

동일인 인식을 위한 컬러 공간의 탐색 및 결합

남영호[†], 김민기^{**}

Color Space Exploration and Fusion for Person Re-identification

Young-Ho Nam[†], Min-Ki Kim^{**}

ABSTRACT

Various color spaces such as RGB, HSV, log-chromaticity have been used in the field of person re-identification. However, not enough studies have been done to find suitable color space for the re-identification. This paper reviews color invariance of color spaces by diagonal model and explores the suitability of each color space in the application of person re-identification. It also proposes a method for person re-identification based on a histogram refinement technique and some fusion strategies of color spaces. Two public datasets (ALOI and ImageLab) were used for the suitability test on color space and the ImageLab dataset was used for evaluating the feasibility of the proposed method for person re-identification. Experimental results show that RGB and HSV are more suitable for the re-identification problem than other color spaces such as normalized RGB and log-chromaticity. The cumulative recognition rates up to the third rank under RGB and HSV were 79.3% and 83.6% respectively. Furthermore, the fusion strategy using max score showed performance improvement of 16% or more. These results show that the proposed method is more effective than some other methods that use single color space in person re-identification.

Key words: Color Space, Fusion, Person Re-identification

1. 서 론

폐쇄회로텔레비전(CCTV)은 특정 구역에 대한 감시는 물론 현대 사회에서 발생하는 다양한 범죄를 사전에 예방하거나 사후 범인 검거를 위한 수단으로 활용되고 있다. 초기의 CCTV는 단순히 영상을 녹화하는 기능으로 출발하였으나, 디지털 영상처리 기술이 접목되면서 사람, 사물 등 특정 객체 및 행동을 자동으로 검출하고 식별하는 지능형 CCTV로 발전되고 있다. CCTV에 나타난 영상을 감시하는 것은 누구나 수행할 수 있는 단순한 일이지만, 장시간 여러 대의 카메라에서 획득된 영상들을 모니터링 해야

하는 경우에는 상황이 달라진다. 한 사람이 2대 이상의 모니터를 동시에 감시하게 되면 12분이 경과했을 때 특정 상황의 45%를 눈치 채지 못하며, 22분이 경과했을 때 95%를 놓치게 된다고 한다[1].

다수의 카메라가 네트워크로 연동된 환경에서 특정인의 동선을 파악하기 위해서는 동일인 인식(re-identification) 기술이 필요하다. 동일인 인식은 서로 다른 카메라에서 획득한 영상으로부터 동일인을 효과적으로 식별하는 방법을 말한다. 일반적으로 CCTV는 원거리에서 영상을 획득하기 때문에 사람의 얼굴을 식별할 수 있을 정도의 해상도를 제공하지 않는다. 이러한 이유로 동일인 인식에 대한 기존 연

* Corresponding Author : Min-Ki Kim, Address: (660-701) Jinju-daero 501, Jinju, Gyeongnam, Korea, TEL : +82-55-772-1393, FAX : +82-55-772-1389, E-mail : mkkim@gnu.ac.kr

Receipt date : Aug. 31, 2016, Revision date : Oct. 4, 2016
Approval date : Oct. 25, 2016

[†] Dept. of Computer Science in Gyeongsang National Univ. Engineering Research Institute
(E-mail : yhnam@gnu.ac.kr)

^{**} Dept. of Computer Science in Gyeongsang National Univ. Engineering Research Institute

구들은 원거리에서 특정인을 식별할 수 있는 전반적인 신체 외양, 걸음걸이, 그리고 착용한 의상에 대한 시각적 정보를 이용한다. 의상에 대한 시각적 정보는 여타의 특징에 비하여 식별력이 우수하고 짧은 시간 범위 내에서 안정성을 유지할 수 있기 때문에 이에 대한 많은 연구들이 이루어져 왔다.

의상에 나타난 시각적 정보는 컬러(color)와 텍스처(texture)로 표현할 수 있는데, 동일한 사람이라 하더라도 서로 다른 CCTV에 나타난 외형적 정보는 날씨나 촬영 시간에 따른 조도의 변화, 사람의 자세(pose)나 카메라의 시점(viewpoint)에 따른 형태 변형, 그리고 다른 사람이나 물체에 의한 부분적 가려짐(occlusion)이 발생하여 컬러와 텍스처 정보에 다양한 변형이 발생한다[2, 3]. 동일인 인식에 대한 기존의 연구들은 이러한 변형을 최소화하기 위하여 다양한 컬러 공간을 이용하고 있다. 조도의 변화에 따른 RGB 컬러 공간에서의 색상 변화를 줄이기 위하여 Nakagima 외[4]는 조도의 강도(intensity)를 이용하여 RGB 컬러를 정규화 하였으며, Bouma 외[5]는 RGB 채널별로 픽셀의 명도 값과 채널별 평균과의 차이를 표준편차로 나눈 표준점수 값으로 색상을 변환한 컬러 공간을 이용하였다. Park 외[6]는 HSV 컬러 공간에서 색상(hue)을 10개의 빈으로 구성하여 색 정보를 기술하였으며, Liu 외[7]와 Du[8]의 연구 결과에서 HSV 컬러 공간이 RGB 컬러 공간에 비하여 동일인 식별에 효과적인 것으로 나타났다. Lin 외[9]는 RGB 컬러 공간에서 픽셀의 색상 값을 무시하고 각 채널별로 명도 값의 상대적인 순위만을 이용하여 색 정보를 기술하였다.

컬러 정보는 텍스처에 비하여 저해상도 영상은 물론 사람의 자세나 카메라의 방향에 따른 변형에도 안정적인 특징 추출이 가능한 장점이 있지만, 조도에 따른 컬러 변화가 크게 나타나는 단점이 있다. 기존 연구들은 이러한 단점을 극복하고자 다양한 컬러 공간을 이용하고 있으나 특징 기술자, 정합 방법, 실험 데이터 등 실험 조건이 상이하여 어떤 컬러 공간이 동일인 인식에 효과적인지 판단하기 어렵다. 본 연구에서는 동일한 실험 조건하에서 조도 변화에 강인한 컬러 공간을 탐색한다. 그리고 서로 다른 컬러 공간에서 추출한 컬러 정보를 결합하여 동일인을 인식하는 방법을 제시한다. 본 논문의 구성은 다음과 같다. 2장에서는 대각선 모델(diagonal model)을 기반으로

기존 연구에서 도입되었던 컬러 공간에 대한 컬러 불변성을 살펴보고, 3장에서는 색차 기반의 히스토그램 정교화, 특징 정합, 그리고 컬러 공간의 결합 방법에 대하여 설명한다. 그리고 4장에서는 공개된 데이터베이스를 이용하여 다양한 컬러 공간의 특성과 제안된 동일인 인식 방법의 타당성을 제시한다. 끝으로 5장에서 결론과 향후 과제를 제시한다.

2. 영상의 컬러 표현

2.1 대각선 모델

영상 f 는 관찰자의 시점과 무관하게 표면에서 완전 확산이 이루어지는 람버티안 반사(Lambertian reflection)를 가정할 때 식 (1)과 같이 모델링될 수 있다[9, 10].

$$f(x) = \int_{\omega} e(\lambda)\rho_k(\lambda)s(x, \lambda)d\lambda, \text{ where } k \in \{R, G, B\} \quad (1)$$

여기서 $e(\lambda)$ 는 광원의 컬러, $\rho_k(\lambda)$ 는 카메라의 민감도 함수, $s(x, \lambda)$ 는 공간 좌표 x 에서의 표면 반사를 나타낸다. 그리고 ω 는 가시광선을 의미한다. 여기에 확산광(diffuse light)을 모델링한 $A(\lambda)$ 를 추가하면, 식 (2)와 같이 영상 f 에 대한 모델링을 확장할 수 있다.

$$f(x) = \int_{\omega} e(\lambda)\rho_k(\lambda)s(x, \lambda)d\lambda + \int_{\omega} A(\lambda)\rho_k(\lambda)d\lambda \quad (2)$$

조도의 변화는 식 (3)의 대각선 모델(diagonal model)로 나타낼 수 있다. $f^u(x)$ 는 알려지지 않은 광원에서 촬영된 영상이고, $D^{u,c}$ 는 알려지지 않은 광원 u 에서 얻어진 컬러를 표준 조도(canonical illumination) c 에서의 대응되는 컬러로 사상하는 대각 행렬이다. 즉, 대각선 모델은 임의의 환경에서 촬영된 영상을 대각 행렬에 의하여 표준 조도에서 촬영된 영상 $f^c(x)$ 로 변환하기 위한 모델이다.

$$f^c(x) = D^{u,c}f^u(x) \quad (3)$$

영상을 R, G, B 채널로 나타낼 때, $f^u(x)$ 의 픽셀을 $f^c(x)$ 의 픽셀로 변환하는 식은 식 (4)와 같이 나타낼 수 있고, 여기에 확산광을 추가하면 식 (5)와 같이 대각선 모델을 확장할 수 있다.

$$\begin{pmatrix} R^c \\ G^c \\ B^c \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} R^u \\ G^u \\ B^u \end{pmatrix} \quad (4)$$

$$\begin{pmatrix} R^c \\ G^c \\ B^c \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} R^u \\ G^u \\ B^u \end{pmatrix} + \begin{pmatrix} o_1 \\ o_2 \\ o_3 \end{pmatrix} \quad (5)$$

식 (4)의 대각선 모델에서 대각 행렬의 모든 요소가 동일한 경우($a=b=c$)는 광원의 밝기 변화(intensity change)를 의미하며, 대각 행렬의 요소가 서로 다른 경우($a \neq b \neq c$)는 광원의 컬러 변화(color change)를 나타낸다. 식 (5)의 확장 모델에서 대각 행렬 요소가 모두 1이고 채널별 확산광이 모두 동일한 경우($o_1 = o_2 = o_3$)는 광원의 밝기 이동(intensity shift)을 나타낸다. 밝기 변화는 광원의 밝기의 차이뿐만 아니라 음영에 의한 밝기 변화를 내포하며, 밝기 이동은 백색광의 산란을 포함한 확산광에 기인한다. 확장된 모델에서 대각 행렬의 모든 요소가 동일한 경우($a=b=c$)는 광원의 밝기 변화와 밝기 이동이 동시에 발생한 상태를 나타내며, 대각 행렬의 요소가 서로 다른 경우($a \neq b \neq c$)는 광원의 컬러 변화와 이동이 발생한 완전한 대각선 모델이다.

2.2 컬러 공간

RGB 컬러 모델로 표현된 영상은 광원의 밝기, 이동, 컬러 변화에 따라 화소 값이 변하기 때문에 이를 최소화하기 위하여 다양한 정규화 방식과 컬러 공간이 제안되었다. rg-정규화는 식 (6)에 제시된 바와 같이 화소의 R, G, B 성분을 전체 성분의 합으로 나누어 정규화한 것으로, 조도의 강도 변화에 불변하는 특징을 갖는다[4]. $r+g+b=1$ 이므로 b 는 중복된 정보가 된다. 따라서 정규화된 컬러 공간은 r 과 g 의 두 성분으로 나타낼 수 있으며, 이 컬러 공간은 밝기 변화에 불변하는 특성을 갖는다.

$$\begin{pmatrix} r \\ g \\ b \end{pmatrix} = \begin{pmatrix} R/(R+G+B) \\ G/(R+G+B) \\ B/(R+G+B) \end{pmatrix} \quad (6)$$

회색계(grey world) 정규화는 식 (7)과 같이 각 채널별 화소의 평균값으로 R, G, B 명도 값을 나누어 정규화한 것으로, 조도의 강도 변화에는 불변성을 만족하지 않지만 조도의 컬러 변화에는 불변하는 특성을 갖는다[11]. 식 (8)에 제시된 아파인(affine) 정규화는 영상의 채널 별 화소 값에 대한 평균과 표준편차를 이용하여 RGB 컬러 공간을 정규화한다[5, 12]. μ_c 와 σ_c 는 각 채널의 평균과 표준편차를 나타낸 것으로 채널별로 정규화를 수행하기 때문에 조도의 강

도 변화와 이동뿐만 아니라 조도의 컬러 변화와 이동에도 불변하는 특성을 갖게 된다.

$$\begin{pmatrix} r' \\ g' \\ b' \end{pmatrix} = \begin{pmatrix} R/\mu_R \\ G/\mu_G \\ B/\mu_B \end{pmatrix} \quad (7)$$

$$\begin{pmatrix} R' \\ G' \\ B' \end{pmatrix} = \begin{pmatrix} (R-\mu_R)/\sigma_R \\ (G-\mu_G)/\sigma_G \\ (B-\mu_B)/\sigma_B \end{pmatrix} \quad (8)$$

로그-색도(log-chromaticity) 컬러 공간은 녹색(G) 화소를 기준으로 각각 적색(R)과 청색(B) 화소의 비율에 로그 함수를 취한 것으로 식 (9)와 같이 표현된다[13]. ξ_1 과 ξ_2 는 조도의 밝기와 이동 변화에 불변하는 특성을 갖는다.

$$\xi_1 = \ln \frac{R}{G}, \quad \xi_2 = \ln \frac{B}{G} \quad (9)$$

HSV 컬러 공간은 색상(hue), 채도(saturation), 명도(value)로 컬러를 지정하는 방법이다. 명도(V)와 채도(S)는 조도 변화에 가변적이거나, 색상(H)은 두 채널간의 차이 값을 이용하므로 조도의 강도와 이동 변화에 불변하는 특성을 갖는다.

$$\begin{aligned} V &= \max(R, G, B) \\ S &= \begin{cases} (V - \min(R, G, B))/V, & \text{if } V \neq 0 \\ 0, & \text{otherwise} \end{cases} \\ H &= \begin{cases} 60(G - B)/(V - \min(R, G, B)), & \text{if } V = R \\ 120 + 60(B - R)/(V - \min(R, G, B)), & \text{if } V = G \\ 240 + (R - G)/(V - \min(R, G, B)), & \text{if } V = B \end{cases} \end{aligned} \quad (10)$$

대응(opponent) 컬러 공간은 적색과 녹색, 청색과 황색, 흑색과 흰색이라는 서로 대립되는 3개의 채널을 이용하여 컬러 공간을 구성한다[14]. O_3 는 밝기(luminance)를 나타내는 채널로 조도 변화에 가변적이거나, O_1 과 O_2 는 조도의 이동 변화에 불변하는 특성을 갖는다.

$$\begin{pmatrix} O_1 \\ O_2 \\ O_3 \end{pmatrix} = \begin{pmatrix} (R - G)/\sqrt{2} \\ (R + G - 2B)/\sqrt{6} \\ (R + G + B)/\sqrt{3} \end{pmatrix} \quad (11)$$

3. 동일한 인식을 위한 특징 기술 및 정합

동일인 인식에 사용되는 특징 기술과 정합 방법은 매우 다양하다. 본 논문에서는 컬러 기반의 특징을 이용하여 동일인 식별을 수행할 때, 이에 적합한 컬러 공간을 탐색하기 위하여 컬러 정보를 효과적으로 기술할 수 있는 히스토그램 기반의 특징 기술자를

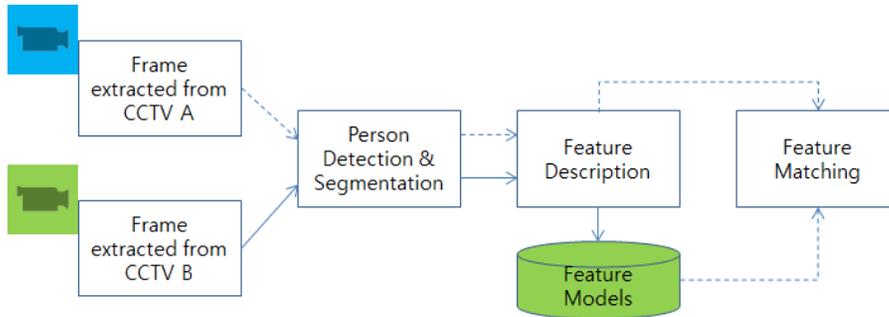


Fig. 1. Overall process of person re-identification.

이용한다. 컬러 히스토그램 기술자(CHD, color histogram descriptor)는 영상에 나타나는 컬러의 빈도를 나타낸 것으로, 카메라의 시점 변화에 강인하고 처리 속도가 빠른 장점이 있기 때문에 동일인 인식을 위한 특징 기술 방법으로 널리 이용되어 왔다[5, 6, 9, 15, 16].

3.1 인식 대상 분할

CHD는 특정 영역에 나타나는 컬러의 분포를 나타낸 것으로, 컬러 정보를 추출하기 위한 영역을 정의하는 것은 동일인 식별을 위한 중요한 처리 과정 중 하나이다. Fig. 1은 동일인 인식을 위한 전반적인 과정을 나타낸 것이다. 실선으로 나타낸 것은 인식 대상이 되는 사람들을 영상에서 분할하여 특징을 추출한 후 특징 모델을 구축하는 과정이고, 점선으로 나타낸 것은 인식할 대상을 분할하여 특징을 추출한 후 정합하는 과정을 나타낸 것이다. 결국 동일인 인식은 특징 모델에 저장된 특징 벡터와 인식 대상에서 추출한 특징 벡터의 정합에 의해 동일인을 인식하는 것이다.

CCTV 영상에서 사람을 검출하고, 검출된 사람 영역에서 머리, 몸통, 팔, 다리 등으로 영역을 정확하게 분할하는 문제는 도전적인 문제로 본 연구의 범위를 벗어난다. 본 연구에서는 검출된 사람 영상과 실루엣 영상으로부터 동일인을 인식할 수 있는 효과적인 특징 기술과 정합 방법에 대하여 기술한다. Fig. 2-(a)와 Fig. 2-(b)는 각각 검출된 사람 영상과 실루엣 영상을 60×136 크기로 정규화 한 것이고, Fig. 2-(c)는 배경 색에 의한 촬영 효과를 제거하기 위하여 사람 영상을 실루엣 영상으로 AND 연산에 의해 마스킹한 결과이다. Fig. 2-(d)는 미리 정의해 놓은 윈도우

에 의해 상체와 하체 부분으로 영역을 분할한 것으로, 관찰과 실험을 통해 머리, 상체(몸통과 팔), 허리, 하체(다리), 발 영역을 각각 4:11:6:11:2의 비율로 단순하게 수평 분할하였다. 본 연구에서는 5개 영역 중 상체와 하체 두 영역에서 특징을 추출한다.

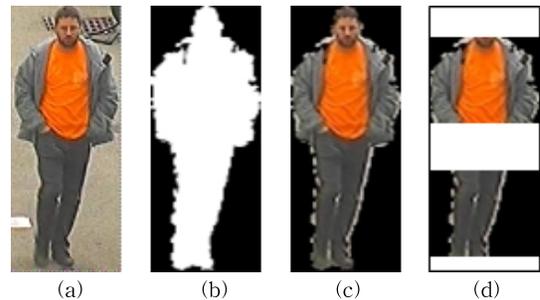


Fig. 2. Detected person image: (a) normalized image, (b) silhouette image, (c) masked image, (d) upper and lower parts.

3.2 색차 기반의 히스토그램 정교화

CHD는 히스토그램을 추출한 영역의 전역적인 컬러 분포만을 나타낼 뿐 컬러의 지역적인 배치는 무시되는 단점이 있다. 이러한 단점을 해소하기 위하여 히스토그램 정교화의 개념이 제시되었다[17]. 히스토그램 정교화는 화소의 위치나 주변 화소들 간의 관계에 따라 영상 내에 존재하는 화소들을 여러 개의 군집으로 분할하여 각 군집별로 히스토그램을 구하는 방식이다. 본 연구에서는 색차 벡터(CDV, color difference vector)[18]의 개념을 적용하여 히스토그램을 정교화 한다. 색차 벡터는 3×3 영역에서 중심 화소와 이웃 화소간의 색차에 따라 중심 화소를 분류하는 것으로, 동일한 색상이 넓게 분포하는 공간적 응집도가

높은 영역과 좁은 영역에 다수의 색상이 나타나거나 예지가 나타나는 공간적 응집도가 낮은 영역으로 화소를 분류하게 된다. 응집도가 높은 영역과 낮은 영역에서 i 번째 색상을 갖는 화소의 개수를 각각 α_i, β_i 라 하면, 색차 벡터는 $\langle (\alpha_1, \beta_1), (\alpha_2, \beta_2), \dots, (\alpha_n, \beta_n) \rangle$ 으로 기술된다.

색차를 이용하여 히스토그램을 정교화 하는 과정은 다음과 같다. 컬러 영상은 카메라의 RGB 센서에 의해 입력된 아날로그 신호를 각각 256레벨의 디지털 값으로 변환(즉, 양자화)한 것으로, RGB 컬러 영상으로부터 컬러 히스토그램을 구성한다면 256^3 개의 빈(bin)이 필요할 뿐만 아니라 조도 변화에 지나치게 민감하여 컬러 분포가 불안정하게 나타나게 된다. 따라서 이러한 문제를 해소하기 위하여 먼저 컬러 영상을 컬러 공간에 따라 독립적인 채널들로 분리한 후 각 채널 영상에 대하여 양자화 레벨을 축소한다. 다음으로 양자화 레벨이 축소된 영상에서 임의의 화소와 그 이웃 화소간의 명도 값 분포에 따라 해당 화소를 분류한다. 화소를 분류하는 기준으로 명도 값의 절대 값 차이가 충분히 큰 경우와 충분한 차이가 나타나는 빈도를 이용한다. 식 (12)와 식(13)에서 $p(x, y)$ 와 $p_i(x, y)$ 는 각각 중심 화소와 이웃 화소의 명도 값을 나타낸다. 중심 화소인 $p(x, y)$ 에 이웃한 화소는 총 8개가 존재하지만, 본 연구에서는 빠른 처리를 위하여 상, 하, 좌, 우의 4개의 이웃 화소($k=4$)를 사용하였다.

$$\frac{1}{k} \sum_{i=1}^k |p(x, y) - p_i(x, y)| > th_1 \tag{12}$$

$$\sum_{i=1}^k F(p(x, y) - p_i(x, y)) > th_2, \text{ where } F(\delta) = \begin{cases} 1, & \text{if } |\delta| > th_1 - 1 \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

식 (12)는 중심 화소와 이웃 화소간의 명도 값의 차이가 충분히 크게 나타나는지를 판단하기 위한 조건을 나타낸 것이고, 식 (13)은 중심 화소와 이웃 화소간의 명도 값 차이가 나타나는 빈도가 충분하게 큰 지를 판단하기 위한 조건을 나타낸 것이다. 중심 화소가 식 (12)와 식 (13)에 제시된 조건 중 적어도 하나를 충족시킬 때 중심 화소가 주변 화소들과는 다른 속성을 나타낸 것이므로, 중심 화소를 공간적 응집도가 낮은 영역으로 분류하고, 그렇지 않으면 공간적 응집도가 높은 영역으로 분류한다. 두 임계값 th_1 과 th_2 는 중심 화소를 분류하기 위한 임계값으로

너무 작게 설정하면 조도의 변화에 지나치게 민감하게 반응하게 되며, 너무 크게 설정하면 색차 정보를 적절하게 반영할 수 없다. 컬러 공간을 구성하는 각 채널의 양자화 레벨이 작아지면 임계값 th_1 도 작게 설정되며, 임계값 th_2 는 양자화 레벨과는 무관하고 중심 화소와 비교되는 이웃 화소의 개수에 따라 설정하게 된다. 두 임계값은 실험을 통해 설정하게 되는데, RGB 각 채널을 8레벨로 축소한 상태에서는 임계값 th_1 과 th_2 를 모두 1로 설정하였을 때 중심 화소의 공간적 응집도가 효과적으로 분류되는 것으로 나타났다.

3.3 특징 정합

특징 정합은 질의 영상(query image)에서 추출한 색차 벡터와 참조 영상(reference image)에서 추출한 색차 벡터의 비교를 통해 이루어진다. 질의 영상과 참조 영상에서 추출한 색차 벡터를 각각 V_Q, V_R 이라 할 때, 두 특징 벡터간의 거리는 식 (14)로 계산할 수 있다. 두 영상의 유사도 비교는 각각의 영상에서 추출한 특징 벡터의 비교를 통하여 이루어지는데, 특징 공간에서 두 특징 벡터의 거리가 가깝다는 것은 두 특징 벡터가 유사함을 의미하는 것이다. (α_i^q, β_i^q) 와 (α_i^r, β_i^r) 는 각각 질의 영상과 참조 영상에서 추출한 색차 벡터의 요소들로, 응집도가 높은 군집과 낮은 군집에 나타는 i 번째 색의 빈도를 나타낸다.

$$D(V_Q, V_R) = \sum_{i=1}^n (|\alpha_i^q - \alpha_i^r| + |\beta_i^q - \beta_i^r|) \tag{14}$$

두 군집에 포함된 화소의 개수는 서로 다르므로, 식 (14)에 의한 특징 거리는 군집의 크기에 비례하는 가중치를 갖게 된다. 그러나 일반적으로 색차가 크게 나타나는 화소의 개수는 그렇지 않은 화소에 비하여 출현 빈도는 낮지만 영상에 나타나는 중요한 특징을 내포하고 있다. 따라서 군집별 화소의 개수에 무관하게 동일한 가중치를 갖도록 각 군집별 히스토그램의 합이 1이 되도록 식 (15)에 의해 정규화한 후, 식 (16)과 같이 대응되는 영역별로 정규화된 히스토그램간의 특징 거리의 합으로 전체 특징 거리를 계산한다.

$$N(c_i) = c_i / \sum_{i=1}^n c_i \tag{15}$$

$$D'(V_Q, V_R) = \sum_{i=1}^n (|N(\alpha_i^q) - N(\alpha_i^r)| + |N(\beta_i^q) - N(\beta_i^r)|) \tag{16}$$

3.4 컬러 공간의 결합

하나의 컬러 공간이 여타의 컬러 공간이 갖는 장점을 모두 갖추고 있지 못하므로, 복수의 컬러 공간에서 특징을 추출하여 효과적으로 결합한다면 동일인 인식 성능을 개선할 수 있다. 본 연구에서는 실험을 통해 동일인 인식에 적합한 것으로 나타난 RGB와 HSV 컬러 공간으로부터 특징 추출 및 정합을 통해 생성된 특징 거리를 결합하는 방식을 취하였다. 특징 거리를 결합하는 방식으로 최소 점수(min score), 최대 점수(max score), 단순 합(imple sum), 가중 합(weighted sum) 방식 등이 있다[19]. 최소 점수와 최대 점수는 각각 특징 거리가 최소, 최대인 거리를 결합된 특징 거리에 설정하는 방식이고, 단순 합은 특징 거리의 산술적인 합을 결합된 특징 거리로 산출하는 방식이다. 가중 합은 각 특징 거리별로 다른 가중치를 곱한 후 합을 구하게 되는데, 가중치 산정은 식 (17)을 적용하였다. $Avg(D'_k(V_Q, V_R))$ 는 컬러 공간 k 에서 얻은 특징 거리의 평균을 나타내는 것으로, 각 컬러 공간에서 추출한 특징 벡터의 차원이 다름으로 인하여 생기는 특징 거리의 근원적인 차이를 보정하는 효과를 갖는다.

$$W_k = \frac{\sum_{k=1}^2 (Avg(D'_k(V_Q, V_R))) - Avg(D'_k(V_Q, V_R))}{\sum_{k=1}^2 (Avg(D'_k(V_Q, V_R)))} \quad (17)$$

4. 실험 결과 및 분석

4.1 컬러 공간의 불변성 측정

어떤 컬러 공간이 조도의 변화에 대하여 강인한 특성을 나타내는지 실험적으로 확인하기 위하여 학술연구의 목적으로 제작된 두 가지 공개 데이터를 이용하여 실험을 수행하였다. ALOI (Amsterdam Library of Object Image) 데이터는 실내에서 조명 조건을 변화시켜 가면서 물체를 촬영한 영상으로 [20], 본 실험에서는 물체 50개로부터 획득한 총 600개의 영상을 실험 대상으로 하였다. Fig. 3에 제시된 바와 같이 물체 1개당 조명 변화에 의한 12개의 영상이 존재한다. ImageLab 데이터는 실외에 설치된 CCTV에서 촬영된 영상으로부터 사람을 검출하여 해당 영역을 분할한 영상이다[21]. 50명에 대하여 CCTV가 피사체의 앞쪽, 뒤쪽, 좌측, 우측에 위치한



Fig. 3. Sample images from ALOI dataset.

곳에서 촬영한 보행자를 분할한 총 200개로 구성되어 있다. Fig. 4에서 사람의 크기가 달리 보이는 것은 카메라와 보행자와의 거리가 일정하지 않기 때문이다. 상단의 여성 보행자의 경우는 카메라의 시점과 무관하게 색상 정보에 큰 차이를 보이지 않지만, 하단에 나타난 남성의 경우 카메라의 시점이 앞, 뒤, 좌, 우로 변환에 따라 컬러 정보의 분포와 비율에 있어 큰 변화가 나타남을 알 수 있다.

ALOI 데이터를 이용한 실험에서 첫 번째 영상을 참조 영상, 나머지 영상들을 질의 영상으로 설정한 후, 각각의 컬러 공간으로부터 컬러 히스토그램을 추출하여 물체 인식을 수행하였다. 컬러 공간의 특성에 따라 양자화 레벨에 차이가 있으나, 가급적 특징 벡터의 차원에는 큰 차이가 없도록 양자화 레벨을 설정



Fig. 4. Sample images from ImageLab dataset: (a) front, (b) back, (c) left, (d) right.

하였다. HSV 컬러 공간은 색 정보가 H채널에 집중되어 있으므로 H채널은 16레벨, S와 V는 각각 4레벨로 양자화 레벨을 설정하였고, 정규화된 rg와 로그-색도($\xi_1 \xi_2$) 컬러 공간은 두 채널의 양자화 레벨을 모두 16으로 설정하였다. 이 외의 컬러 공간들은 각 채널별 양자화 레벨을 8로 설정하였다.

Table 1은 각 컬러 공간별 물체 인식에 대한 누적 인식률의 평균을 나타낸 것이다. 여기서 Rank N은 질의 영상에서 추출한 특징 벡터를 참조 영상에서 추출한 특징 벡터들과 비교하여 거리에 따라 오름차순으로 N순위까지 나열하였을 때 여기에 올바른 인식 결과가 포함된 비율을 나타낸 것이다. RGB와 HSV 컬러 공간보다는 영상의 밝기나 채널별 평균 등을 이용하여 RGB 컬러 공간을 정규화한 것이 더 우수한 성능을 제공하였다. 특히 대립되는 색상 차를 이용한 대응 컬러 공간이 조도 변화에 대한 불변성이 높은 것으로 나타났다.

ImageLab 데이터를 이용한 실험에서는 정면(front) 영상을 참조 영상으로 하고, 나머지 영상들을 질의 영상으로 설정하였다. 각 컬러 공간별 양자화 레벨을 ALOI 데이터를 이용한 실험과 동일하게 설정하여 CHD를 추출하고 인식 실험을 수행하였다. Table 2

는 각 컬러 공간별 동일한 인식에 대한 누적 인식률의 평균을 나타낸 것이다. 동일한 인식에 대한 실험 결과는 ALOI 데이터를 이용한 실험과는 달리 정규화한 컬러 공간보다는 RGB와 HSV 컬러 공간에서 추출한 CHD가 동일한 인식에 보다 효과적인 것으로 나타났다. 카메라의 시점을 변화시키지 않고 실내에서 촬영된 ALOI 영상들은 형태 변화 없이 단지 조명에 의한 컬러만 변경되기 때문에 정규화를 통해 어느 정도 컬러 변화에 대한 불변성이 유지되었다. 그러나 ImageLab은 카메라의 시점이 변하기 때문에 형태 변화가 크게 나타나고 실외에서 촬영되었기 때문에 조도 환경의 변화가 크게 발생함으로 인하여 컬러 정규화 과정에서 오히려 정보의 왜곡이 발생한 것으로 판단된다.

4.2 제안된 방법의 성능 측정

ImageLab 영상들은 조도 변화는 물론 카메라의 시점 변화에 따른 다양한 변형이 수반되었으며, 동일한 인식 과정에서 정규화된 컬러 공간으로는 이러한 변형을 효과적으로 처리하는데 한계를 보였다. 동일한 인식에 대한 실험 결과 RGB와 HSV 컬러 공간이 여타의 컬러 공간에 비하여 우수한 성능을 보였으므로, 본 연구에서는 색차 정보를 기반으로 히스토그램을 정교화 방식을 RGB와 HSV 컬러 공간에 적용하여 동일한 인식을 수행하였다. 그리고 기존 방식과 동일한 인식에 대한 성능을 비교하기 위하여 CHD 방식과 형태 문맥 기술자(SCD, shape context descriptor)[7] 방식을 구현하여 실험하였다. SC(shape context)는 서로 다른 영역에서 추출한 임의의 두 점에 대한 컬러 공간상의 관계를 극좌표상에 누적한 히스토그램이다. SCD는 RGB 컬러 공간을 로그-색도의 2차원 공간으로 변환한 후, 상체와 하체 영역에 속하는 임의의 두 점간의 거리($\log r$)와 방향(θ)에 따라 2차원 극좌표에 빈도를 누적하여 특징을 기술한다. Fig. 5는 정면 영상을 참조 영상으로 하고, 나머지 영상들을 질의 영상으로 설정한 실험에서, 질의 영상이 우측 영상이었을 때의 누적 인식률 특성(CMC, cumulative matching characteristic) 곡선들을 보여주고 있다. CDV를 이용한 제안하는 방식이 RGB와 HSV 컬러 공간 모두에서 CHD 방식에 비하여 높은 인식률은 보였으며, SCD 방식에 비해서도 동일한 인식에 보다 효과적인 것으로 나타났다.

Table 1. Object recognition rates on ALOI dataset

Color Space	Rank1	Rank2	Rank3	Rank10
RGB	55.8	67.5	73.3	91.8
rg	59.6	68.2	74.9	88.4
$r'g'b'$	72.0	80.5	88.0	96.7
$R'G'B'$	72.2	80.9	87.5	96.7
$\xi_1 \xi_2$	68.4	78.5	83.6	97.3
HSV	45.3	55.8	63.1	87.1
$O_1 O_2 O_3$	78.2	87.1	90.5	96.7

Table 2. Re-identification rates on ImageLab dataset

Color Space	Rank1	Rank2	Rank3	Rank10
RGB	40.7	50.7	56.7	81.3
rg	17.3	22.7	27.3	40.7
$r'g'b'$	32.0	39.3	43.3	69.3
$R'G'B'$	31.0	39.3	43.3	70.0
$\xi_1 \xi_2$	5.3	11.3	14.7	35.3
HSV	40.7	55.3	60.0	81.3
$O_1 O_2 O_3$	11.3	18.7	23.3	62.0

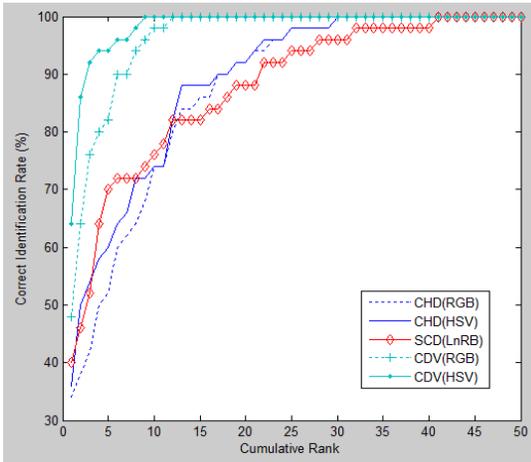


Fig. 5. Performance comparison using CMC curves.

Table 3에서 CDV(RGB)와 CDV(HSV)는 제안된 방법을 단일 컬러 공간에 적용하여 얻은 누적 인식률의 평균이고, 나머지는 RGB와 HSV 두 개의 컬러 공간에서 얻은 특징 거리의 결합을 통해 얻은 누적 인식률의 평균이다. 1순위 인식률을 기준으로 컬러 공간의 결합 방식에 의한 인식률을 비교해 보면, 최소 점수에 의한 결합 방식을 제외하고는 모두 인식 성능이 개선되었다. 최소 점수에 의한 결합 방식은 서로 다른 색이 조도 변화에 의하여 한 개의 컬러 공간에서라도 유사하게 나타나면 인식 오류로 이어질 수 있기 때문인 것으로 판단된다. 이와 반대로 나머지 결합 방식은 두 개의 컬러 공간상의 거리를 더하거나 최대 거리를 이용함으로써 이러한 오류가 나타날 가능성이 줄어들게 된다.

Table 4는 컬러 공간의 결합으로 인하여 얻은 성능 향상 비율을 식 (18)을 이용하여 계산한 결과이다. 식 (18)에서 R_s 는 단일 컬러 공간을 이용하였을 때의

Table 3. Re-identification results using fusion of color spaces

Color Space & Fusion	Rank1	Rank2	Rank3	Rank10
CDV(RGB)	52.0	71.3	79.3	96.7
CDV(HSV)	57.4	76.5	83.6	97.8
Min Score	53.3	72.0	79.3	98.7
Max Score	66.7	85.3	89.3	99.3
Simple Sum	62.7	82.7	89.3	98.7
Weighted Sum	62.7	82.0	88.7	98.7

Table 4. Performance improvement by fusion of color spaces

	CDV(RGB)	CDV(HSV)
Min Score	2.5%	-7.1%
Max Score	28.3%	16.2%
Simple Sum	20.6%	9.2%
Weighted Sum	20.6%	9.2%

인식률, R_f 는 두 개의 컬러 공간을 결합하였을 때의 인식률을 의미한다. 본 실험에서는 최대 점수를 이용한 결합 방식이 가장 높은 성능 향상을 보였으며, RGB와 HSV 단일 컬러 공간을 이용한 방식에 비하여 각각 28.3%, 16.2%의 성능 향상을 보여 컬러 결합 방식이 효과적임을 확인할 수 있었다. 그러나 최대 점수 방식이 단순 합이나 가중 합 방식에 비하여 다른 데이터 세트에 대해서도 가장 좋은 결과를 나타낼 수 있다고 단정하기는 어렵다. 제안된 방법은 두 개의 컬러 공간에서 추출된 특징 벡터를 이용하기 때문에 단일 컬러 공간을 이용하는 방법에 비하여 약 2배의 처리 시간이 소요되나, 한 개의 컬러 공간에서 특징을 추출하고 정합하는데 까지 1ms 이내로 처리가 가능하므로 인식 성능 향상을 위해 감내할 수 있는 수준이다.

$$I(s, f) = \frac{(R_f - R_s)}{R_s} \times 100 (\%) \quad (18)$$

5. 결 론

본 연구에서는 동일인 인식에 적합한 컬러 공간을 탐색하기 위하여 다양한 컬러 공간을 적용하여 물체 인식 및 동일인 인식 실험을 수행하였다. 채널별 컬러 값의 평균이나 표준편차를 이용하여 컬러를 정규화하거나 대립 색의 차이를 이용하는 컬러 공간은 영상에서 컬러의 통계적 분포가 크게 변하지 않는 조건에서 우수한 불변성을 보였다. 그러나 CCTV 영상에서와 같이 조도 변화는 물론 카메라의 시점 변화가 나타나는 환경에서는 오히려 정규화 과정에서 왜곡이 발생하여 인식의 정확도를 떨어뜨리는 것으로 나타났다. 이에 반하여 RGB와 HSV 컬러 공간은 조도 변화보다는 카메라 시점에 따른 형태 변화에 강한 특성을 나타내었다. 이러한 실험 결과를 토대로, RGB와 HSV 컬러 공간에 색차 기반의 히스토그램 정규화 기법을 적용하여 3순위까지의 누적 인식률을

각각 79.3%, 83.6%로 끌어 올릴 수 있었다. 그리고 두 컬러 공간에서 얻은 특징 거리를 최대 점수로 결합하였을 때, RGB와 HSV 단일 컬러 공간을 이용한 방식에 비하여 각각 28.3%, 16.2%의 성능 향상을 보여 컬러 결합 방식이 효과적임을 확인할 수 있었다. 본 연구에서는 미리 정의해 놓은 두 영역에 나타나는 전역적인 컬러 정보를 이용하므로, 카메라 시점의 변화에 따른 형태 변화에 적응하는데 한계가 있다. 따라서 향후 연구에서는 검출된 사람 영역에서 머리, 몸통, 팔, 다리 등으로 세부 영역을 분할하는 연구가 필요하다.

REFERENCE

- [1] N. Dadashi, *Automatic Surveillance and CCTV Operator Workload*, Master's Thesis of University of Nottingham, 2008.
- [2] X. Li, A. Wu, M. Cao, J. You, and W. Zheng, "Towards More Reliable Matching for Person Re-identification," *Proceedings of IEEE International Conference on Identity, Security and Behaviour Analysis*, pp. 1-6, 2015.
- [3] R. Zhao, W. Ouyang, and X. Wang, "Person Re-identification by Saliency Matching," *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2528-2535, 2013.
- [4] C. Nakajima, M. Pontil, B. Heisele, and T. Poggio, "Full-Body Person Recognition System," *Journal of Pattern Recognition*, Vol. 36, No. 9, pp. 1997-2006, 2003.
- [5] H. Bouma, S. Borsboom, R.J.M. den Hollander, and M. Worring, "Re-identification of Persons in Multi-camera Surveillance Under Varying Viewpoints and Illumination," *Proceedings of Sensors and Command, Control, Communications, and Intelligence Technologies for Homeland Security and Homeland Defense XI*, SPIE Vol. 8359, pp. 1-10, 2012.
- [6] U. Park, A.K. Jain, I. Kitahara, K. Kogure, and N. Hagita, "ViSE: Visual Search Engine Using Multiple Networked Cameras," *Proceedings of International Conference on Pattern Recognition*, pp. 1204-1207, 2006.
- [7] I. Kviatkovsky, A. Adam, and E. Rivlin, "Color Invariants for Person Reidentification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 35, No. 7, pp. 1622-1634, 2013.
- [8] S. Bak, E. Corvee, F. Bremond, and M. Thonnat, "Person Re-identification Using Spatial Covariance Regions of Human Body Parts," *Proceedings of IEEE International Conference on Advanced Video and Signal Based Surveillance*, pp. 435-440, 2010.
- [9] Z. Lin and L.S. Davis, "Learning Pairwise Dissimilarity Profiles for Appearance Recognition in Visual Surveillance," *Advances in Visual Computing*, LNCS Vol. 5358, pp. 23-34, 2008.
- [10] K. Sande, T. Gevers, and C. Snoek, "Evaluating Color Descriptors for Object and Scene Recognition," *IEEE Transactions on Pattern Recognition and Machine Intelligence*, Vol. 32, No. 9, pp. 1582-1596, 2010.
- [11] J. M. Buenaposada and L. Baumela, "Variations of Grey World for Face Tracking," *Image Processing and Communications*, Vol. 7, No. 3-4, pp. 51-61, 2001.
- [12] D. Cong, L. Khoudour, C. Achard, C. Meurie, and O. Lezoray, "People Re-identification by Spectral Classification of Silhouettes," *Signal Processing*, Vol. 90, pp. 2362-2374, 2010.
- [13] D. Berwick and S. Lee, "A Chromaticity Space for Specularity, Illumination Color- and Illumination Pose-Invariant 3-D Object Recognition," *Proceedings of the IEEE International Conference on Computer Vision*, pp. 165-170, 1998.
- [14] R.M. Anwer, D. Vazquez, and A.M. Lopez, "Opponent Colors for Human Detection," *Pattern Recognition and Image Analysis*, LNCS Vol. 6669, pp. 363-370, 2011.
- [15] C. Nakajima, M. Pontil, B. Heisele, and T.

Poggio, "Full-Body Person Recognition System," *Journal of Pattern Recognition*, Vol. 36, No. 9, pp. 1997-2006, 2003.

[16] N. Gheissari, T.B. Sebastian, P.H. Tu, and J. Rittscher, "Person Reidentification Using Spatiotemporal Appearance," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1528-1535, 2006.

[17] G. Pass and R. Zabih, "Histogram Refinement for Content-Based Image Retrieval," *Proceedings of the IEEE Workshop on Applications of Computer Vision*, pp. 96-102, 1996.

[18] M. Kim, "Image Retrieval Using Histogram Refinement Based on Local Color Difference," *Journal of Korea Multimedia Society*, Vol. 18, No. 12, pp. 1453-1461, 2015.

[19] J. Kittler, M. Hatef, R.P.W. Duin, and J. Matas, "On Combining Classifiers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No. 3, pp. 226-239, 1998.

[20] Amsterdam Library of Object Images, <http://aloi.science.uva.nl> (accessed Aug., 1, 2015).

[21] D. Baltieri, R. Vezzani, and R. Cucchiara, "3D Body Model Construction and Matching for Real Time People Re-Identification," *Proceedings of Eurographics Italian Chapter Conference*, pp. 65-71, 2010.



남 영 호

1989년 경상대학교 전자계산통계학과(이학사)
 1991년 중앙대학교 전자계산학과(공학석사)
 1994년 중앙대학교 컴퓨터공학과(공학박사)

1996년~2010년 경상대학교 컴퓨터교육과 교수
 2011년~현재 경상대학교 컴퓨터과학과 교수
 관심분야 : 병렬처리, 영상처리, 임베디드시스템



김 민 기

1989년 2월 중앙대학교 전자계산학과 이학사
 1994년 8월 중앙대학교 전자계산학과 이학석사
 1998년 2월 중앙대학교 컴퓨터공학과 공학박사

1998년~1999년 KRIC 선임연구원
 1999년~2000년 KERIS 책임연구원, 연구위원
 2000년~2010년 경상대학교 컴퓨터교육과 부교수
 2011년~현재 경상대학교 컴퓨터과학과 교수
 2013년~2014년 미국 University of Notre Dame 방문학자
 2014년~현재 경상대학교 공학연구원(ERI) 자동화·컴퓨터연구센터 책임연구원
 관심분야 : 생체인식, 패턴인식, 영상처리