

감시 시스템을 위한 얼굴 인식 기반의 검색

이종욱*, 박승진**, 이한성***, 박대희**

*고려대학교 전산학과

**고려대학교 컴퓨터정보학과

***한국전자통신연구원 휴먼인식기술연구팀

e-mail : {eastwest9, valkeylie}@korea.ac.kr, mohan@etri.re.kr, dhpark@korea.ac.kr

A Face Recognition Based Retrieval for Surveillance System

Jong-uk Lee*, Seung-jin Park**, Han-sung Lee***, Dai-hee Park**

*Dept. of Computer and Information Science, Korea University

**Electronics and Telecommunications Research Institute, Korea

요 약

본 논문에서는 CCTV 감시 환경에서 얼굴 이미지를 이용하여 영상에 저장된 범죄 용의자 또는 특정한 시간대에 출입한 사람들을 검색할 수 있는 시스템을 설계 및 구현하였다. 제안된 시스템은 감시 영상을 효율적으로 검색하기 위하여 사람의 얼굴이 나타난 장면을 기반으로 감시 영상을 분할하였으며, 최근 얼굴 인식 분야에서 성공적인 업적을 보여주고 있는 신호 처리 분야의 SRC를 이용하여 얼굴 검색 모듈을 구성하였다. 자체 제작한 KUFDB(Korea University Face Database)와 CCTV 환경의 얼굴 인식 기반 검색 시스템 환경을 캠퍼스 내에서 모의 구축하여 제안된 시스템의 성능을 실험적으로 검증하였다.

1. 서론

지난 2001년 발생한 9.11 테러 사태는 미국을 중심으로 세계 각국이 개인 및 공공의 안전을 위하여 보다 강화된 국토보안(homeland security) 기술의 개발에 많은 투자를 하게 되는 계기가 되었으며, 이와 더불어 CCTV(Close-Circuit Television)를 이용한 영상 감시에 대한 수요가 급속하게 증가하고 있다[1].

이와 같은 특정 지역의 보안 및 감시를 위한 영상 감시 시스템에서 가장 중요한 부분 중 하나는 관리자가 관심 있는 장면을 신속하고 정확하게 찾아주는 영상 검색 요소이다. 영상 검색을 위한 방법들은 이미지 검색 방법과 유사하며 크게 다음과 같은 두 가지 방법을 이용한 검색 질의 방법을 사용한다[2]: 1) 키워드에 의한 질의(Query-by-keyword): 이미지 또는 비디오 데이터에 대한 정보를 키워드 또는 메타데이터(metadata)를 이용하여 주석 처리한 후, 사용자가 관심 있는 키워드를 입력하면 시스템은 메타데이터 정보들을 확인하여 적합한 결과를 출력한다; 2) 예제에 의한 질의(Query-by-example): 이미지 또는 비디오 데이터를 검색에 사용하기 위해서는 그 자체를 표현할 수 있는 적합한 특징 집합으로 변환하여 기술한다. 사용자는 찾고자 하는 정보와 관련한 이미지 또는 비디오를 입력하면 시스템은 질의로 주어진 영상을 해당하는 특징 집합으로 변경하여, 데이터베이스에 저장된 내용 중 유사한 것을 찾아 그 결과를 보여준다[2].

감시 영상 데이터베이스에 저장된 영상을 사람이 직접 확인하는 것은 한계가 있으므로, 빠르고 정확한 범죄 용의자 또는 특정 인물 검색을 지원하기 위해서는 사람이 등장한 장면에 대한 기술이 반드시 필요하다. 저장된 감시 영상에서 사람이 등장했음을 확인할 수 있는 방법은 크게 움직임 탐지[3]와 얼굴 탐지[4]가 있다. 이 중 얼굴 탐지 방법을 통해 획득한 얼굴 이미지는 추후 탐지된 사람의 신원을 식별할 수 있는 중요한 생체 정보로 사용될 수 있다. 또한, 얼굴이 탐지되었을 때 해당 정보를 이용하여 메타데이터를 작성하고 영상을 분할하는 것은 보다 효과적으로 감시 영상 검색 시스템에 사용될 수 있다.

대용량의 이미지 데이터베이스로부터, 예제에 의한 질의의 결과를 보여주기 위한 내용 기반 이미지 검색은 1990년 이래로 멀티미디어 데이터베이스 학회의 중요한 연구 과제 중 하나이다[5]. 특히, 얼굴 이미지 검색은 얼굴 사진 검색, 얼굴 인식 기반의 감시 시스템과 같은 실용적인 응용 분야에서 의미 있는 연구 이슈이며, 얼굴 이미지 검색은 기하학적인 외형이 같은 안면 특징으로 인한 얼굴 이미지의 유사성 때문에 여전히 풀기 어려운 문제이다[5]. 이러한 문제는 전통적인 내용 기반 이미지 검색에 비해 얼굴 이미지 검색을 보다 어렵게 만드는 요소로써, 본 논문에서는 얼굴 이미지 검색의 문제를 얼굴 인식 방법을 사용하여 풀고자 한다.

신호 처리 분야에서 효과적인 압축과 복원을 위해 전통적으로 연구되던 SR(Sparse Representation)이 최근 얼굴 인식 분야에서도 성공적인 업적을 보여주고

있음이 보고되고 있다[6-7]. SR 을 이용한 교사 학습 방법, 즉 SRC(Sparse Representation Classifier)는 다양한 조명의 변화, 원 이미지에 다른 이미지가 겹치거나 변장한 경우, 변질된 경우, 저해상도 등 식별 성능을 저해하는 다양한 환경에서도, 정면 얼굴 이미지를 차원 축소 방법을 사용하지 않고서도 매우 높은 얼굴 인식률을 보장한다고 보고되고 있다[8].

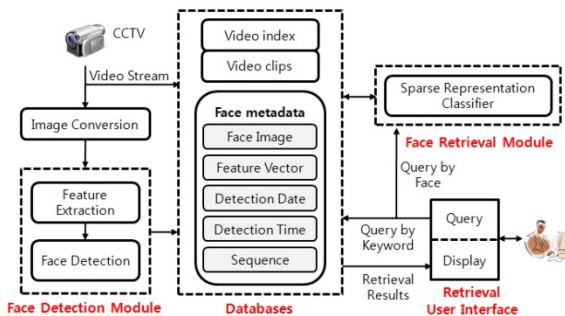
본 논문의 선행 연구로써, 이종욱 등[9]은 얼굴 인식 기반의 범죄 용의자 탐지 및 식별 시스템을 설계 및 구현하였다. 본 논문에서는 선행 연구의 완성도를 높이는 차원에서 얼굴 인식 기반의 범죄 용의자 탐지 및 식별 시스템에 얼굴에 의한 질의(Query-by-face) 검색이 가능한 부 시스템을 제안하고자 한다. 즉, 본 논문에서는 감시 및 보안을 위해 설치한 CCTV 에서 취득한 영상을 효과적으로 검색하는 시스템으로, 영상에서 탐지된 얼굴 정보로 데이터베이스를 실시간으로 구성하고, 사후 얼굴 이미지로 질의를 했을 경우, SRC 에 기반한 얼굴 인식 기법으로 얼굴 이미지 검색을 수행한다. 또한, 자체 제작한 KUFED 와 CCTV 환경의 얼굴 인식 기반 검색 시스템 환경을 캠퍼스 내에서 모의 구축하여 제안된 시스템의 성능을 실험적으로 검증한다.

본 논문의 구성은 다음과 같다. 2 장에서는 본 논문에서 제안하는 감시 시스템을 위한 얼굴 인식 기반의 검색 시스템의 전체적인 구성도와 얼굴 검색 모듈을 자세하게 소개한다. 3 장에서는 실험 결과 및 성능 분석을 서술하고, 4 장에서는 결론 및 향후 연구과제에 대하여 논한다.

2. 감시 시스템을 위한 얼굴 인식 기반의 검색 시스템

2.1 얼굴 인식 기반의 검색 시스템 구조

본 논문에서 제안하는 CCTV 환경에서 얼굴 인식 기반의 감시 영상 검색 시스템 구조는 그림 1 과 같다.



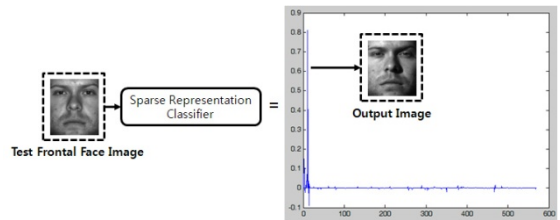
(그림 1) 얼굴 인식 기반의 검색 시스템 구조

제안된 시스템은 크게 얼굴 탐지 모듈, 탐지된 얼굴 정보를 저장하는 데이터베이스, 얼굴 검색 모듈, 사용자 인터페이스로 구성되며, 다음과 같은 절차로 수행된다: 1) CCTV 에서 영상을 취득한다; 2) 취득한 영상을 이미지로 변환한다; 3) 얼굴 탐지 모듈은 감시 영상에서 얼굴 부분을 탐지하고, 얼굴 인식을 위하여

얼굴 이미지 정보를 벡터로 변환한다; 4) 탐지된 얼굴 이미지, 특징벡터, 얼굴 탐지 날짜, 시간 정보 등을 메타데이터화 하고, 원 영상에서 얼굴이 탐지된 영상 부분을 분할하여 비디오 클립으로 데이터 베이스에 저장한다; 5) 얼굴 검색 모듈은 얼굴 인식 분야에서 그 성능이 이미 검증된 SRC 알고리즘을 이용하여 입력 이미지와 저장된 얼굴 이미지들의 유사도 값에 따른 순서로 정렬하여 얼굴 이미지들을 사용자 인터페이스를 통해 출력한다.

2.2 SRC 를 이용한 얼굴 이미지 검색

SR 을 이용한 교사 학습 방법, 즉 SRC 는 다양한 조명 변화, 원 이미지에 다른 이미지가 겹치거나 변장한 경우, 원 이미지가 변질된 경우, 저해상도 등 식별 성능을 저해하는 다양한 환경에서도, 정면 얼굴 이미지를 차원 축소 방법을 사용하지 않고서도 매우 높은 얼굴 인식률을 보장한다고 보고되고 있다[8]. 그림 2 는 SRC 를 이용하여 얼굴 인식을 테스트한 결과를 보여준다. 정면 얼굴 이미지를 SRC 에 적용한 결과, 입력 얼굴 데이터가 속하는 클래스에서만 상대적으로 높은 소속값을 가지며, 속하지 않는 클래스에서는 대부분 0 값을 갖거나 낮은 값을 갖는 것을 확인할 수 있다. 이 때 0 값을 갖는 클래스들의 수가 매우 많으므로 이를 sparse representation 이라 한다[8].



(그림 2) SRC 를 이용한 얼굴 인식의 예

SR 기반의 얼굴 이미지 검색 문제는 다음과 같이 수식으로 표현된다[7-8]. 일반적으로 n 차원을 가지는 고차원 데이터 행렬 A 는 데이터 포인트들의 집합으로 $A = \{a_1, \dots, a_n\}$ 로 표현된다. A 에 속하는 한 점 $a_i \in A$ 은 그 점과 이웃한 점들과의 선형 조합(linear combination)으로 표현된다. 임의의 클래스에 속하는 데이터 포인트들의 집합 $\{a_1, \dots, a_n\}$ 가 주어졌다면, 같은 클래스에 속하는 새로운 데이터 포인트 a^* 는 $\{a_1, \dots, a_n\}$ 의 선형 조합으로 표현된다.

$$a^* = \beta_1 a_1 + \dots + \beta_n a_n. \quad (1)$$

즉, n 개의 학습 샘플 $\{a_1, \dots, a_n\}$ 가 주어졌을 때, 선형 조합은 선형 부분 공간(linear subspace) W 를 생성(span)하며, 새로운 데이터 포인트 a^* 는 포인트가 속하는 클래스에 가장 근사한 부분 공간에 놓이게 된다.

$$W = span\{a_1, \dots, a_n\}. \quad (2)$$

I 개의 클래스를 갖는 학습 샘플이 주어졌다면, 패턴 인식의 기본 방법론은 새로운 테스트 샘플을 학습 샘플

플 클래스를 이용하여 이에 상응하는 클래스로 정확하게 분류하는 것이다. i 번째 클래스에 속하는 n_i 학습 샘플들은 행렬 $A = [v_{i,1}, v_{i,2}, \dots, v_{i,n_i}] \in \mathbb{R}^{m \times n_i}$ 의 열로 정렬이 되며, 이것은 각각의 학습 샘플 집합 $A = [A_1, A_2, \dots, A_k]$ 행렬로 표현된다. 선형 표현 가설 하에서, 테스트 샘플 $y \in \mathbb{R}^m$ 는 학습 샘플들에 의해 생성된 선형 부분 공간에서 근사 된다. 이것은 다음과 같은 행렬식으로 표현된다.

$$y = Ax \in \mathbb{R}^m, \quad (3)$$

위 식에서 x 는 계수 벡터(coefficient vector)이다. 클래스 i 에 속하는 샘플 y 의 계수 벡터 x 는 i 와 관련된 학습 데이터 값을 제외하고는 0을 갖고, 다음과 같이 표현된다.

$$x = [0, \dots, 0, \beta_{i,1}, \beta_{i,2}, \dots, \beta_{i,n_i}, 0, \dots, 0]^T \in \mathbb{R}^n \quad (4)$$

(4)는 $y = Ax$ 선형시스템의 방정식을 풀면 x 값을 얻을 수 있으며, 다음과 같은 x 의 최적해(optimization)를 구하는 문제로 변경할 수 있다[7-8].

$$\hat{x}_0 = \arg \min_x \|x\|_0 \quad \text{subject to } y = Ax, \quad (5)$$

식 (5)를 이용해 해를 찾는 것은 NP-hard 문제이며 해를 찾기 어렵다. 그러나 만약 x 의 해가 충분히 sparse하다면, l^0 -노름(norm) 최소화 문제 (5)는 다음의 convex relaxed optimization 문제를 이용하여, 근사해인 l^1 -노름 최소화 문제로 풀 수 있다[7-8].

$$\hat{x}_1 = \arg \min_x \|x\|_1 \quad \text{subject to } y = Ax. \quad (6)$$

본 논문에서는 [8]에서 사용한 알고리즘을 얼굴 이미지 검색 문제에 적용하였다.

알고리즘 1. SRC 를 이용한 얼굴 이미지 인식

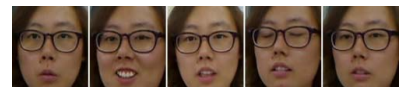
1. **Input:** a matrix of training samples
 $A = [A_1, A_2, \dots, A_k] \in \mathbb{R}^{m \times n}$ for k classes, a test sample $y \in \mathbb{R}^m$.
 2. Normalize the columns of A to have unit l^2 -norm.
 3. Solve the l^1 -minimization problem:
 $\hat{x}_1 = \arg \min_x \|x\|_1 \quad \text{subject to } y = Ax$.
 4. Compute the residuals $r_i(y) = \|y - A\delta_i(\hat{x}_1)\|_2$ for $i = 1, \dots, k$.
 5. **Output:** $\text{identity}(y) = \arg \min_i r_i(y)$
-

얼굴 이미지 검색을 위해 테스트 이미지를 SRC에 적용하면 그와 유사한 특징을 가지는 얼굴 이미지가 0이 아닌 높은 값을 갖게 된다. 이러한 값을 순차적으로 정렬하여 그에 해당하는 얼굴 이미지와 관련 정보들을 검색 결과로써 관리자에게 보여준다. 영상에서 탐지된 얼굴 이미지는 행렬 A 의 열벡터(column vector)로 표현된다. 따라서, 감시 영상에서 새롭게 탐지된 얼굴 이미지는 특징 벡터로 변환되고, 열벡터로 만든 다음 행렬에 추가된다. 또한 SRC는 기계학습

입장에서 보면, 적극적 학습자(eager learner)라기 보다는 게으른 학습자(lazy learner)로써 선행 학습을 하는 적극적 학습자와는 달리 풀어야 할 문제가 발생했을 경우에만 반응하는 문제 해결자이다. 결국, SRC는 과도한 학습을 요구하는 기계학습 알고리즘의 고질적인 문제점을 피할 수 있을 뿐만 아니라, 점진적 갱신의 특성으로 인하여 감시 영상에서 계속적으로 탐지되는 얼굴 데이터베이스의 증가 문제에도 유연하게 대처할 수 있다.

3. 실험 및 결과 분석

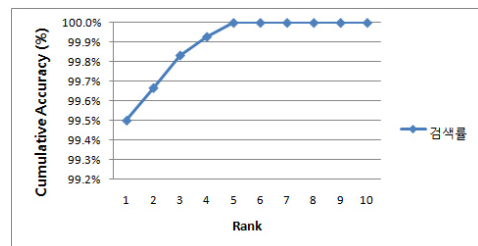
본 논문에서 제안한 감시 시스템을 위한 얼굴 인식 기반의 검색 시스템을 실험적으로 평가하기 위하여, KUFD 및 캠퍼스 내에서 범칙 용의자 또는 특정인을 탐지 및 검색하는 시스템 환경을 구축하였다. 제안된 시스템은 MATLAB과 sparse representation solver인 SparseLab[10]을 이용하여 구현하였다. 일정한 조명이 비추는 실내에서 인터넷 화상 통신용 웹캠(webcam)을 이용하여 탐지된 다양한 얼굴 표정의 영상들 중 정면에 가까운 얼굴 이미지를 수집한 뒤, 얼굴 영역만을 잘라내어 얼굴 이미지 데이터베이스를 제작하였다. 총 28명, 각 30장의 컬러 이미지로 구성하였으며, 크기는 140x140의 JPG 형식이다. KUFD 샘플 이미지들은 그림 3과 같다.



(그림 3) KUFD 샘플 이미지

얼굴 인식 기반의 검색 실험 시, KUFD의 칼라 이미지를 흑백 이미지로 변환하고, 90x90으로 크기를 조정 후, 벡터의 노름(norm)이 1인 단위 벡터로 정규화하였다.

KUFD를 이용한 얼굴 인식 기반의 검색 실험 결과는 그림 4와 같다. KUFD의 총 28명, 한 명당 30장의 얼굴 이미지에서, 각각 15장을 영상에서 탐지되어 저장되어 있는 얼굴 이미지로 사용하였다. 그리고 나머지 클래스 당 15장을 관리자가 질의로 사용하는 테스트 데이터로 실험에 적용하였다. 얼굴 질의를 이용한 검색 결과 첫 번째 추천 이미지의 정확도가 99.5%라는 높은 성능을 기록하였다.



(그림 4) KUFD를 사용한 얼굴 검색 누적 정확도

특정 지역을 감시하는 CCTV 환경에서 감시 시스템을 위한 얼굴 인식 기반의 검색 시스템을 캠퍼스 내

에서 모의 구축한 본 논문의 사용자 인터페이스는 그림 5 와 같다. CCTV 관리자는 실시간으로 감시 영상과 얼굴 영역이 탐지되는 것을 확인할 수 있다. 얼굴 이미지를 질의로 사용하여 가장 유사한 얼굴 이미지 목록들을 확인할 수 있으며, 관련된 영상 클립, 탐지된 날짜, 시간 정보 등을 확인할 수 있도록 구성하였다.



(그림 5) 얼굴 질의 기반의 검색 시스템 인터페이스

4. 결론

본 논문에서는 선행 연구[9]의 완성도를 높이는 차원에서 얼굴 인식 기반의 범죄 용의자 탐지 및 식별 시스템에 얼굴에 의한 질의 검색이 가능한 부 시스템을 제안하였다. 즉 본 논문에서는 감시 및 보안을 위해 설치한 CCTV 에서 취득한 영상을 효과적으로 검색하는 시스템으로, 영상에서 탐지된 얼굴 정보로 데이터베이스를 실시간으로 구성하고, 사후 얼굴 이미지로 질의를 했을 경우, SRC 에 기반한 얼굴 인식 기법으로 얼굴 이미지 검색을 수행한다. 또한, 자체 제작한 KUFDF 와 CCTV 환경의 얼굴 인식 기반 검색 시스템 환경을 캠퍼스 내에서 모의 구축하여 제안된 시스템의 성능을 실험적으로 검증하였다.

향후 연구과제로는 본 연구에서 제안된 프로토타입의 감시 시스템을 위한 얼굴 인식 기반의 검색 시스템을 실 세계에서 구현하고자 한다.

참고문헌

- [1] 차건상, 신용태, “CCTV 설치 증가에 따른 개인 영상 정보 보호 주요 이슈”, 정보과학회지, Vol. 27, No. 12, pp. 119-126, 2009.
- [2] C. Cotsaces, N. Nikolaidis and I. Pitas, “Face-based Digital Signatures for Video Retrieval,” IEEE Transactions on Circuits and Systems for Video Technology, Vol. 18, No. 4, pp. 549-553, 2008.
- [3] 문성룡, 신성, “컨텍스트 기반의 지능형 영상 감시 시스템 구현에 관한 연구”, 전자공학회논문지, Vol. 47, No. 3, pp. 11-12, 2010.
- [4] T. Shan, S. Chen, C. Sanderson and B. C. Lovell, “Towards Robust Face Recognition for Intelligent-CCTV based Surveillance using One Gallery Image,” IEEE Conference on

Advanced Video and Signal Based Surveillance, pp. 470-475, 2007.

- [5] D. B. Megherbi and Y. Miao, “A Distributed Technique for Recognition and Retrieval of Faces with Time-varying Expressions,” IEEE International Conference on Computational Intelligence for Measurement Systems and Application, pp. 8-13, 2009.
- [6] L. Qiao, S. Chen and X. Tan, “Sparsity Reserving Projections with Applications to Face Recognition,” Journal of Pattern Recognition, Vol. 43, No. 1, pp. 331-341, 2010.
- [7] Y. Ji, T. Lin and H. Zha, “Mahalanobis Distance based Non-negative Sparse Representation for Face Recognition,” International Conference on Machine Learning and Applications, pp. 41-46, 2009.
- [8] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry and M. Yi, “Robust Face Recognition via Sparse Representation,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 31, No. 2, pp. 210-227, 2009.
- [9] 이종욱, 강봉수, 이한성, 박대회, “얼굴 인식 기반의 범죄 용의자 탐지 및 식별 시스템”, 정보과학회 추계학술대회, 2010. (게재예정)
- [10] Stanford SparseLab, <http://sparselab.stanford.edu/>