

일반논문-04-09-2-07

방송 비디오 등장인물 자동 분석 시스템

김기남^{a)}, 김형준^{a)}, 김회율^{a)*}

Automatic Cast-list Analysis System in Broadcasting Videos

Ki-Nam Kim^{a)}, Hyoung-Joon Kim^{a)} and Whoi-Yul Kim^{a)*}

요 약

본 논문에서는 등장인물 검출 및 인식과 함께 등장인물의 출연 구간 분석이 가능한 시스템을 제안한다. 드라마, 스포츠와 같은 방송 비디오는 그 특성상 인물이 중심이 되며 각 시점에 등장하는 주요 인물은 방송용 비디오의 중요한 특징이 된다. 본 논문에서는 비디오에서 등장하는 주요 인물을 자동으로 분석하는 ACAV(Automatic Cast-list Analysis in Videos) 시스템을 제안한다. ACAV 시스템은 등장인물을 자동 검출하여 인물 DB에 등록하는 FAGIS(FAce reGIStation)와 생성된 인물 DB을 이용하여 등장인물을 분석하는 FACOG(FAce reCOGnition)로 구성된다. 기존의 상용화된 등장인물 분석 시스템인 FaceIt과의 성능 비교를 통해 ACAV의 성능을 검증하였다. 얼굴 검출 실험에서 ACAV의 얼굴 검출률은 84.3%로 FaceIt 보다 약 30% 높았고, 얼굴 인식 실험에서도 ACAV의 얼굴 인식률은 75.7%로 FaceIt 보다 27.5% 높은 성능을 보였다. ACAV 시스템은 방송 멀티미디어 공급자를 위한 대용량 비디오 관리 시스템으로 이용될 수 있으며 일반 사용자를 대상으로 한 PVR(Personal Video Recorder), 모바일 폰 등의 비디오 관리 시스템으로도 이용될 수 있다.

Abstract

In this paper, we propose a system that can analyze appearance interval of casts by detecting and recognizing casts in broadcasting videos. The cast is one of the most important characteristics in broadcasting videos such as drama and sports. In this paper, we propose the ACAV(Automatic Cast-list Analysis in Videos) system that analyzes cast-list automatically in video. The ACAV system consists of FAGIS(FAce reGIStation) which registers detected faces into the face DB and FACOG(FAce reCOGnition) that analyzes the cast-list in video sequence using the face DB. We evaluate performance of the ACAV system by comparing with FaceIt, one of the most well-known commercial systems for the cast-list analysis. The ACAV shows face detection and recognition rates of 84.3% and 75.7% that are about 30% and 27.5% higher than those of FaceIt, respectively. The ACAV system can be applied to mass broadcasting videos management system for broadcasters and video management system of PVR(Personal Video Recorder) and mobile phone for the public.

Keywords : Cast-list analysis, video analysis, face detection, face recognition

I. 서 론

최근 인터넷 등 네트워크의 발달로 언제 어디서나 자신

이 원하는 정보에 접근할 수 있게 되었고 이를 통해 사용자에게 다양한 멀티미디어 정보가 제공되고 있다. 이러한 환경 속에서 멀티미디어 정보의 양은 폭발적으로 증가하고 있으며 이에 따라 자신이 원하는 정보를 효율적으로 검색하고 관리하는 기술이 요구되고 있다. 특히 방송 멀티미디어 공급자 측면에서도 나날이 누적되는 대용량 비디오 정보를 관리하고 활용하는 기술 개발이 시급한 실정

a) 한양대학교 전자통신전파공학과

Division of Electrical and Computer Engineering, Hanyang University

* 이 논문은 한양대학교 과학기술분야 특별교비 장학사업과 KBS

기술연구소 측의 지원을 받아 연구되었습니다.

이다.

드라마, 스포츠와 같은 방송 비디오는 그 특성상 인물이 중심이 되며 각 시점에 등장하는 주요 인물은 방송용 비디오의 중요한 특징이 된다. 따라서 방송용 비디오의 중요한 특징인 등장인물을 분석하여 효율적인 비디오 관리 시스템을 개발할 수 있다. 등장인물 분석을 통한 비디오 관리 시스템은 방송 멀티미디어 공급자를 위한 대용량 비디오 관리 시스템으로 이용될 수 있으며 일반 사용자를 대상으로 한 PVR(Personal Video Recorder), 모바일 폰 등의 비디오 관리 시스템으로도 이용될 수 있다. 등장인물 정보를 이용한 비디오 관리 시스템 개발을 위해서는 정확한 등장인물 분석이 선행되어야 한다. 이를 위해 효율적이고 정확한 얼굴 검출 및 인식 시스템 개발이 필수적이다.

방송 비디오 관리를 위한 대표적인 프로그램으로 Virage사에서 개발된 VideoLogger를 들 수 있다. VideoLogger는 비디오 분석을 위한 여러 플러그 인을 지원하며, 특히 등장인물 분석을 위한 플러그 인으로서 FaceIt을 제공한다^[1]. FaceIt은 Visionics사에서 개발된 것으로, 얼굴 등록을 위한 외부 프로그램인 SmartFaces와 등록된 얼굴들을 중심으로 등장인물을 분석하는 플러그 인 두 부분으로 크게 나눌 수 있다^[2]. SmartFaces는 사용자로부터 입력 받은 비디오에서 얼굴을 검출하고 등록하는 기능으로

이루어져 있다. VideoLogger 플러그 인은 SmartFaces로부터 생성된 인물 DB를 이용하여 입력 비디오에서 등장인물을 분석한다. 그러나 SmartFaces는 사용자가 등록을 원하는 인물이 출연하는 시점에서 특정 동작을 수행해야 하는 등의 기능적인 문제점을 가지고 있으며 VideoLogger 플러그 인의 등장인물 분석 또한 만족할 만한 결과를 보여주지 않는다. 이러한 문제점 이외에도 VideoLogger 플러그 인은 등장인물의 출연 구간 분석 기능을 제공하지 못하며 단순히 등장인물의 출연 시간만을 보여 주는 문제점을 가지고 있다.

본 논문에서는 등장인물 검출 및 인식과 함께 등장인물의 출연 구간 분석이 가능한 시스템을 제안한다. 제안된 ACAV(Automatic Cast-list Analysis in Videos) 시스템은 그림 1과 같이 등장인물 등록 프로그램인 FAGIS(FACE reGISTRATION)와 등장인물 분석 프로그램인 FACOG(FACE reCOGNITION)로 구성된다. 입력 비디오에서 등장인물을 등록하고 실시간으로 등장인물을 분석하기 위해서는 효율적인 얼굴 검출 및 인식 알고리즘이 필요하다. FAGIS와 FACOG에서는 얼굴 검출을 위해 연산량이 적은 얼굴 후보 영역 검출 알고리즘과 판별 능력이 우수한 SVM을 함께 사용하였다^[7]. 또한 정확한 얼굴 인식을 위해 기저 벡터가 일정하고 특징값 추 출이 빠른

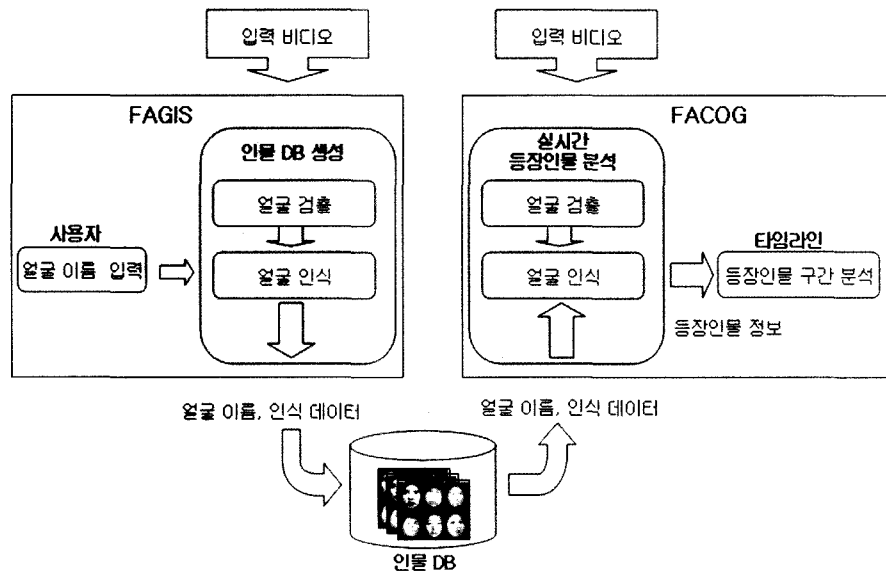


그림 1. ACAV 시스템 개요
Fig. 1. The overall structure of ACAV system

얼굴 구성요소 기반 DCT(Discrete Cosine Transform) /LDA(Linear Discriminant Analysis)를 사용하였다^[8].

FAGIS는 얼굴 검출 기능과 검출된 얼굴을 인물 DB에 등록하는 기능을 갖는다. FAGIS는 입력 비디오를 재생하면서 얼굴을 검출하고 검출된 얼굴이 이미 인물 DB에 등록되어 있으면 등록된 이름을 표시하고, 그렇지 않으면 '미등록'을 표시한다. 검출 결과가 잘못된 경우 사용자는 수동으로 수정하여 정정된 정보를 인물 DB에 등록할 수 있다. FACOG에서는 FAGIS를 통해 구축된 인물 DB를 이용하여 등장인물을 검출하고 등장구간을 분석하여 타임라인에 보여준다. FAGIS는 입력 비디오에서 사용자의 특별한 동작 없이도 자동으로 얼굴을 검출하며 사용자는 검출된 결과들을 일괄적으로 수정 확인 할 수 있다. 따라서 대용량의 방송 비디오인 경우 등록 작업 효율이 높아진다. FACOG는 등장인물의 출연 구간을 분석하여 타임라인으로 출력한다. 타임라인에는 등록된 인물의 얼굴 영상과 등장구간이 그래프로 표시되며 사용자는 이를 이용하여 효율적으로 등장인물을 분석할 수 있다.

본 논문의 구성은 다음과 같다. 2장과 3장에서는 ACAV 시스템에 사용된 얼굴 검출 알고리즘과 얼굴 인식 알고리즘에 대해 기술하였고, 4장에서 ACAV 시스템의 기능에 대해 설명하였다. 5장에서는 ACAV 시스템의 얼굴 검출 및 인식 성능을 FaceIt과 비교 분석하였고, 6장에서 결론을 맺었다.

II. 얼굴 검출 알고리즘

본 논문에서는 얼굴 검출을 위해 연산량이 적은 얼굴 후보 영역 검출 알고리즘과 판별 능력이 우수한 SVM^[3]을 함께 사용하였다. 얼굴 검출 알고리즘은 그림 2와 같이 얼굴 후보 영역 검출 단계와 얼굴 검증 단계로 나눌 수 있다.

얼굴 검출을 위해 입력 비디오의 프레임영상을 24×24 크기의 윈도우로 스캔한다. 이때 다양한 크기의 얼굴을 검출하기 위해 프레임 영상을 1.2 배씩 줄이면서 스캔하는 방법을 사용하였다. 24×24 크기의 윈도우 영상은 얼굴 검출을 위한 입력 영상으로 사용되며, 얼굴 후보 영역 검출 단계와 얼굴 검증 단계를 통해 실제 얼굴 영역으로 판단된다.

1. 얼굴 후보 영역 검출

본 논문에서는 살색 정보, Harr feature, Hausdorff distance를 사용하여 얼굴 후보 영역을 검출하였다. 입력 영상의 색상 정보를 이용하여 살색 영역을 추출하고 추출된 영역을 얼굴 후보 영역으로 판단하였다. 입력 영상의 살색 영역 추출을 위해서는 살색 분포 모델이 필요하다. 이를 위해 임의의 영상에서 살색 영역만을 수동으로 분리하고 이 영역의 색상을 살색 영역이 잘 분리되는 YUV 색상 공간으로 변환하였다^{[4][6]}. 변환된 YUV 성분 중 인종에 따라 급격히 변하는 Y 성분을 제외한 UV 성분만으로 2차원 살색

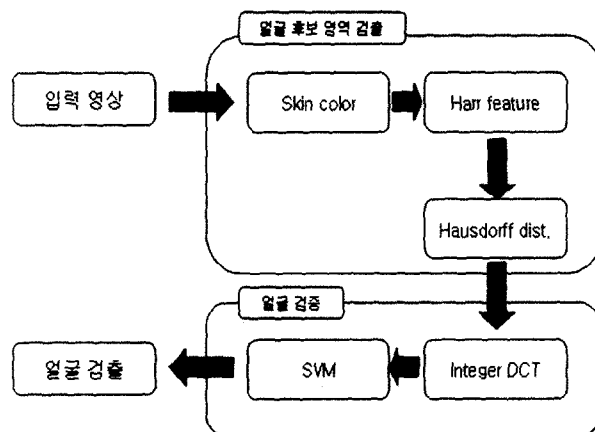


그림 2. 얼굴 검출 알고리즘의 전체 구성
Fig. 2. The overall structure of face detection algorithm

분포 모델을 생성한다^[5]. 입력 영상의 살색 영역 추출을 위해서 입력 영상을 YUV 색상 공간으로 변환한다. 변환된 색상의 UV 성분과 생성된 살색 분포 모델을 비교하여 살색 분포 모델에 해당되는 픽셀 수가 전체 픽셀 수의 80% 이상이 되면 입력 영상을 얼굴 후보 영역으로 판단한다. 그러나 살색 분포 모델을 이용한 얼굴 후보 영역 검출 방법은 목이나 팔과 같이 얼굴과 유사한 살색 분포를 가지는 영역도 얼굴 영역으로 판단한다. 따라서 얼굴 특징을 잘 나타내는 Harr feature를 이용하여 목과 팔과 같이 얼굴이 아닌 영역을 배제시킨다.

얼굴은 그 특성상 눈 근처의 영역은 비교적 그레이 레벨이 낮으며 이외의 영역은 비교적 그레이 레벨이 높다. Harr feature는 이러한 얼굴의 특징을 표현한 방법으로 한 영역의 모든 픽셀 값의 합과 다른 영역의 픽셀 값을 합한 것의 차로 표현된다^[9]. 그림 3과 같이 일반적으로 얼굴의 특징을 잘 나타내는 feature는 여러 가지가 있을 수 있다. 생성 가능한 feature들 중, 본 논문에서는 실험적으로 얻어진 그림 4와 같은 4개의 feature를 사용해서 얼굴 후보 영역을 검출하였다.

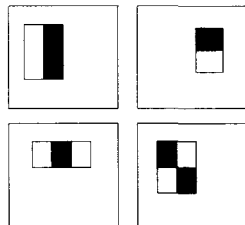


그림 3. Harr features
Fig. 3. Harr features

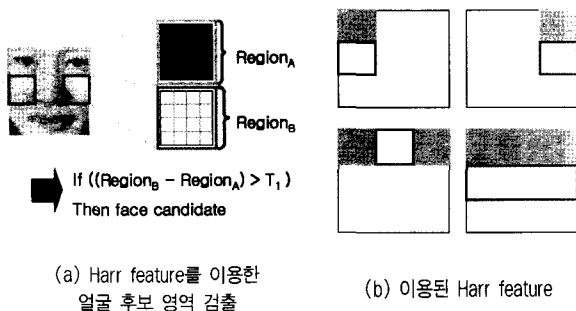


그림 4. Harr feature를 이용한 얼굴 후보 영역 검출
Fig. 4. Detection of face candidate regions using Harr features:
(a) face candidate detection using Harr feature, (b) used Harr features

얼굴 후보 영역 검출을 위해 사용된 feature들은 얼굴 회전에 영향을 덜 받고 눈썹 영역까지 포함되도록 일반적인 Harr feature에 비해 크게 구성되었다. 영역 B(RegionB)와 영역 A(RegionA)와의 차이가 임계치 이상인 영역은 Hausdorff distance를 통해 최종 얼굴 후보 영역으로 판단된다.

Hausdorff distance는 객체 윤곽선 모양의 유사도를 나타낸다^[10]. Hausdorff distance를 계산하기 위해서는 얼굴 윤곽선 모델이 필요하다. 이를 위해 MPEG-7 얼굴 인식 기술자 실험^[13]에 사용되었던 데이터 셋 중 Altkom 데이터 셋과 MPEG 데이터 셋의 정면 얼굴 715장을 평균하였다. 평균 얼굴 영상이 구해지면 이 평균 영상의 에지 영상을 구한다. 이렇게 구해진 평균 얼굴 영상의 에지 영상을 얼굴 영상을 대표하는 얼굴 윤곽선 모델로 사용하였다. 얼굴 윤곽선 모델을 객체 A, 입력 영상의 에지 영상을 객체 B라고 한다면 두 객체 A, B 사이의 Hausdorff distance는 식 (1)과 같이 구할 수 있다. 이를 위해 식 (2)를 이용하여 객체 A에서 바라본 A와 B의 거리 $h(A, B)$ 를 구하고 같은 방법으로 B에서 바라본 A와 B의 거리 $h(B, A)$ 를 구한다. 구해진 $h(A, B)$ 와 $h(B, A)$ 중 최대값이 얼굴 윤곽선 모델과 입력 영상의 Hausdorff distance가 된다. Hausdorff distance가 임계값 이상이 되면 이 영역은 최종 얼굴 후보 영역으로 판단된다.

$$H(A, B) = \max\{h(A, B), h(B, A)\} \tag{1}$$

$$h(A, B) = \frac{1}{|A|} \sum_{a \in A} \min_{b \in B} \|a - b\| \tag{2}$$

2. 얼굴 검증

검출된 후보 영역에 대해 마지막으로 SVM을 이용하여 실제 얼굴 영역인지 검증한다. 이때 SVM의 계산시간을 단축시키기 위해서 Integer DCT(Discrete Cosine Transformation)를 이용하였고 SVM에 입력되는 특징 벡터의 차원을 효율적으로 줄였다^[7]. Integer DCT와 SVM을 이용한 얼굴 검증 알고리즘은 그림 5와 같다.

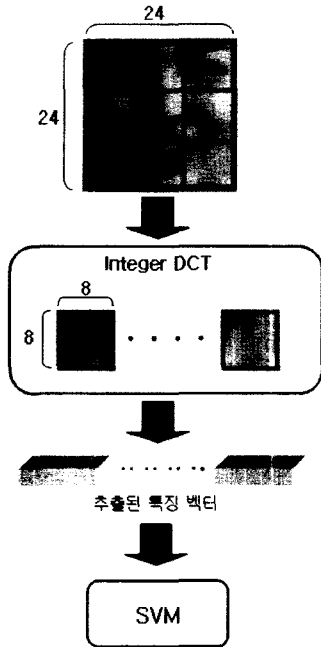


그림 5. 얼굴 검증 알고리즘
Fig. 5. Face verification algorithm

얼굴 검증을 위해 얼굴 후보 영역을 8×8 크기의 블록으로 분할한다. 분할된 9개의 블록에 대하여 Integer DCT를 수행하고 저주파 DCT 계수 12개를 추출한다. 이때 DC성분을 제외하여 조명 성분의 영향을 줄인다. 블록마다 추출된 12개의 DCT 계수들을 이용해서 108 차원의 특징 벡터를 생성한다. 생성된 특징 벡터는 SVM의 입력으로서 사용되며 SVM을 통해 얼굴 후보 영역은 실제 얼굴 영역으로 판단된다.

III. 얼굴 인식 알고리즘

얼굴 인식 성능 향상을 위해서는 검출된 얼굴 영상을 얼굴 인식에 적합한 위치로 변환하는 얼굴 정규화 과정이 필요하다. 따라서 본 논문에서는 SVM을 이용하여 두 눈의 위치를 검출하고 이를 정해진 위치로 변환하여 정규화를 수행하였다. 또한 정확한 얼굴 인식을 위해 기저 벡터가 일정하고 특징값 추출이 빠른 얼굴 구성요소 기반 DCT (Discrete Cosine Transform)/LDA(Linear Discriminant Analysis)를 사용하였다.

1. 얼굴 정규화 알고리즘

얼굴 정규화를 위해서는 눈, 코, 입과 같은 얼굴 특징점이 정확히 검출되어야 한다. 본 논문에서는 SVM을 이용하여 두 눈의 위치를 검출하고 이를 정해진 위치로 변환하여 얼굴 정규화를 수행하였다. 얼굴 정규화를 위한 전체 알고리즘은 그림 6과 같다.

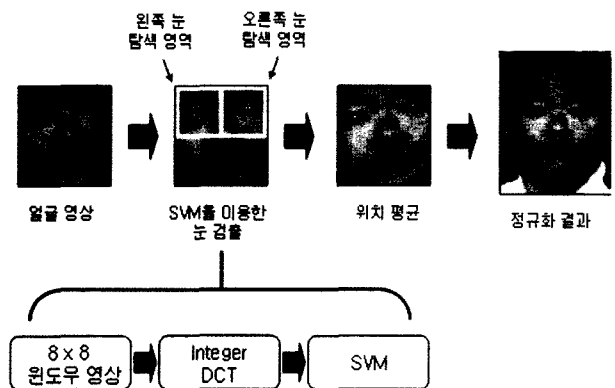


그림 6. 눈 위치 정보를 이용한 얼굴 정규화
Fig. 6. Face alignment using eye position

얼굴 정규화를 위해 입력 영상에서 SVM을 이용하여 두 눈의 위치를 검출한다. 이때 효율적인 눈 검출을 위해 왼쪽 눈을 검출할 경우에는 왼쪽 위 1/4 영역만을, 오른쪽 눈을 검출할 경우에는 오른쪽 위 1/4 영역만을 탐색하였으며 얼굴을 검증할 때와 마찬가지로 Integer DCT와 SVM을 사용하였다^[7]. SVM을 이용하여 검출된 결과는 눈 주변에 여러 점으로 나타나므로 검출된 눈 위치들의 평균을 통해 눈 위치를 결정하고 결정된 눈 위치를 미리 정해진 위치로 변환하여 얼굴 정규화를 수행한다.

2. 얼굴 인식 알고리즘

정확한 얼굴 인식을 위해 기저 벡터가 일정하고 특징값 추출이 빠른 DCT/LDA를 사용하였다^[8]. 또한 얼굴의 국부적인 특징을 이용하기 위해 얼굴 구성요소 별로 DCT/LDA 특징값을 추출하였다^[11]. 본 논문에서는 전체 얼굴 영상과 왼쪽 눈 영역, 오른쪽 눈 영역, 코 영역을 얼굴 구성요소로 사용하였다. 각 얼굴 구성요소는 그림 7과 같이 정의하였다.

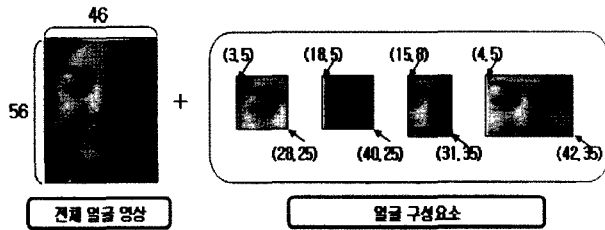


그림 7. 얼굴 구성요소 추출 영역
Fig. 7. The regions of facial components

얼굴 구성요소들 중 입은 다른 얼굴 구성요소에 비해 표정 변화에 민감하므로 제외하였다. 이렇게 정의된 얼굴 구성요소를 기반으로 한 DCT/LDA의 특징값 추출과정은 그림 8과 같다.

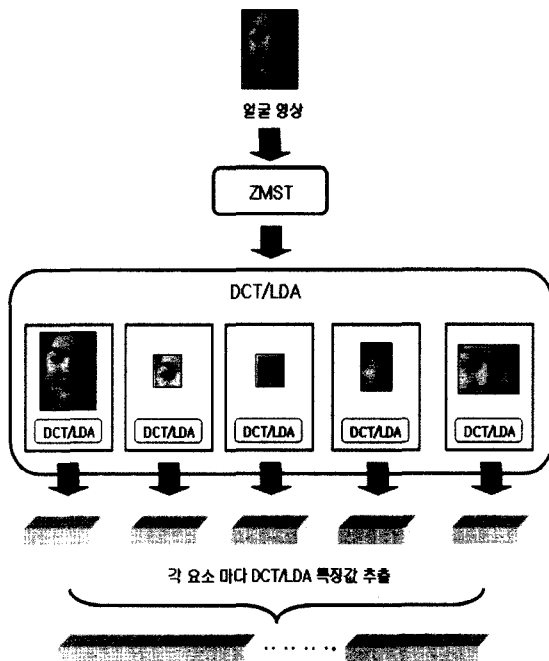


그림 8. 얼굴 구성요소 기반 DCT/LDA 특징값 추출 과정
Fig. 8. Feature extraction using component based DCT/LDA

DCT/LDA를 수행하기 전 조명 성분의 영향을 줄이기 위해 ZMST(Zero Mean Unit Variance)를 수행한다^[12]. ZMST는 영상의 밝기 성분에 평균을 빼고 표준편차로 나누어 밝기 성분을 정규화한다. ZMST가 끝나면 전체 얼굴 영상 및 왼쪽 눈, 오른쪽 눈, 코 영역에 대해 DCT 계수를 추출한다. 이렇게 추출된 DCT 계수를 가지고 LDA를

학습하여 LDA의 기저 벡터를 추출한다. 추출된 기저 벡터를 통해 얼굴 구성요소들의 DCT/LDA 특징값을 추출하고 이를 이용하여 최종 특징값을 만든다.

입력 영상에서 얼굴 구성요소 기반 DCT/LDA 특징값이 추출되면 추출된 특징값과 DB에 등록된 얼굴 영상의 특징값에 대한 유사도를 식 (3)과 같이 계산하여 가장 짧은 거리를 가지는 얼굴을 입력 영상의 얼굴로 인식하게 된다.

$$d(\mathbf{V}_{unknown}, \mathbf{V}_{database}) = \sum_{k=0}^N (\mathbf{v}_{unknown,k} - \mathbf{v}_{database,k})^2$$

$V_{unknown}$: 입력 얼굴 영상의 DCT/LDA 특징값
 $V_{database}$: DB에 등록된 얼굴 영상의 DCT/LDA 특징값
 N : DCT/LDA 특징값 사용 개수

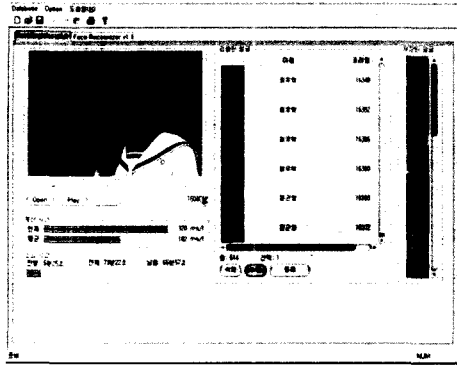
IV. ACAV 시스템

제안된 ACAV 시스템은 그림 9와 같이 등장인물 등록 프로그램인 FAGIS와 등장인물 분석 프로그램인 FACOG로 구성된다.

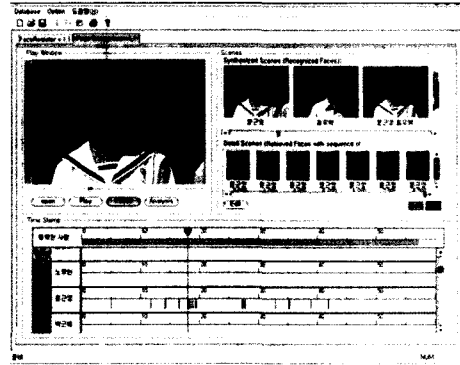
FAGIS는 얼굴 검출 기능과 검출된 얼굴을 인물 DB에 등록하는 기능을 갖는다. FAGIS는 입력 비디오를 재생하면서 얼굴을 검출하고 검출된 얼굴이 이미 인물 DB에 등록되어 있으면 등록된 이름을 표시하고, 그렇지 않으면 '미등록'을 표시한다. 검출 결과가 잘못된 경우 사용자는 수작업으로 수정하여 정정된 데이터를 인물 DB에 등록할 수 있다. FACOG에서는 FAGIS를 통해 구축된 인물 DB를 이용하여 등장인물을 검출하고 등장구간을 분석하여 하단의 타임라인에 보여준다.

이에 비해 이미 상용화 되어 있는 Visionics사의 FaceIt 시스템은 그림 10과 같이 구성되어 있다.

FaceIt은 그림 10과 같이 등장인물 등록 프로그램인 SmartFaces와 등장인물 분석 프로그램인 VideoLogger 플러그 인으로 구성된다^[2]. SmartFaces는 사용자로부터 입력 받은 비디오에서 얼굴을 검출하고 등록하는 기능으로 이루어져 있다. SmartFaces에서 얼굴을 등록하기 위해서는 우선 비디오를 등록하고 싶은 장면까지 재생시킨 후 해당 장면이 나오는 시점에서 사용자가 등록 버튼을 클릭해야 한다. 등록 버튼을 클릭한 후 별도의 등록 창에서 사용자가



(a) 등장인물 등록 프로그램



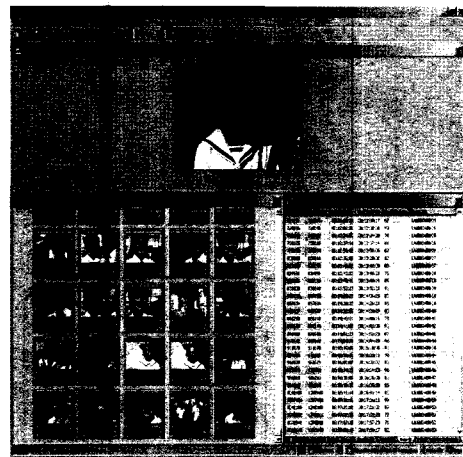
(b) 등장인물 분석 프로그램

그림 9. ACVA 시스템

Fig. 9. ACVA system (a) The characters register program, (b) The characters analysis program



(a) 등장인물 등록 프로그램



(b) 등장인물 분석 프로그램

그림 10. Facelt 시스템

Fig. 10. Facelt system: (a) The characters register program, (b) The characters analysis program

등장인물의 이름을 입력하면 검출된 얼굴은 XML의 형태로 저장된다. 등록된 얼굴을 확인하거나 수정하기 위해서는 별도의 수정 창을 열어야 한다. VideoLogger 플러그인에서는 SmartFaces를 통해 생성된 XML 정보를 이용하여 비디오에서 등장하는 인물을 검출하고 그 등장인물 정보를 분석하여 Text의 형태로 보여준다.

대용량 방송 비디오에서 등장인물을 효율적으로 분석하기 위해서는 작업 효율을 높일 수 있는 GUI가 필수적이며 결과 또한 단순한 Text 형식이 아닌 분석이 용이한 그래프 형식으로 표시되어야 한다. 이러한 측면에서 ACVA 시스템은 다음과 같은 장점을 가진다. FAGIS는

입력 비디오에서 사용자의 특별한 동작 없이도 자동으로 얼굴을 검출하며 사용자는 검출된 결과들을 일괄적으로 수정 확인 할 수 있다. 따라서 대용량의 방송 비디오인 경우 등록 작업 효율이 높아진다. FACOG는 결과를 단순한 Text의 형식뿐만 아니라 타임라인으로도 출력한다. 타임라인에는 등록된 인물의 얼굴 영상과 등장구간이 그래프로 표시되며 사용자는 이를 이용하여 효율적으로 등장인물을 분석할 수 있다. 이러한 기능적인 장점 이외에도 ACVA 시스템은 입력 비디오의 프레임 영상에서 단 하나의 얼굴만 검출하는 FaceIt에 비해 다수의 얼굴을 검출할 수 있는 장점이 있으며 이를 이용하여 정확한 등

장인물 분석을 수행할 수 있다.

V. 실험 결과

제안된 시스템의 성능 검증을 위해 Visionics 사의 FaceIt 과 비교 평가 하였다. 성능 비교를 위한 실험 비디오는 가을동화 6회분과 KBS 9시 뉴스 2003년 11월 29일 분을 사용하였다. 본 논문에서는 정확한 성능 측정을 위해 각 비디오에서 정면 얼굴이 포함된 160 프레임과 얼굴을 포함하지 않은 760 프레임을 수작업으로 추출하였다. 이 프레임들을 이용하여 정면 얼굴들로만 이루어진 FaceVideo와 얼굴을 포함하지 않은 영상들로 이루어진 NonFaceVideo를 제작하였다. 그림 11은 FaceVideo와 NonFaceVideo의 일부 프레임을 보여준다.

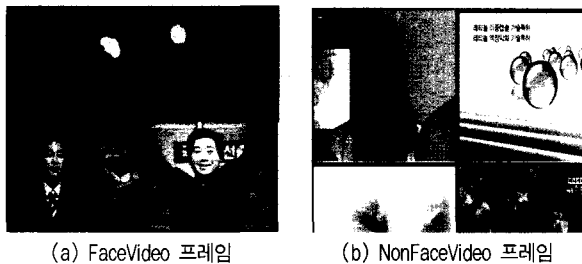
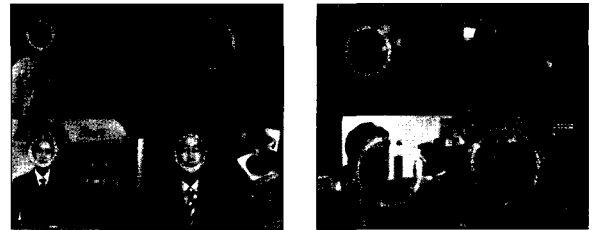


그림 11. 실험 영상
Fig. 11. Test videos: (a) FaceVideo frames, (b) NonFaceVideo frames

등장인물 분석 시스템의 얼굴 검출률은 높아야 하며 동시에 얼굴이 아닌 영역을 얼굴로 판단하는 FA(False Alarm)는 낮아야 한다. 따라서 FaceVideo에서의 얼굴 검출률과 FaceVideo와 NonFaceVideo에서의 FA를 통해 얼굴 검출 성능을 비교하였다. 정확한 측정을 위해 검출 영역에서 실제 얼굴이 차지하는 비율이 1/2 이상이 되고 눈, 코, 입이 모두 포함된 결과만을 얼굴 검출 결과로 판단 하였다. 배경을 얼굴 영역으로 검출한 결과와 검출된 영역에 얼굴의 일부가 포함되었다고 눈, 코, 입이 모두 포함되지 않았거나 얼굴이 차지하는 비율이 1/2 이하인 결과는 FA로 판단하였다. 그림 12는 이와 같은 기준을 사용해서 얼굴 검출로 판단된 결과와 FA로 판단된 결과를 보여준다.



(a) 얼굴 검출의 예 (b) False Alarm의 예

그림 12. FaceIt에서 얼굴 검출과 False Alarm의 예
Fig. 12. The examples of face detection and False Alarm in FaceIt:
(a) The examples of face detection, (b) The examples of False Alarm

가을 동화 6회 분에서 얼굴 검출 성능은 표 1과 같다. 가을동화에서 ACAV의 얼굴 검출률은 84.3%로 FaceIt의 53.8%보다 약 30% 높았고, FA 또한 ACAV가 두 배 이상 낮게 검출 되었다. 표 2는 KBS 뉴스 9시에서 ACAV와 FaceIt의 얼굴 검출 성능을 나타내고 있다. 뉴스 비디오에서도 ACAV의 얼굴 검출률이 73.5%로 FaceIt의 44.2%보다 약 30% 높았고 FA 또한 낮게 검출되었다.

표 1. 가을동화 6회에서 얼굴 검출 성능 비교
Table 1. Comparison of face detection performance in 'Fall Story' episode 6

	얼굴 검출률	False Alarm
FaceIt	53.8%	12개
ACAV	84.3%	5개

표 2. KBS 뉴스 9에서 얼굴 검출 성능 비교
Table 2. Comparison of face detection performance in KBS news 9

	얼굴 검출률	False Alarm
FaceIt	44.2%	17개
ACAV	73.5%	10개

얼굴 인식 성능 평가를 위해 가을 동화 3회의 등장인물을 각 시스템에 등록하였다. 등장인물 등록 프로그램인 FAGIS와 SmartFaces를 사용하여 등록하였으며 총 6명의 등장인물을 4장씩 등록하여 인물 DB를 생성하였다. 생성된 인물 DB를 사용하여 가을 동화 6회 분에서의 얼굴 인식률을 측정하였다. 정확한 얼굴 인식을 위해서 등장인물마다 다양한 표정의 얼굴 영상을 등록하였다. 그림 13은 등록 프로그램을 통해 등록된 얼굴 영상의 예를 보여준다.



그림 13. 등록된 얼굴 영상의 예
Fig. 13. The examples of registered face images: (a) Facelt (b) ACAV

등록된 얼굴 영상을 이용하여 가을 동화 6회 분에서 측정된 Facelt과 ACAV의 얼굴 인식률은 표 3과 같다. 얼굴 인식률은 ACAV가 75.7%로 Facelt의 48.2% 보다 약 30% 높은 성능을 보였다.

표 3. 얼굴 인식 성능 비교
Table 3. Comparison of face recognition performance

실험 영상	Facelt	ACAV
가을 동화 6회	48.2%	75.7%

얼굴 정규화 방법의 차이가 Facelt과 ACAV의 얼굴 인식 성능에 큰 영향을 준 것으로 판단된다. 그림 14와 같이 Facelt은 두 눈의 위치가 영상의 중심 정도에 위치하고 얼굴 부분만이 아닌 헤어스타일까지 반영되도록 얼굴 정규화를 수행한다. 이에 비해 ACAV에서는 두 눈의 위치가 영상의 중심 위 부분에 위치하고 헤어스타일을 최대한 반영되지 않도록 얼굴 정규화를 수행한다. 드라마나 스포츠와 같은 방송 비디오에서 등장인물의 헤어스타일은 그 변화 정도가 매우 심하며 본 실험의 '가을동화'도 매 회마다 그 변화 정도가 심하다. 따라서 헤어스타일을 최대한 배제한 얼굴 정규화 방법이 얼굴 인식 성능 향상에 큰 영향을 미친 것으로 해석할 수 있다.



그림 14. 얼굴 정규화 이미지
Fig. 14. Aligned face images: (a) Facelt, (b) ACAV

VI. 결론

본 논문에서는 등장인물을 자동으로 분석하는 시스템을 제안하였다. 제안된 ACAV 시스템은 등장인물을 검출하여 인물 DB에 등록하는 FAGIS와 생성된 인물 DB을 이용하여 등장인물을 분석하는 FACOG로 구성된다. 상용화된 등장인물 분석 시스템인 FaceIt과의 성능 비교를 통해 ACAV의 성능을 검증하였다. 실험에서 ACAV의 얼굴 검출률과 얼굴 인식률은 모두 FaceIt 보다 약 30% 높게 측정되었고 얼굴 정규화 방법의 차이가 FaceIt과 ACAV의 얼굴 인식 성능에 큰 영향을 준 것으로 판단된다. ACAV 시스템은 방송 멀티미디어 공급자를 위한 대용량 비디오 관리 시스템으로 이용될 수 있으며 일반 사용자를 대상으로 한 PVR, 모바일 폰 등의 비디오 관리 시스템으로도 이용될 수 있다. 향후 연구로는 ACAV의 얼굴 인식 성능을 향상시키기 위한 연구가 필요하다.

참고 문헌

- [1] Virage Inc., "VideoLogger 6.0 getting start guide," Virage Inc., 2003.
- [2] Virage Inc., "Media Analysis Plug-in for Face Recognition 3.6," Virage Inc., 2003.
- [3] E. Osuna, R. Freund and F. Girosi, "Support vector machines: Training and applications," Technical report, MIT, 1997. 3.
- [4] H. Wu, Q. Chen and M. Yachida, "Face Detection From Color Images Using a Fuzzy Pattern Matching Method," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 21, No. 6, pp. 557-563, 1999. 6.
- [5] H. L. Wang and S. F. Chang, "A Highly Efficient System for Automatic Face Region Detection in MPEG Video," IEEE Trans. Circuits and Systems for Video Technology, Vol. 7, No. 4, pp. 615-628, 1997. 8.
- [6] D. Cahi and K. N. Ngan, "Face segmentation using skin color map," IEEE Trans. Circuit and Systems for Video Technology, Vol 9, No 4, pp. 551-564, 1999. 6.
- [7] 박현선, 김경수, 김희정, 정병희, 하명환, 김희율, "Integer DCT와 SVM을 이용한 실시간 얼굴 검출," 대한전자공학회 하계학술대회, Vol. 26, No. 1, pp. 2112-2115, 2003. 7.
- [8] 이훈진, 김형준, 김희정, 하명환, 정병희, 김희율, "DCT/LDA를 이용한 얼굴 인식의 성능 향상," 제16회 신호처리합동학술대회, Vol. 16, No. 1, pp. 854-857, 2003. 9.
- [9] P. Viola, M. Jones, "Rapid object detection using a boosted cascade of simple features," Proc. IEEE Computer Society Conference, Vol. 1, No. 1, pp. 511-518, 2001.

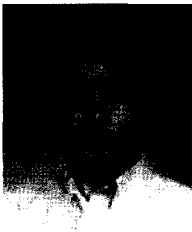
[10] O. Jesorsky, K. J. Kirchberg and R. Frischholz, "Robust Face Detection Using the Hausdorff Distance," Audio and Video-Based Person Authentication, Vol. 2091, pp. 90-95, 2001.

[11] H. Kim, T. Kim, J. Lee, W. Hwang and S. Kee, "Component-based LDA Face Descriptor," ISO/IEC JTC1/SC21/WG11 M8243, 2002. 5.

[12] Y. Li, "Linear Discriminant Analysis and its application to Face Identification," Ph. D. Thesis, University of Surrey, 2000. 9.

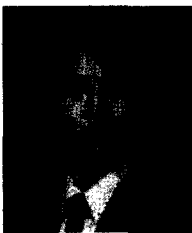
[13] M. Bober, et al., "Description of MPEG-7 Visual Core Experiments," ISO/IEC JTC1/SC29/WG11 N4925, 2002. 10.

저 자 소 개



김 기 남

- 2003년 : 한양대학교 전자전기공학부 졸업 (공학사)
- 2003년 ~ 현재 : 한양대학교 대학원 전자통신전파공학과 석사과정
- 주관심분야 : 패턴 인식, 얼굴 인식 등



김 형 준

- 1999년 : 한양대학교 전자전자통신전파공학과 졸업 (공학사)
- 2001년 : 한양대학교 대학원 전자통신전파공학과 졸업 (공학석사)
- 2003년 ~ 현재 : 한양대학교 대학원 전자통신전파공학과 박사과정
- 주관심분야 : 패턴 인식, 얼굴 인식, 멀티미디어 검색 등



김 희 율

- 1980년 : 한양대학교 전자공학과 졸업 (공학사)
- 1983년 : Pennsylvania State University 전기공학과 졸업 (공학석사)
- 1989년 : Purdue University 전기공학과 졸업 (공학박사)
- 1989년 9월 ~ 1994년 2월 : University of Texas 조교수
- 1994년 ~ 현재 : 한양대학교 전자전기컴퓨터공학부 정교수
- 주관심분야 : 영상처리, 컴퓨터비전, 패턴 인식, 머신비전, MPEG-7 등