 피부와 비피부 영역의 구분에 가장 적합한 컬러 공간은 정규화한 RGB 즉, rgb였다. YIQ컬러 공간도 경우에 따라서는 좋은 성능을 보이기는 하였지만, rgb가 거의 모든 상황에서 좋은 성능을 보였다.

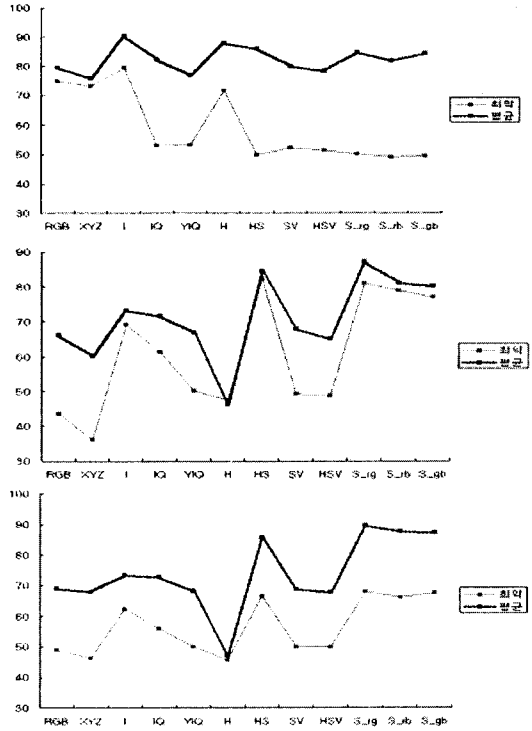


그림 5. 카메라별 검출 성공률의 요약 : (상) SK-2146AIN, (중) CCN-261 IA, (하) 1394 Web Cam

참고 문헌

- [1] Vladimir I. Pavlovic, Rajeev Sharma and Thomas S. Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction", IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 19, NO. 7, July 1997
- [2] Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing 2/E, Prentice Hall, September 2003