

특집논문-06-11-1-02

## 인물 기반 사진 색인을 위한 인물 특징값 개발에 관한 연구

양 승 지<sup>a)</sup>, 서 경 석<sup>a)</sup>, 노 용 만<sup>a)†</sup>, 김 상 균<sup>b)</sup>

### Exploiting Person-identity Features for Person-based Photo Indexing

Seungji Yang<sup>a)</sup>, Kyong Sok Seo<sup>a)</sup>, Yong Man Ro<sup>a)†</sup>, and Sang-Kyun Kim<sup>b)</sup>

#### 요 약

본 논문에서는 일련의 디지털 사진들을 특정한 인물에 기반하여 브라우징하는 것을 돕기 위한 인물 기반 사진 색인 방법을 제안한다. 기존의 인물 기반 영상 색인의 경우, 얼굴 특징값만을 인물 특징값으로 이용하였다. 이에 반하여, 제안하는 방법의 주된 목적은 얼굴의 주변 정보인 상황 및 옷 정보를 얼굴 정보와 함께 조합하여 색인 과정에 활용함으로써 인물 기반 사진 색인의 성능을 높이는 것이다. 얼굴 특징값과 옷 특징값을 효과적으로 조합하기 위해, 인물을 상황 별로 클러스터링하는 방법이 함께 제안된다. 본 논문에서 제안하는 방법의 효용성을 검증하기 위해 1120장의 사진 데이터베이스를 활용하여 실험을 수행하였다. 실험 결과에서, 기존의 방법에서처럼 얼굴 특징값만을 이용하여 색인을 수행한 경우 약 70%의 평균 성능을 보였으며, 제안된 방법을 이용하여 색인을 수행한 경우 약 92%의 평균 성능을 보여 제안된 방법이 인물 별 사진 색인에 효과적임을 확인할 수 있었다.

#### Abstract

In this paper, a novel approach is addressed to facilitate the browsing of large collection of digital photos associated with specified person(s) in the photos. The goal of the proposed method is to exploit additional person-identity features as incorporating facial regions and peripheral clothes region associated with them. For more effective incorporation of the clothes and facial features, situation-based photo clustering is also proposed. To evaluate the efficacy of the proposed method experiment was performed with 1120 generic home photos. The experiment results showed that the proposed method outperformed the conventional method using only face feature as showing the average performance of about 92% contrary to the average performance of about 70% in the conventional method.

Keyword: 인물 기반 사진 분류, 상황 기반 사진 분류, 얼굴 인식, 내용 기반 특징, MPEG-7

### I. 서 론

최근 디지털 카메라의 보급으로, 디지털 사진의 양이 증

가하고 있다. 특히, 최근 저장 장치의 집적화로 용량이 증가하고 있고, 화질을 크게 손상시키지 않는 디지털 영상 압축 기술이 개발됨에 따라, 사람들은 자신의 일생을 디지털 사진으로 기록하고 저장할 수 있게 되었다. 이에 따라, 많은 양의 사진들을 보다 효과적으로 관리하기 위한 장치와 도구가 필요하게 되었고, 효율적인 디지털 사진 앨범 방식에 대한 사용자의 요구가 증가하고 있다.

a) 한국정보통신대학교 공학부 영상비디오시스템연구실  
Image and Video Systems Lab., School of Engineering, Information and Communications Univ. (ICU)  
b) 삼성종합기술원 컴퓨팅연구실  
Computing Lab., Samsung Advanced Institute of Technology (SAIT)  
† 교신저자 : 노용만(yro@icu.ac.kr)

기존의 디지털 사진 앨범은 사용자가 촬영한 사진을 디지털 카메라나 메모리 카드 등으로부터 사용자의 로컬 저장 장치로 옮겨오고, 이를 관리하기 위한 간단한 기능만을 제공한다. 사용자들은 사진 앨범을 이용하여 사진들을 시간/인물/카테고리 별로 색인하길 원하지만, 이러한 과정의 대부분이 사용자의 수동 작업에 의해 이뤄짐에 따라, 사용자들로 하여금 많은 불편을 야기시키고 있다. 최근 몇몇 연구에서 디지털 사진 앨범의 사용자 요구 사항을 조사하였다<sup>[1][2]</sup>. 사진을 촬영시간 순으로 정렬하는 것만으로도 앨범을 구성하는 것이 유용하다고 조사되었으나, 사용자의 요구 사항을 보다 충실히 만족시키기 위해서는 내용 기반 검색이나 이벤트 혹은 인물 기반 사진 색인 등의 기능이 추가되어야 한다고 밝히고 있다.

특히, 인물 기반 사진 색인은 앨범의 기능 가운데 가장 중요한 기능의 하나로써, 사진을 다른 사람과 공유하거나, 사진을 인물 별로 브라우징하는데 유용하게 사용된다. 인물 색인 과정에서 가장 중요한 부분은 얼굴 검출 및 얼굴 인식 부분이다. 즉, 사진에서 얼굴 영상을 검출하고 이를 데이터베이스의 얼굴 영상들과 매칭하여 해당 얼굴을 인식하는 과정이다. 이를 위해서 영상 처리 방법에서 패턴 인식을 수행할 때와 거의 같은 과정을 거치게 되는데, 영상 획득, 전처리, 얼굴 추출, 얼굴 인식 등의 과정을 거치게 된다.

영상으로부터 얼굴을 검출하거나 얼굴을 인식하는 방법에 대한 연구는 지금까지 매우 많이 이루어졌다<sup>[3][4][5]</sup>. 얼굴 인식 분야는 특히 보안 시스템을 위주로 발전하였다. 예를 들어, 지능형 무인 감시 시스템이나 출입자 통제 시스템, 그리고 범죄 용의자 검색등의 시스템을 위한 얼굴 추출 및 인식에 대한 연구가 주로 수행되었다. 이러한 응용 분야에서는 인물의 얼굴을 보다 정확하게 추출하고, 외부 조명이나, 인물의 표정 및 포즈에 대한 강인한 얼굴 인식 방법을 개발하는데 연구의 초점이 맞춰져 있다.

이에 반하여, 사진에서 얼굴을 추출하거나 인식하는 과정은 일반적으로 보안 시스템에서 획득된 영상으로부터 얼굴을 추출하거나 인식하는 과정보다 어렵다. 보안 시스템에서는 고정된 공간에 설치된 고정된 카메라를 사용하여 영상을 획득하기 때문에, 영상으로부터 추출된 얼굴 영상은 비교적 정형화된 배경을 가지고 있다. 이에 반하여 사진은

사용자가 카메라를 가지고 다니면서 찍은 영상이기 때문에, 사진 속의 배경은 일반적으로 매우 자주 변하고 다양하다. 뿐만 아니라, 사용자가 카메라를 조작하는 방법(예를 들어, 카메라 줌의 사용이나 플래시의 사용 등) 또는 사진을 찍을 당시의 카메라의 방향에 따라 고정된 공간에서도 다양한 배경을 가질 수 있으며, 외부 조명이나 조도의 변화가 훨씬 크다.

그럼에도 불구하고, 개인의 디지털 사진 모음에서 인물 기반 색인은 세 가지 특성을 가진다. 하나는 색인 대상 인물의 범위가 가족이나 친구 등의 작은 그룹으로 한정된다는 점이다. 다른 하나는 사진에서 관심의 인물은 대부분 정면 얼굴을 포함하고 있다는 점이다. 또한, 대부분의 디지털 사진은 Exif(Exchangeable Image File) 형식을 따르고 있다<sup>[6]</sup>. Exif는 일본 전자공업발달협회에서 만든 표준 사진 파일 형식으로, 사진의 픽셀 정보뿐만 아니라, 사진을 찍을 당시의 시간 정보 등의 촬영 정보 및 카메라 상태를 저장하고 있다. 본 논문에서는 이러한 관찰에 기반하여, 기존에 이용되던 인물의 얼굴 특징값뿐만 아니라, 촬영 정보, 사진을 찍을 당시의 상황 정보와 인물들이 입은 옷 정보 등의 다양한 부가 정보를 인물 식별 특징값으로 이용함으로써 사진을 인물 별로 자동 색인하는 방법을 제안한다.

## II. 본 론

일반적으로 개인이 촬영한 디지털 사진은 다양한 배경을 가진다. 이는, 소형 디지털카메라를 이용하여 장소나 시간에 구애 받지 않고, 언제 어디서나 사진 촬영이 가능하기 때문이다. 그러나, 고정된 장소에서 여러 장의 사진을 촬영하는 패턴이 일반적이기 때문에, 근접한 시간 내에 촬영된 사진은 유사한 배경 및 전경을 포함하는 경우가 많다. 본 논문에서는 비슷한 장소에서 촬영한 사진들의 모음을 하나의 상황 클러스터(situation cluster)로 정의하고, 다음의 세 가지 특성을 관찰하였다. (1) 같은 상황 클러스터에 속한 사진들은 촬영 시간이 비슷하고 비슷한 배경을 가진다. (2) 특정 사람들이 같은 상황 클러스터에 속한 사진들 내에 자

주 촬영되었을 가능성이 높다. (3) 또한, 상황 클러스터 내의 사진들 간의 시간 간격이 좁다면 촬영된 인물들은 같은 옷을 입고 있을 가능성이 매우 높다. 본 논문에서는 이러한 관찰에 기반하여, 색인할 사진을 상황 별로 분류하고 인물의 얼굴 정보뿐만 아니라 옷 정보를 인물 식별 특징값으로 이용하고자 한다.

그림 1은 본 논문에서 제안하는 사진의 인물 색인 방법에 대한 절차를 나타낸 것이다. 먼저, 입력 사진열을 촬영 시간에 기반하여 오름차순으로 정렬한다. 사진의 촬영 시간 정보는 사진의 Exif로부터 추출할 수 있다. 다음으로, 정렬된 사진들을 순차적으로 검사하여, 사진들간의 시간

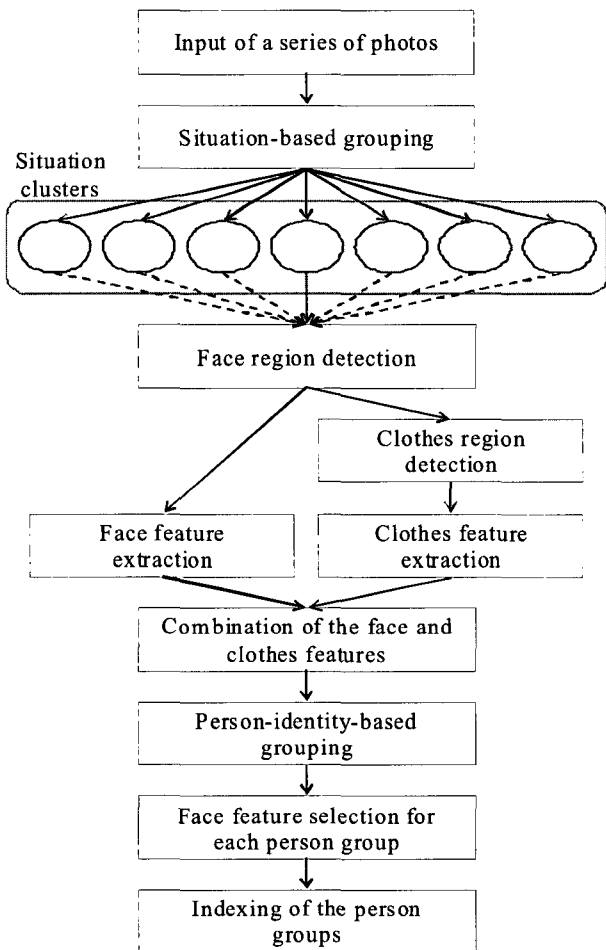


그림 1. 제안된 방법의 절차  
Fig. 1. Overall procedure of the proposed method

차이 및 시각 특성의 유사도를 측정한다. 측정된 유사도 값은 각 사진에서 상황 변화가 발생되었는지의 여부를 결정하기 위해 사용된다. 상황 기반 클러스터들이 결정되면, 각 사진으로부터 각 인물의 얼굴 영역을 검출하고 대응되는 인물의 옷 영역을 검출한다. 검출된 얼굴영역과 옷 영역으로부터 특징값을 추출하고, 주어진 상황 클러스터 내에서 비슷한 얼굴 및 옷 특징값들을 클러스터링한다. 생성된 인물 클러스터는 동일한 인물들로 구성되어 있다고 가정하고, 각 인물클러스터에 속한 인물들의 얼굴 특징값들은 평균을 내어 클러스터 대표 특징값을 생성한다. 마지막으로, 각 인물 클러스터의 특징값과 트레이닝된 인물 특징값과 비교하여 색인하는 과정을 수행한다. 그림 2는 기존의 인물 기반 영상 색인 방법과 제안하는 방법의 차이를 나타낸다. 그림 2-(a)는 기존의 인물 기반 영상 색인 방법으로, 추출된 얼굴 특징값을 데이터베이스 내에 얼굴 특징값과 비교하여 색인하는 과정만을 수행한다. 그림 2-(b)는 제안하는 인물 기반 영상 색인 방법으로 비슷한 상황에서 촬영된 사진들 가운데 비슷한 얼굴 특징값과 옷 특징값을 함께 조합하여 하나의 인물 클러스터를 생성하고 생성된 클러스터를 대표하는 하나의 대표 얼굴 특징값을 생성한 후, 해당 인물 클러스터 내에 속하는 인물들을 데이터베이스 내에 얼굴 특징값과 비교하여 색인하는 과정이다.

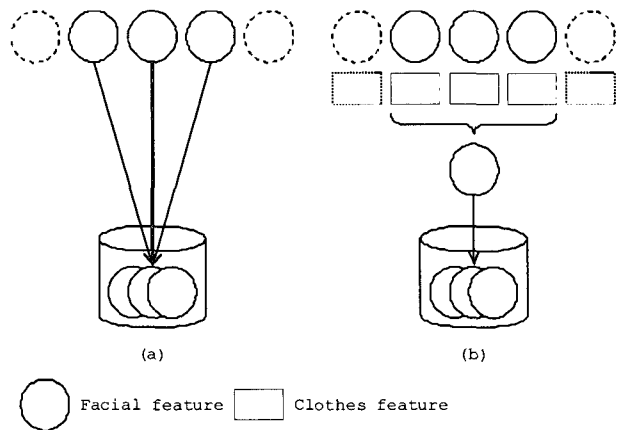


그림 2. 기존 방법과 제안하는 방법의 차이, (a) 기존의 인물 기반 영상 색인 방법, (b) 제안하는 방법  
Fig. 2. Comparison between traditional person-based image indexing method and the proposed method

## 1. 상황 기반 사진 클러스터링

색인할 사진열이 입력되면 먼저 사진을 상황에 기반하여 클러스터링하는 과정을 수행한다. 사진을 상황에 기반하여 클러스터링하기 위해 입력된 N개의 사진열을 촬영 시간에 대해 오름차순으로 정렬한다. 정렬된 사진열에서, 인접된 사진들 간의 상황 유사도를 시간 특징값 및 색상, 질감, 모양 등의 내용 기반 특징값을 이용하여 측정한다. 인접한 두 사진, j번째 사진과 j-1번째 사진 간의 유사도 차이는 다음과 같은 수학적식에 의해 계산된다.

$$d(j, j-1) = \exp \left[ d_s(j, j-1) \cdot \sum_{f \in F} \{w_f \cdot d_f(j, j-1)\} \right], \quad (1)$$

여기서,  $d_s$ 는 시간 특징값의 유사도 거리값이고,  $d_f$ 는 내용 기반 특징값의 유사도 거리값이다. 임의의 두 사진 간의 시간 특징의 유사도 거리값은 초 단위로 환산된 시간값의 차이이다. 내용 기반 특징의 유사도 거리값은 내용 기반 특징값의 집합 F에 속하는 특징값 f에 대해 측정된 유사도 거리값  $d_f$ 를 합산한 값으로, 특징값 f에 상응하는 가중치  $w_f$ 를 적용하여 합산된다. 시간 및 내용 기반 특징값은 서로 독립적이므로 최종 유사도 거리값은 이 두 값을 곱하여 표현한다. 임의의 두 유사도 거리값을 곱하여 얻는 결과는 두 유사도 거리값들 가운데 어느 한 값의 변화에 크게 의존한다. 따라서, 시간 유사도 값이나 내용 기반 유사도 값 가운데, 어느 한 값이 크면 상황 변화가 발생했다고 간주하기 쉽다. 이에 반하여, 두 유사도 거리값들을 더하여 얻는 결과는 시간 유사도 값이나 내용 기반 유사도 값이 상대적으로 모두 커야 상황 변화가 발생했다고 간주하기 쉽다. 일반적으로, 내용 기반 특징의 유사도가 크더라도 시간 유사도는 작을 수 있고, 반대의 경우도 있기 때문에, 이러한 경우를 고려하기 위해 시간 특징 유사도 값과 내용 기반 특징 유사도 거리값을 곱하였다. 또한, 지수 함수를 사용하여, 동일 상황 간의 시간 및 내용 기반 특징 차이를 보다 좁히고, 서로 다른 상황 간의 차이를 크게 하여, 상황 변화 지점을 검출하기 용이하게 하였다. 본 논문에서는, 입력된 N개의 사진열에서, 인접한 두 개의 사진에 대해 측정된 유사도 거리값 d가 임계값 보다

클 경우에 상황 변화가 발생되었다고 간주한다. 여기서, 임계값은 실험적으로 결정되었다. 이러한 비교를 N개의 사진열에 대해 N-1번 수행하여, 첫 번째 사진을 제외한 N-1개의 사진에서 상황 변화가 발생했는지에 대한 여부를 측정한다. 결과적으로, 입력된 N개의 사진열은 M개의 상황 클러스터로 그룹되며, 각 사진은 단 하나의 상황 클러스터에 속한다. 여기서, 임의의 사진 i가 상황 클러스터 j에 속한다면,  $1 \leq i \leq M$ ,  $1 \leq j \leq N$ ,  $M \leq N$ 의 조건을 만족한다.

## 2. 인물 특징 추출

얼굴 인식(face recognition)을 위해 선행되어야 할 기본 과정은 얼굴 영역을 검출하는 일(face detection)이다. 지금까지 얼굴 검출에 대한 많은 연구가 수행되어 왔다<sup>[3][4][5]</sup>. 본 논문에서는 얼굴 영역 검출 방법에 대해 다루지 않는 대신에, 얼굴 영역이 임의의 방법에 의해 신뢰적으로 검출되었다고 가정한다. 사진에서 인물의 얼굴 영역들이 검출되면, 검출된 얼굴 영역에 기반하여 대응되는 인물의 옷 영역을 검출한다. 그림 3은 인물의 얼굴 영역에 상응하는 옷 영역을 검출하는 방법을 도시한다. 그림에서와 같이, 얼굴 영상이 가로  $w_f$  픽셀, 세로  $h_f$  픽셀의 크기로 정규화되었다면, 옷 영역은 인물의 목에 해당하는 부분을 제외시키

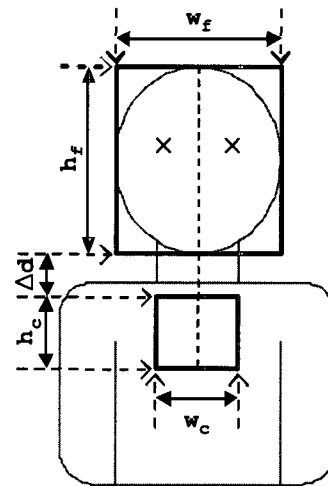


그림 3. 인물의 얼굴 영역에 상응하는 옷 영역 검출  
Fig. 3. Detection of corresponding clothes region to face region

기 위해 옷 영역보다  $\Delta u$ 의 크기만큼 차이를 두고 가로  $w_c$  픽셀, 세로  $h_c$  픽셀 크기로 옷 영역을 검출한다. 여기서, 옷 영역의 크기가 얼굴 영역보다 크지 않도록 설정한다( $w_c < w_f, h_c < h_f$ ). 이는 옷 영역을 검출할 때, 가능한 한 다른 인물이나 객체의 간섭을 피하기 위함이다.

얼굴 및 옷 영역이 검출되면, 해당 영역으로부터 내용 기반 인물 특징값으로 얼굴 특징값과 옷 특징값을 추출한다. 본 논문에서는 MPEG-7의 Advanced Face Recognition Descriptor (AFRD)를 이용하여 얼굴 특징값을 추출한다<sup>[10][11]</sup>. AFRD는 총 48개의 특징값으로 구성되며, 이 가운데 24개의 값은 얼굴 영상 전체 영역에서 추출한 푸리에(Fourier) 계수값으로 구성되며, 나머지 24개의 값은 얼굴 영상의 중앙 부분에서 추출한 24개의 푸리에 계수값으로 구성된다.

일반적으로 옷의 특징은 색상과 질감으로 표현될 수 있다. 본 논문에서는, 옷의 색상 특징을 추출하기 위해, MPEG-7의 Color Structure Descriptor(CSD)를 이용하며, 옷의 질감 특징을 추출하기 위해, MPEG-7의 Edge Histogram Descriptor(EHD)를 이용한다<sup>[11]</sup>. CSD는 영상의 색상 값 빈도수를 측정하여 전역적인 색상 특성만을 고려하는 색상 히스토그램 방법을 보완하기 위해, 영상의 지역적인 색상 구조 특성을 고려하는 방법이다. 색상 값의 수는 {256, 128, 64, 32} 가운데 하나로 결정된다. EHD는 다섯 종류의 에지(edge) 성분의 45도, 90도, 135도, 180도, 무방향(non-directional) 에지 성분 빈도수를 영상으로부터 추출하여 특징값으로 이용한다. 이를 위해, 전체 영상을 16개의 서브-블록(sub-block) 영상으로 나누고, 각 서브-블록 영상에서 다섯 종류의 에지 성분에 대한 빈도수를 측정하여 총 80개(= 16 x 5)의 특징값을 구성한다.

상기에서와 같이 인물 특징값이 추출되면, 상황 I에 속하는 사진  $j$ 에 존재하는 인물  $k$ 의 특징값은  $\vec{f}_{ijk} = \{\vec{f}_{face}, \vec{f}_{color}, \vec{f}_{texture}\}$ 와 같이 표현된다. 여기서,  $\vec{f}_{face}, \vec{f}_{color}, \vec{f}_{texture}$ 는 각각 얼굴 특징값 벡터, 옷의 색상 특징값 벡터, 옷의 질감 특징값 벡터를 나타낸다. 두 인물의 특징값  $\vec{f}$ 와  $\vec{f}'$ 의 유사도 비교는 유클리디안(Euclidean) 거리 측정 방법인  $d(\vec{f}, \vec{f}') = \sum_{e=1}^E \|f_e - f'_e\|^2$

을 이용한다. 여기서,  $E$ 는 특징값의 차원을 나타내고,  $f_e$ 는  $e$ 번째 특징값을 나타낸다. 이를 이용하여 임의의 두 인물의 얼굴 특징값인  $\vec{f}_{face}$ 와  $\vec{f}'_{face}$ 의 유사도 거리값은  $d_{face}(\vec{f}_{face}, \vec{f}'_{face})$ 로 표현되고, 옷 색상 특징값  $\vec{f}_{color}$ 와  $\vec{f}'_{color}$ 의 유사도 거리값은  $d_{color}(\vec{f}_{color}, \vec{f}'_{color})$ 로 표현되며, 질감 특징값의  $\vec{f}_{texture}$ 와  $\vec{f}'_{texture}$ 의 유사도 거리값은  $d_{texture}(\vec{f}_{texture}, \vec{f}'_{texture})$ 로 표현된다. 임의의 상황 클러스터 I 내에 존재하는 사진  $j$ 와  $j'$ 이 있을 때, 사진  $j$ 에 존재하는 인물  $k$ 의 특징값인  $\vec{f}$ 와 사진  $j'$ 에 존재하는 인물  $k'$ 의 특징값인  $\vec{f}'$ 간의 유사도 거리값  $d_{combined}(\vec{f}(i, j, k), \vec{f}'(i, j', k'))$ 은 다음 식에서와 같이 계산한다.

$$\begin{aligned} d_{combined}(\vec{f}(i, j, k), \vec{f}'(i, j', k')) &= w_{face} \cdot \tilde{d}_{face} + w_{color} \cdot \tilde{d}_{color} + w_{texture} \cdot \tilde{d}_{texture} \\ d_{combined}(\vec{f}(i, j, k), \vec{f}'(i, j', k')) &= w_{face} \cdot \tilde{d}_{face} + w_{color} \cdot \tilde{d}_{color} + w_{texture} \cdot \tilde{d}_{texture} \end{aligned} \quad (2)$$

여기서,  $\tilde{d}$ 는 0.0에서 1.0까지의 값으로 정규화된 유사도 거리값이다. 이는 얼굴 특징값과 옷 특징값의 유사도 거리값이 서로 다른 범위를 가지기 때문에, 0.0에서 1.0까지의 값으로 정규화된 유사도 거리값을 얻어 각 특징값을 조합하기 위한 것이다.  $w_{face}, w_{color}, w_{texture}$ 는 각각 얼굴 특징값의 중요도, 옷의 색상 특징값의 중요도, 옷의 질감 특징값의 중요도를 나타내며,  $w_{face} + w_{color} + w_{texture} = 1.0$ 이 되도록 한다.

### 3. 인물 기반 사진 클러스터링 및 색인

일반적으로 같은 상황 클러스터 내에 인물들은 비슷한 환경 하에서 촬영되었을 가능성이 높기 때문에, 외부 조도의 변화나 옷 정보의 변화가 적을 것이다. 본 논문에서는 이러한 관찰에 기반하여, 각 인물 특징값을 바로 색인하지 않고, 색인 과정 이전에 유사한 인물 특징값을 클러스터링하는 단계를 거친다. 즉, 색인 과정 이전에 주어진 상황 클러스터 내에서 동일한 인물 특징값을 클러스터링하는 과정

을 수행한다. 이는 상황 클러스터 내에서 얻을 수 있는 인물 특징 정보를 최대한 활용하여 색인 과정에서 발생하는 에러를 줄이는데 그 목적이 있다.

임의의 상황 클러스터 내에 몇 명의 인물 클러스터가 존재하는지에 대한 선행 정보(prior knowledge)가 존재하지 않기 때문에, **unsupervised** 클러스터링 방법을 이용해야 한다. 본 논문에서는 대표적인 **unsupervised** 클러스터링 방법 가운데 하나인 **agglomerative** 클러스터링 방법을 이용하여 인물 기반 사진 클러스터링을 수행한다<sup>7)</sup>. 이 방법의 단계별 내용은 다음과 같다.

- (1) 먼저, 임의의 상황에서 검출된 모든 인물 특징들을 단독으로 구성된 인물 클러스터로 분할한다. 만일 임의의 상황 클러스터 내에 S개의 인물 특징값이 존재한다면 S개의 클러스터가 생성되며 각 클러스터 내에는 하나의 인물 특징값들이 속하게 된다.
- (2) 임의의 두 인물 클러스터 P, Q내의 인물 특징값의 수를 각각  $n_p$  와  $n_q$  라고 할 때, 두 인물 클러스터 간의 유사도 거리값  $d_{cluster}(P, Q)$ 는 다음과 같이 계산된다.

$$d_{cluster}(P, Q) = \frac{\sum_{p \in P} \sum_{q \in Q} d_{combined}(f_p, f_q)}{n_p \cdot n_q}, \quad (3)$$

여기서 p, q는 P, Q 인물 클러스터 내에 속하는 인물들을 나타내며,  $f_p, f_q$ 는 각각 P, Q 인물 클러스터 내에 속하는 인물 특징값을 나타낸다.

- (3) 유사도 거리값이 가장 작은, 즉 가장 유사한 두 개의 인물 클러스터를 찾고, 선택된 두 인물 클러스터들을 하나의 인물 클러스터로 합친다.
- (4) 어떤 클러스터 간의 유사도 거리값  $d_{cluster}(P, Q)$ 도 정해진 임계값보다 작지 않을 때까지, (2), (3) 과정을 반복한다. 여기서, 임계값을 작게 하면 클러스터의 수가 많아지고, 크게 하면 클러스터의 수가 상대적으로 적어진다.

인물 기반 사진 클러스터링이 완료되면, 모두 R개의

인물 클러스터  $r_1, r_2, r_3, \dots, r_R$ 가 생성된다. 다음으로 각 인물 클러스터의 얼굴 특징값을 평균화하여 클러스터 r의 대표 얼굴 특징값  $\hat{f}_r$ 을 얻는다. 여기서 얼굴 특징값만을 인물 클러스터의 대표 특징값으로 이용하는 이유는, 인물 색인 과정에서는 촬영 시간에 관계없이 유사도 정도를 비교해야 하기 때문이다. 즉, 과거에 트레이닝된 인물 특징값들과 입력된 새로운 인물 특징값을 비교하여 유사도가 가장 큰 인물 클러스터를 결정하는 과정이기 때문이다. 마지막으로, 인물 클러스터 대표 얼굴 특징값  $\hat{f}_r$ 을 이미 트레이닝된 인물 특징값 모델들과 아래와 같이 비교한다.

$$v_r = \arg \max_{v \in V} \Theta_v(\hat{f}_r) \quad (4)$$

여기서,  $\Theta_v$ 는 트레이닝된 전체 인물 집합 V에 속하는 인물 v의 결정 함수(decision function)이다. 본 논문에서는 결정 함수로 SVM(support vector machine)을 이용하였다<sup>8)12)</sup>. SVM은 주어진 테스트 데이터의 클래스를 결정하기 위한 이진 분류 함수를 생성하는 방법 가운데 하나이다. 결정 함수  $\Theta_v$ 의 입력값은 인물 클러스터 대표 얼굴 특징값  $\hat{f}_r$ 이며, 출력 값은 해당 인물 v에 대한  $\hat{f}_r$ 의 신뢰도를 정수로 표현한 값이다. 음수일 경우, 절대값이 커질수록 해당 인물 v에 대한  $\hat{f}_r$  신뢰도가 낮다는 의미이며, 양수일 경우 절대값이 커질수록 해당 인물 v에 대한  $\hat{f}_r$  신뢰도가 높다는 의미이다. 결과적으로, 신뢰도가 가장 큰 인물 집합  $v_r$ 을 해당 인물 클러스터  $\hat{f}_r$ 의 인물로 색인한다.

그림 4는 본 논문에서 제안하는 인물 기반 사진 색인의 예를 도식화한 것이다. 그림의 예에서, 상황 클러스터 j내에 모두 5장의 사진이 있으며, 총 8개의 얼굴 영역이 검출되었다. 검출된 얼굴 영역을 이용하여, 옷 영역을 추출하고, 얼굴 및 옷 특징값을 추출한다. 다음으로, **agglomerative** 클러스터링 방법을 이용하여, 3개의 인물 클러스터로 군집화하고, 각각의 인물 클러스터의 얼굴 특징값을 평균화하여 대

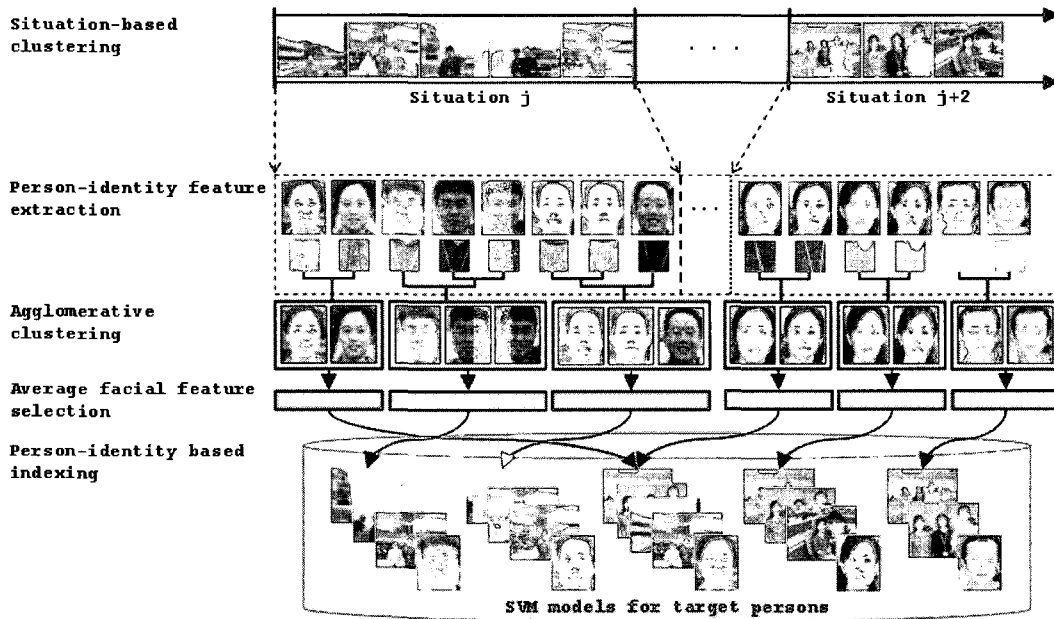


그림 4. 제안된 인물 기반 사진 색인의 예  
 Fig. 4. Example of the proposed person-identity based photo indexing

표 특징값을 추출한다. 각 인물 클러스터의 대표 특징값은 이미 트레이닝된 5명의 인물 특징 모델과 비교함으로써 가장 근사한 인물로 색인한다.

### III. 실험 및 결과

본 논문에서 제안된 방법을 검증하기 위하여, 1120장의 디지털 사진을 대상으로 실험을 수행하였다. 본 데이터베이스는 MPEG-7 Visual Core Experiment 3(VCE-3)의 공식 데이터베이스로 인물 기반 사진 클러스터링을 검증하기 위한 실험 데이터베이스이며<sup>[9]</sup>, 모두 27명의 인물 사진들로 구성되어 있다. 또한, 각 사진의 인물에 대한 얼굴 영역 정보를 포함하고 있으며, 얼굴 영역에 대한 정보는 2개의 눈 위치로 표현되어 있다. 본 실험에서는 1120장의 사진들 가운데 684장의 얼굴 영상을 트레이닝(training) 영상으로 이용하였다. 트레이닝 방법은 상기 기술하였듯이, SVM을 이용하였고, SVM의 커널(kernel)은 선형(linear) 커널과 비선형 커널 가운데 하나인 Gaussian RBF(radial basis func-

tion) 커널을 이용하여 비교 실험하였다. 표 1은 각 인물에 대한 훈련 및 테스트 영상 정보를 나타낸다. 27명의 인물에 대한 트레이닝 영상이므로, 각 인물당 사용된 영상의 수는 약 25장이다. SVM 훈련을 위하여 각 인물의 훈련 영상은 'positive sample'과 'negative sample'로 분류하였다. 'Positive sample'은 임의의 인물이 포함된 사진을 나타내고, 'negative sample'은 해당 인물을 제외한 26명의 인물들이 포함된 사진을 나타낸다. 테스트 영상으로는 451 장의 얼굴 영상을 이용하였고, 훈련 영상과 마찬가지로 27 명의 인물을 포함하고 있으므로, 각 인물당 약 17장의 영상으로 구성되었다. 얼굴 특징값으로는 MPEG-7의 AFRD를 이용하였고, 옷 특징값으로는 옷의 질감 정보를 추출하기 위해 MPEG-7의 EHD를 이용하였으며, 옷의 색상 정보를 추출하기 위해 MPEG-7의 CSD를 이용하였다. 또한, CSD를 이용하는 과정에서 외부 밝기 변화에 강인한 특징값 추출을 위해, MPEG-7의 Illumination Invariant Color Descriptor (IICD)를 적용하였다. 인물의 얼굴 영역은 MPEG-7 AFRD의 제안에 따라,  $w_f = 46$ ,  $h_f = 56$ 으로 정규화하였다. 옷 영역을 추출하는 과정에서 인물의 목 부분을 제외하기 위

표 1. 훈련 및 테스트 영상 정보

Table 1. Training and test dataset information

| Person index         | Initial name of target persons | Training images             |                             | Testing images           |
|----------------------|--------------------------------|-----------------------------|-----------------------------|--------------------------|
|                      |                                | # of positive facial images | # of negative facial images | # of ground truth images |
| 1                    | JCK                            | 36                          | 36                          | 24                       |
| 2                    | KDHb                           | 34                          | 34                          | 23                       |
| 3                    | HWJ                            | 34                          | 34                          | 21                       |
| 4                    | PJH                            | 34                          | 34                          | 21                       |
| 5                    | JJS                            | 32                          | 32                          | 22                       |
| 6                    | KMY                            | 32                          | 32                          | 21                       |
| 7                    | SCY                            | 30                          | 30                          | 19                       |
| 8                    | KCH                            | 28                          | 28                          | 19                       |
| 9                    | KDHa                           | 28                          | 28                          | 18                       |
| 10                   | LJH                            | 28                          | 28                          | 17                       |
| 11                   | SHS                            | 26                          | 26                          | 18                       |
| 12                   | HEH                            | 26                          | 26                          | 17                       |
| 13                   | KTK                            | 24                          | 24                          | 17                       |
| 14                   | LY                             | 24                          | 24                          | 17                       |
| 15                   | KJHa                           | 24                          | 24                          | 16                       |
| 16                   | LKE                            | 24                          | 24                          | 15                       |
| 17                   | LSJb                           | 22                          | 22                          | 15                       |
| 18                   | KHW                            | 22                          | 22                          | 14                       |
| 19                   | KWS                            | 22                          | 22                          | 13                       |
| 20                   | UKW                            | 20                          | 20                          | 14                       |
| 21                   | KCS                            | 20                          | 20                          | 14                       |
| 22                   | LKC                            | 20                          | 20                          | 14                       |
| 23                   | SKA                            | 20                          | 20                          | 13                       |
| 24                   | KHY                            | 20                          | 20                          | 13                       |
| 25                   | JJK                            | 18                          | 18                          | 12                       |
| 26                   | JYJ                            | 18                          | 18                          | 12                       |
| 27                   | LHS                            | 18                          | 18                          | 12                       |
| Total # of samples   |                                | 684                         | 684                         | 451                      |
| Average # of samples |                                | 25.3                        | 25.3                        | 16.7                     |

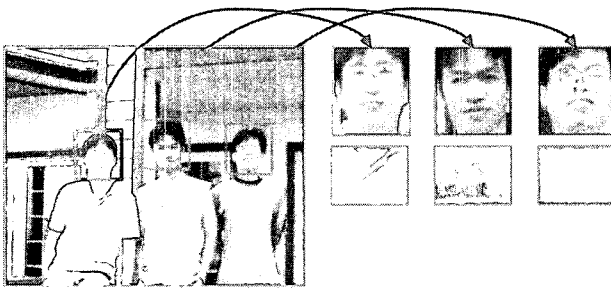


그림 5. 사진으로부터 추출한 얼굴 및 옷 영역의 예

Fig. 5. Example of face and clothes regions extracted from photo

한  $\Delta u$ 는 16으로 고정하였는데, 이 값은 실험에 의해 얻어

진 경험적 수치로서, 얼굴 영역을  $w_f = 46$ ,  $h_f = 56$ 으로 정규화하였을 때, 인물의 턱과 옷 사이에 존재하는 목의 크기가 16 픽셀 정도에 해당하였다. 인물의 옷 영역의 크기인  $w_c$ 와  $h_c$ 는 변화시켜가며 실험을 수행하였다. 그림 5는 사진으로부터 추출된 얼굴 및 옷 영역의 예를 나타낸다. 그림 5의 사진에는 3명의 인물이 포함되어 있으며, 각 인물들로부터 추출된 얼굴 영역 및 해당 인물의 옷 영역을 나타낸다.

상황 기반 클러스터링은 사진의 Exif로부터 추출한 시간 특징값과 5 종류의 MPEG-7 기술자를 이용하여 추출한 비주열 특징값을 이용하였다. 사용된 MPEG-7 기술자는



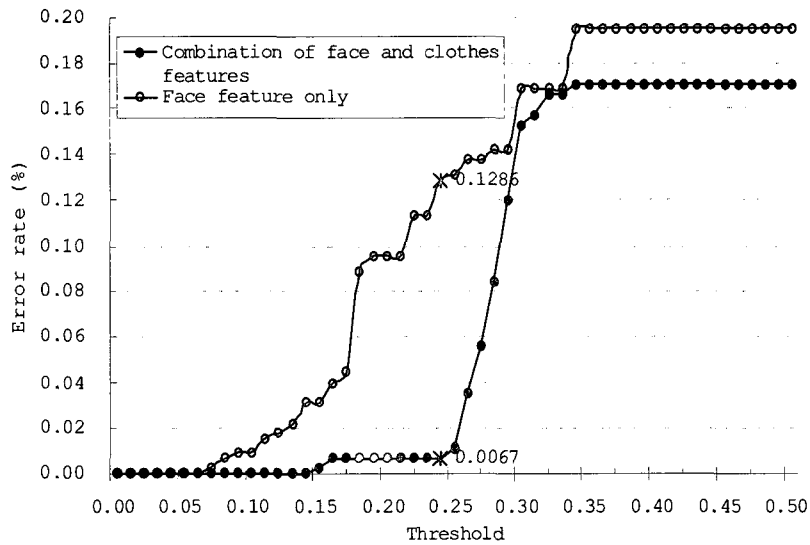


그림 6. 임계값 변화에 따른 인물 기반 사진 클러스터링의 에러율  
 Fig. 6. Error rate according to different threshold in the person-based photo clustering

Color Structure, Color Layout, Scalable Color, Edge Histogram, Homogeneous Texture 기술자이다. 각 특징값의 가중치는 0.2로 동등하게 설정하였다. 상황 변화를 검출하기 위한 임계값은 실험적으로 결정되었으며 0.35로 설정하였다.

451 장의 테스트 영상은 본 논문에서 제안한 방법에 의해 모두 13 개의 상황 클러스터로 나뉘었다. 그림 6은 각 상황 클러스터 내에서 동일 인물들을 agglomerative 클러스터링한 실험 결과를 나타낸다. 그림 6에서, 각 임계값에서의 에러율은 13개의 상황 클러스터 각각에서, 식 (5)에 의해 에러율을 구하고, 모든 상황 클러스터에 대한 평균 에러율을 나타낸다. 군집 클러스터링 방법의 특성상, 군집클러스터링을 멈추는 임계값이 증가할수록 클러스터의 수는 작아지고, 에러율은 증가한다. 임의의 임계값에서 K 개의 인물 클러스터가 존재한다면, 에러율은 다음과 같이 계산한다.

$$\text{error rate} = \frac{1}{K} \cdot \sum_{k=1}^K \left\{ \frac{g(k) - g_{\text{major}}(k)}{g(k)} \right\}, \quad (5)$$

여기서,  $g(k)$ 는 인물 클러스터 k에 속하는 모든 인물의 수를 나타내고,  $g_{\text{major}}(k)$ 는 인물 클러스터 k에서 가장 많은

수를 차지하는 인물의 수를 나타낸다. 본 논문에서는 실험 결과의 클러스터 수가 GT(ground truth) 클러스터의 수인 36에 가장 근사할 때의 임계값을 최종 임계값으로 결정했다. 여기서, GT 클러스터 수는, 상황 기반 클러스터가 주어졌을 때, 에러율이 0%인 경우의 클러스터의 개수를 의미한다. 따라서, 서로 다른 상황 클러스터에 동일한 인물이 있다면, 서로 다른 인물 클러스터를 생성하기 때문에, GT 클러스터의 수가 인물의 수보다 많을 수 있다. 임계값이 약 0.24일 때, 클러스터의 수가 37로 GT 클러스터의 수와 가장 근사하였으며, 이 때의 에러율은 얼굴 특징값만 사용한 경우 약 12.9%, 얼굴 및 옷 특징값을 함께 사용한 경우 약 0.7%를 보였다. 실험 결과에서와 같이, 얼굴 특징값만 사용한 경우에 비해 얼굴과 옷 특징값을 함께 조합하여 사용한 경우, 동일 클러스터 수에서 에러율이 작음을 보였다. 따라서, 인물의 옷 정보가 서로 다른 인물을 식별하는데 도움을 준다고 볼 수 있다.

다음으로, 각 특징값의 가중치 값을 결정하기 위해, 가중치 값의 변화에 따른 색인 성능을 비교하였다. 그림 7은 가중치 값의 변화에 따른 인물 클러스터링 단계에서의 에러율과, 해당 가중치값으로 인물 클러스터링을 수행한 이후 인물 색인을 수행했을 때의 각 인물의 색인 성능의 평균치

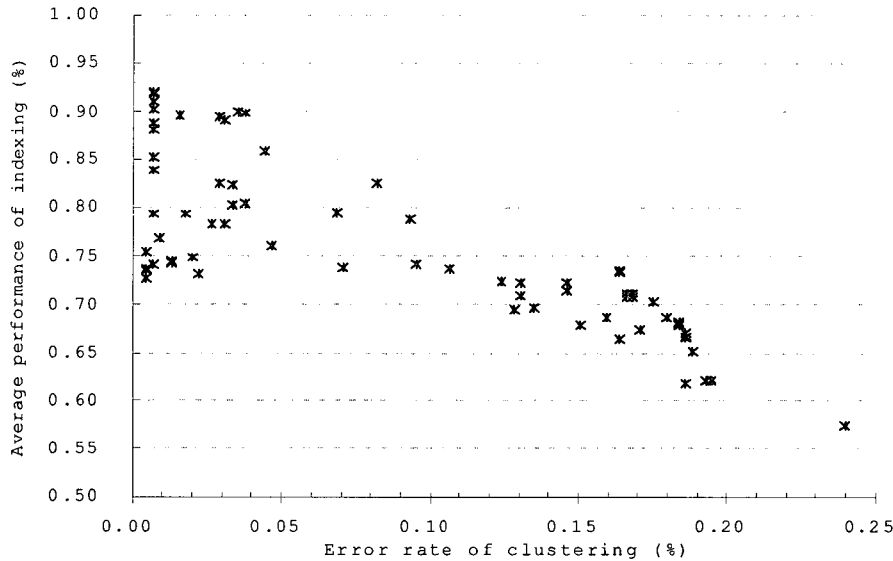


그림 7. 가중치 값 변화에 따른 색인 성능 (여기서, 인물 기반 클러스터링을 위한 임계값은 0.24로 고정하였고, 얼굴 및 옷 특징값을 위한 가중치 값인  $w_{face}$ ,  $w_{color}$ ,  $w_{texture}$ 는 가중치 합이 1.0이 되도록 각각의 가중치 값은 0.0에서 1.0까지 변한다.)

Fig. 7. Indexing performance with different weighting values (where the clustering threshold is fixed to 0.24 and the weighting values for face and clothes ( $w_{face}$ ,  $w_{color}$ , and  $w_{texture}$ ) are set to a value from 0.0 to 1.0 as keeping  $w_{face} + w_{color} + w_{texture} = 1.0$ )

를 나타낸다. 색인 성능을 평가하기 위한 기준은 0.5 (recall + precision)을 이용하였으며, recall과 precision은 다음과 같이 정의 된다.

$$\text{recall} = \sum_v \frac{G_v}{R_v} \text{ and } \text{precision} = \sum_v \frac{G_v}{P_v}, \quad (6)$$

여기서,  $V$ 는 색인할 인물의 수를 나타내며, 본 실험에서는 27명의 인물을 대상으로 하고 있으므로 27이 된다.  $R_v$ 는 테스트 영상들 가운데 인물  $v$ 에 해당하는 true 영상의 수이고,  $G_v$ 는 인물  $v$ 에 해당하는 true 영상들 가운데 인물  $v$ 로 정확히 색인된 positive 영상의 수이며,  $P_v$ 는 테스트 영상들 가운데 인물  $v$ 로 색인된 영상의 수이다. 따라서,  $G_v/R_v$ 는 인물  $v$ 의 recall을 나타내고,  $G_v/P_v$ 는 인물  $v$ 의 precision을 나타낸다. 그림 7에서 인물 색인 성능이 가장 높은 경우는,  $w_{face}$ ,  $w_{color}$ ,  $w_{texture}$ 의 값이 각각 0.59, 0.23, 0.18인 경우였으며, 이때의 평균 색인 성능은 약 92% 정도를 나타냈다. 즉, 얼굴 특징값과 옷 특징값의 중요도의 비율이 약 6:4 정도일 때 가장 높은 색인 성능을 보였다.

다음으로, 상황 클러스터링 및 인물 클러스터링을 수행하지 않고 얼굴 특징값만을 이용하여 인물 색인을 수행했을 때의 성능을 비교 실험하였다. 표 2에서 첫 번째 결과는 상황 클러스터링 및 인물 클러스터링을 수행하지 않고 얼굴 특징값만을 이용하여 색인을 수행했을 때의 결과를 보인다. 각 인물의 SVM 트레이닝 시, 선형 커널을 이용했을 때, recall은 약 53%였고, precision은 약 57%였다. 이에 반하여, Gaussian RBF 커널을 이용했을 때, recall이 약 71%로 향상되었고, precision이 약 73%로 향상되었다. 이는 Gaussian RBF 커널이 가지는 결정 함수(decision function)가 선형 결정 함수에 비해 1:N 식별 능력이 우수한 특징을 가지기 때문으로 분석된다.

표 2에서 두 번째 결과는 상황 클러스터링 및 인물 클러스터링을 수행했을 때의 결과를 보인다. 클러스터링 과정에서 옷 특징값을 이용하지 않고, 얼굴 정보만 이용했을 경우에는 recall이 약 77%로 향상되었으며, precision이 약 78%로 향상되었다. 이 결과는 SVM 커널로 Gaussian RBF 커널을 이용한 결과이다. 다음으로, 클러스터링 과정에서 얼굴 및 옷 특징값을 함께 조합하여 이용한 경우, 옷 영역의 크기에

표 2. 방법에 따른 평균 색인 성능 비교

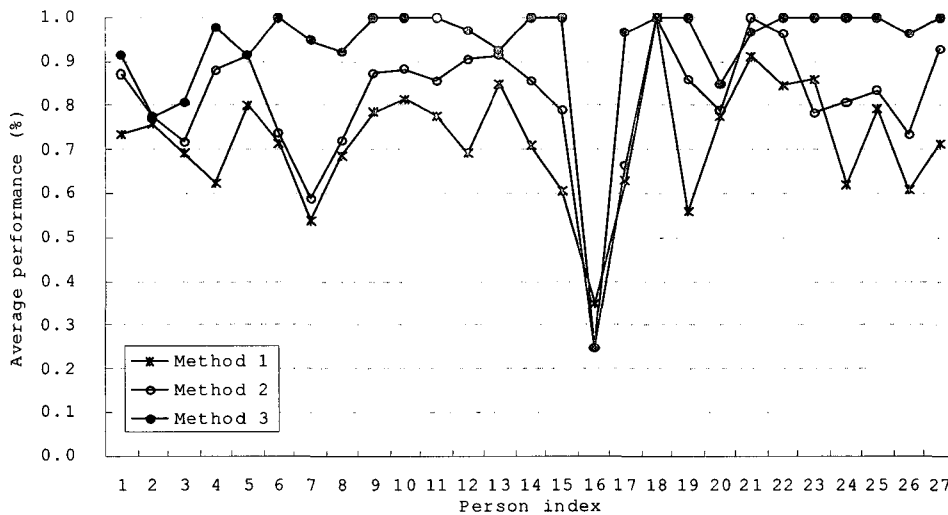
Table 2. Comparison of average indexing performance with different methods

| Method                             |  | Performance              | Recall | Precision |
|------------------------------------|--|--------------------------|--------|-----------|
| Person indexing without clustering | Face feature only  | Linear kernel            | 0.5322 | 0.5659    |
|                                    |  | RBF kernel               | 0.7051 | 0.7340    |
| Person indexing with clustering    | Face feature only  | RBF kernel               | 0.7683 | 0.7829    |
|                                    | Combination of face and clothes features (with different clothes region) | $21(w_c) \times 18(h_c)$ | 0.8141 | 0.8119    |
|                                    |  | $21 \times 21$           | 0.8220 | 0.8160    |
|                                    |  | $24 \times 18$           | 0.8194 | 0.8152    |
|                                    |  | $24 \times 21$           | 0.9287 | 0.9159    |
|                                    |  | $27 \times 21$           | 0.9016 | 0.8903    |
|                                    |  | $24 \times 24$           | 0.8696 | 0.8836    |
|                                    |  | $27 \times 24$           | 0.8937 | 0.9062    |

따라 recall이 약 81% ~ 93%로 향상되었으며, precision이 약 81% ~ 92%로 향상되었다. 옷 영역의 가로 크기( $w_c$ )는 {21, 24, 27}, 옷 영역의 세로 크기( $h_c$ )는 {18, 21, 24}로 변화시키면서 실험을 수행하였다. 실험 결과, 옷 영역의 가로 x 세로가  $24 \times 21$ 인 경우, 성능이 가장 높았으며, 이보다 영역이 작은 경우나 큰 경우에는 성능이 저하됨을 알 수 있었다. 이는 옷 영역의 크기가  $24 \times 21$ 보다 작은 경우, 옷 정보의 특징을 표현하기 어렵다는 점을 나타내고,  $24 \times 21$

보다 큰 경우, 다른 인물이나 주변 객체의 간섭을 받아 옷 정보의 특징을 제대로 표현하기 어렵다는 점을 나타낸다.

그림 8은 인물 별 색인 성능을 세 가지 방법에 따라 비교 도시한 그래프이다. 여기서, 방법 1(method 1)은 상황 기반 클러스터링을 적용하지 않고 얼굴 특징값만을 이용한 경우이고, 방법 2(method 2)는 상황 기반 클러스터링을 적용하고 인물 기반 클러스터링에서 옷 특징값을 이용하지 않고 얼굴 특징값만을 이용한 경우이고, 방법 3(method 3)은 상황 기반



- Method 1: Indexing with only face feature
- Method 2: Indexing with only face feature incorporated with situation clustering
- Method 2: Indexing with face and clothes features incorporated with situation clustering

그림 8. 방법에 따른 인물 별 색인 성능 비교

Fig. 8. Comparison of indexing performance with different methods for each person

클러스터링을 적용하고 인물 기반 클러스터링에서 얼굴 및 옷 특징값을 함께 조합하여 이용한 경우이다. 결과에서, 대부분의 인물들에서 방법 1보다 방법 2가 우수한 성능을 보였으며, 방법 2보다 방법 3이 보다 우수한 성능을 보였다. 그러나, 16번째 인물의 경우, 방법 1이 방법 2와 방법 3보다 우수한 성능을 보였는데, 이는 클러스터링 단계에서, 16번째 인물이 다른 인물의 클러스터에 속하는 에러가 발생되었기 때문으로, 제안된 방법의 단점이라고 할 수 있다. 클러스터의 수를 증가시키면 이러한 형태의 에러를 감소시킬 수 있으나, 클러스터의 수를 너무 많이 증가시키면 얼굴 특징값만을 이용하는 색인 방법과 같아지기 때문에, 색인 성능이 감소한다.

#### IV. 결 론

본 논문에서는 대용량의 디지털 사진들을 특정한 사람에 관계된 사진들로 브라우징하는 것을 돕기 위한 새로운 인물 기반 사진 색인 방법을 제안하였다. 제안된 방법은 인물의 주변 정보인 옷 특징값을 얼굴 특징값과 함께 조합하여 색인 과정에 활용하는 방법으로, 얼굴 특징값과 옷 특징값을 효과적으로 조합하기 위해, 인물을 상황 별로 클러스터링한다. 실험 결과에서, 기존의 얼굴 특징값만을 이용한 경우 약 70% 정도의 색인 성능을 보였으며, 제안된 방법을 이용한 경우 약 92% 정도의 색인 성능을 보여 제안된 방법이 인물 별 사진 색인에 효과적임을 확인할 수 있었다. 향후, 옷 특징값 외에 다른 부가 정보를 인물 색인에 활용하는 방법에 대한 연구가 필요하며, 인물이 성별이나 나이 등을 고려한 색인 방법에 대한 연구도 계속되어야 할 것이다. 또한, 색인할 인물 대상의 범위가 커지거나, 혼란되지 않은

인물이 색인할 대상에 포함되어 있는 경우를 고려한 알고리즘의 개발도 필요하다.

#### 참 고 문 헌

- [1] D. Frohlich, A. Kuchinsky, C. Pering, A. Don, and S. Ariss, "Requirement for photoware," Proc. of the 2002 ACM Conf. on Computer Supported Cooperative Work, pp.166-175, 2002
- [2] Kerry Rodden and Kenneth R. Wood, "How do people manage their digital photographs," Proc. of the SIGCHI Conf. on Human Factors in Computing Systems, pp.409-416, 2003
- [3] M.H. Yang, D. Kriegman, and N. Ahuja, "Detection Faces in Images: A Survey," IEEE Trans. on PAMI, pp.34-58, 2002
- [4] W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey," ACM Computing Surveys, pp.399-458, 2003
- [5] L. Ahang, L. Chen, M. Li, and H. Zhang, "Automated Annotation of Human Faces in Family Albums," Proc. of the 11th ACM Conf. on Multimedia, pp.355-358, 2003
- [6] "Exchangeable image file format for digital still cameras: Exif Version 2.2 (JEITA CP-3451)," Standard of Japan Electronics and Information Technology Industries Association, 2002
- [7] R.O. Duda, P.E. Hart, D.G. Stork, "Pattern Classification," John Wiley & Sons, LTD, 2001
- [8] V.N. Vapnik, "The Nature of Statistical Learning Theory," Springer, 1999
- [9] M. Bober, S.K. Bober, "Description of MPEG-7 Visual Core Experiments," MPEG-7 Visual Group ISO/IEC JTC1/SC29/WG11, N6905, 2005
- [10] T. Kamei, A. Yamada, H. Kim, T.K. Kim, W. Hwang, S.C. Kee, "Advanced Face Descriptor Using Fourier and Intensity LDA Features," ISO/IEC JTC1/SC29/WG11, M8998, 2002
- [11] B.S. Manjunath, P. Salembier, and T. Sikora, "Introduction to MPEG-7," John Wiley and Sons, LTD, 2002
- [12] B. Scholkopf and A. Smola, "Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond," MIT Press. 2002

저 자 소 개



양 승 지

- 2001년 2월 : 강원대학교 정보통신공학과 학사
- 2002년 8월 : 한국정보통신대학교 공학부 영상비디오시스템연구실 석사
- 2002년 9월~현재: 한국정보통신대학교 공학부 영상비디오시스템연구실 박사과정
- 주관심분야 : 이미지/비디오 처리/분석, 색상 처리, 내용 기반 영상 검색/분류, MPEG-7/21



서 경 석

- 2004년 2월 : 광운대학교 전자공학과 학사
- 2005년 2월 : 한국정보통신대학교 공학부 영상비디오시스템연구실 석사
- 주관심분야 : 이미지/비디오 처리 및 분석, 내용 기반 영상 검색/분류, MPEG-7



노 용 만

- 1985년 2월 : 연세대학교 전자공학과 학사
- 1987년 2월 : KAIST 전기공학과 석사
- 1992년 2월 : KAIST 전기공학과 박사
- 1992년~1995년 : Dept. of Radiological Science, University of California, Irvine, 초빙 연구원
- 1996년 : Dept. of Electrical Eng. and Computer Science, University of California, Berkeley 연구원
- 1997년~현재 : 한국정보통신대학교 정교수
- 주관심분야 : 이미지/비디오 처리 및 분석, MPEG-7, 특징인식, 이미지/비디오 인덱싱



김 상 군

- 1991년 : BS, Dept. of Computer Science, Univ. of Iowa
- 1994년 : MS, Dept. of Computer Science, Univ. of Iowa
- 1997년 : Ph.D, Dept. of Computer Science, Univ. of Iowa
- 1997년~현재 : Senior Research Staff, Computing Lab., Digital Research Center, Samsung Advanced Institute of Technology (SAIT)
- 주관심분야 : Image/Video/Audio Analysis and Management, Fast Image Search and Indexing, MPEG-7, Multi-modal Analysis