

Mean Shift Clustering을 이용한 영상 검색결과 개선

권경수* 신윤희** 김영래*** 김은이****

(Kyung Su Kwon, Yunhee Shin, Youngrae Kim, Eun Yi Kim)

요 약 본 논문에서는 감성 공간에서 mean shift clustering과 user feedback을 이용하여 영상 검색 결과를 개선하기 위한 시스템을 제안한다. 제안된 시스템은 사용자 인터페이스, 감성 공간 변환, 검색결과 순위 재지정(re-ranking)으로 구성된다. 사용자 인터페이스는 텍스트 형태의 질의 입력과 감성 어휘 선택에 따른 user feedback에 의해 개선된 검색결과를 보인다. 사용된 감성 어휘는 고바야시가 정의한 romantic, natural, casual, elegant, chic, classic, dandy, modern 등의 8개 어휘를 사용한다. 감성 공간 변환 단계에서는 입력된 질의에 따라 웹 영상 검색 엔진(Yahoo)에 의해 검색된 결과 영상들에 대해 컬러와 패턴정보의 특징을 추출하고, 이를 입력으로 하는 8개의 각 감성별 분류기에 의해 각 영상은 8차원 감성 공간으로의 특징 벡터로 변환된다. 이때 감성 공간으로 변환된 특징 벡터들은 mean shift clustering을 통해 군집화 되고, 그 결과로써 대표 클러스터를 찾게 된다. 검색결과 순위 재지정 단계에서는 user feedback 유무에 따라 대표 클러스터의 평균 벡터와 user feedback에 의해 생성된 사용자 감성 벡터에 의해 검색 결과를 개선할 수 있다. 이때 각 기준에 따라 유사도가 결정되고 검색결과 순위가 재지정 된다. 제안된 시스템의 성능을 검증하기 위해 7개의 질의의 각 400장, 총 2,800장에 대한 Yahoo 검색 결과와 제안된 시스템을 개선된 검색 결과를 비교하였다.

핵심주제어 : 감성 개념, 평균 이동 군집화, 영상 검색 엔진, 결과 개선, 사용자 피드백

Key Words : emotional concepts, mean shift clustering, image search engine, result enhancement, user feedback

1. 서 론

최근 인터넷의 속도와 저장 장치의 용량이 급히 증가함에 따라 웹상에 존재하는 거대한 양의 영상 데이터를 효율적으로 검색하기 위한 방법에 관한 연구가 연구자들에게 많은 관심을 받고 있다. 전통적으로 내용 기반의 영상 검색(CBIR)은 자동으로 영상의 특징을 추출하고, 이를 이용하여 대용량의 영상 데이터베이스에서 효과적으로 영상을 검색할 수 있게 한다^[1]. 내용 기반 영상 검색에서 좋은 성능을 보일지라도 이러한 시스템들은 실시간으로 처리되는 방대한 영상 데이터에서는 높은 계산비용이 초래하게 된다.

이에 따라, 최근 이들의 방법을 결합한 방법들이 등장하고 있는데, 이 시스템들은 속도를 줄이기 위해 먼저 텍스트 정보를 사용하여 영상을 검색한 후, 이를 시각정보를 사용하여 결과의 정확도를 향상시킨다. 이러한 방법들은 사용자 피드백과 클러스터링이다. 사용자 피드백은 CBIR에서 검색 성능을 향상 시키는 중요한 기술이지만, 사용자에 대한 의존도가 높은 단점을 가진다^[2]. 이후에, 색상 및 질감, 또는 의미 있는 Semantic과 같은 개체의 이름, 그리고 인간의 감정 추출 등 시각적 기능을 사용한다.

시각적인 특징에 기반한 영상 검색 결과를 클러스터링 알고리즘을 사용하여 의미적 기반의 검색에 적용하였다^[3]. Semantic은 첫 번째로 물리적인 의미로 객체의 이름이나 사건, 그리고 두 번째로

* 경북대학교 컴퓨터공학과 박사과정
** 건국대학교 신기술융합학과 박사과정
*** 건국대학교 신기술융합학과 석사과정
**** 건국대학교 신기술융합학과 교수

추상적인 의미로 인간의 감성으로 나눌 수 있다. 물리적 의미에 대한 연구는 [3]에서 연구되었으며,

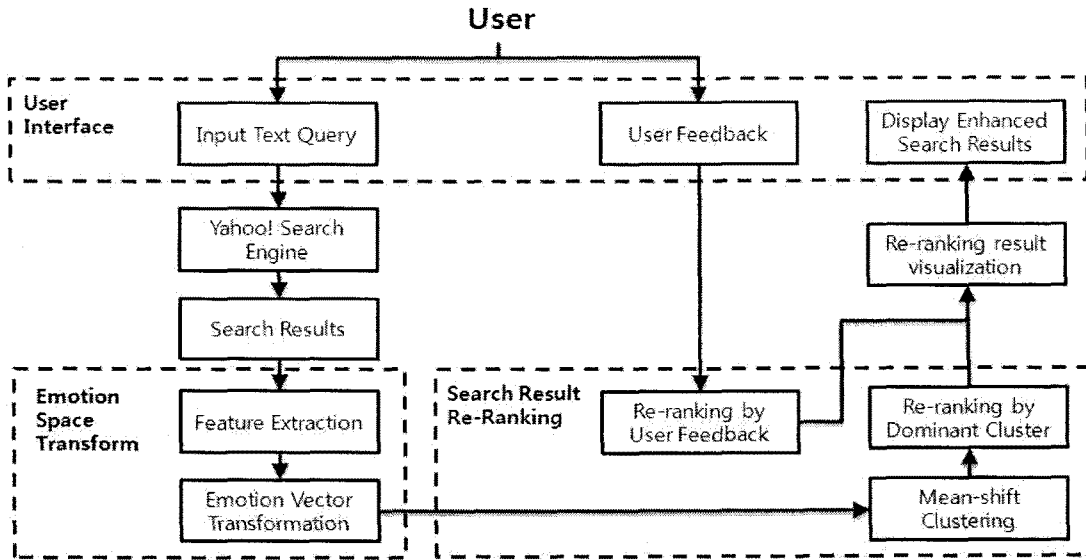


그림 1. 시스템 구조

추상적 의미에 대한 연구는 여전히 미흡하다. 그러나 추상적 의미는 더 의미 있는 결과를 제공할 수 있기 때문에 매우 중요하다.

본 논문에서는 영상 검색결과를 개선하기 위해 감성 공간에서 mean shift clustering과 검색 수행시 user feedback을 이용하여 영상 검색 시스템을 제안한다. 본 논문에서 제안하는 영상 검색 시스템은 그림 1에서 보여준다. 제안된 시스템은 사용자 인터페이스, 감성 공간 변환, 검색결과 순위 재지정(re-ranking)으로 구성된다. 사용자 인터페이스는 텍스트 형태의 질의 입력과 감성 어휘 선택에 따른 user feedback에 의해 개선된 검색 결과를 보인다. 감성 공간 변환 단계에서는 입력된 질의에 따라 웹 영상 검색 엔진(Yahoo)에 의해 검색된 결과 영상들에 대해 컬러와 패턴정보의 특징을 추출하고, 이를 입력으로 하는 8개의 각 감성별 분류기에 의해 각 영상은 8차원 감성 공간으로의 특징 벡터로 변환한다. 검색결과 순위 재지정 단계에서는 사용자의 영상 검색시 user feedback 유무에 따라 mean shift clustering의 결과인 대표 클러스터의 평균 벡터와 user feedback에 의해 생성된 사용자 감성 벡터에 따라 유사도를 통하여 검색 결과를 개선할 수 있다.

제안된 시스템의 성능을 검증하기 위해 7개의 질의의 각 400장, 총 2,800장에 대한 Yahoo 검색 결과와 제안된 시스템을 개선된 검색 결과를 비교하였다.

II. 제안된 시스템

본 논문에서는 사용자가 입력한 질의에 대한 개선된 검색 결과를 얻기 위해 먼저 텍스트 형태의 질의로 Yahoo 영상 검색 엔진을 통해 검색 결과를 얻는다. 얻어진 영상 검색 결과는 감성공간으로의 특징 벡터들로 변환되고, 이들을 mean shift clustering 을 이용한 순위 재지정을 통해 일차적으로 검색 결과를 개선한다. 이후에 user feedback 을 이용하여 사용자의 의도를 파악하고 이에 가까운 영상을 찾기 쉽도록 결과를 재개선한다. 제안된 시스템은 크게 사용자 인터페이스, 감성 공간 변환, 검색 결과 순위 재지정 단계로 구성된다.

2-1. 사용자 인터페이스

개선된 영상 검색 결과를 바탕으로 검색을 수행하기 위해 본 논문에서는 그림 2와 같이 인터페이스 화면을 구성한다. 영상 검색을 위해 제시된 인터페이스 화면은 질의를 입력 받기 위한 질의 입력 박스, 질의에 대한 영상 검색 결과를 보여주는, 사용자의 원하는 감성 어휘의 선택을 돕기 위한 감성 체크 박스, 검색 결과 내에서 페이지를 이동하기 위한 페이지 이동 버튼들로 구성된다.

사용자는 감성 어휘로의 영상 검색을 위해 화면의 우측에 위치한 8개의 감성 체크 박스를 통해

user feedback을 시스템에 줄 수 있다.

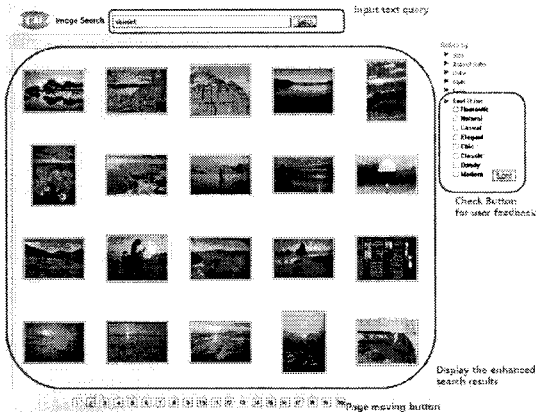


그림 2. 인터페이스 화면 구성

2-2. 감성 공간 변환

검색 결과를 개선하기 위해 영상들로부터 특징 추출과 이들을 이용한 감성 공간 변환을 수행한다. 다음은 감성 공간 변환을 위한 컬러 정보 추출, 패턴 정보 추출, 감성 분류기를 이용한 감성 벡터 생성에 대한 설명이다.

● 컬러정보 추출

고바야시의 이론에 기반하여 입력 영상에서의 대표 컬러 정보를 추출한다^[4]. 고바야시가 제안한 Color image scale은 PCCS 색상 공간에서 컬러 양자화 방법을 사용하여 영상의 대표 컬러를 추출한다. 영상을 구성하는 대표 컬러를 추출하기 위해서 입력 영상을 컬러공간에서 특정 색상을 분석한다. 이에 따라 고바야시의 컬러 공간에서 양자화 방법으로 대표 컬러 정보를 추출하여 감성분류를 위한 특징으로 사용한다.

● 패턴 정보 추출

패턴 정보의 추출을 위해 웨이블릿 변환을 사용한다. 웨이블릿 변환은 입력영상을 각각 수직, 수평적 특징들을 나타내는 계수들로 구성된 4개의 sub-block LL, LH, HL, HH로 나눈다. LL은 원 영상을 변환 레벨만큼 압축한 것으로 모든 웨이블릿의 sub-block 중에서 가장 중요한 정보를 포함하고 있으며 나머지 sub-block은 수직, 수평, 대각선 방향에 대한 에지 정보를 나타내고 있다. 웨이블릿 변환은 LL에서 계속적으로 수행되어 4개씩의 sub-block을 얻을 수 있다. 이를 P번 동안 반복적으로 처리되어 4P개의 sub-block을 만든다. 제안된 방법은 구

해진 각 블록들에 대해 파라미터를 계산 한다^[5]

● 감성 벡터 생성

8차원 감성 공간에서 영상을 표현하기 위해 다층 퍼셉트론(MultiLayer Perceptron, MLP) 기반의 분류기를 사용 한다^[6,7]. 입력 노드의 수는 추출된 특징 벡터의 차원과 같다. 출력 노드의 수는 한 개로 출력 값은 0에서 1사이의 실수로 정규화 되며, 각 분류기를 통해 얻어지는 값들은 {romantic, natural, casual, elegant, chic, classic, dandy, modern} 순으로 나열되어 하나의 8차원 벡터를 구성한다.

2-3. 검색 결과 순위 재지정

검색 결과 개선은 사용자의 feedback의 유무에 따라 두 가지 종류의 유사도 계산에 순위가 재지정 된다. 먼저 사용자의 feedback이 없는 경우 유사도는 mean shift clustering 결과로 찾아진 대표 클러스터의 평균 벡터와 각 영상들의 감성 벡터 사이의 코사인 유사도를 계산하여 그 값에 따라 순위를 재지정한다.

사용자의 feedback이 있는 경우 사용자가 선택한 감성들을 '1'로 그렇지 않은 벡터의 요소들은 '0'으로 설정한 벡터를 이용하여 다음과 같이 유사도를 계산하고 그 값으로 순위를 재지정한다.

$$D = w_1(\mu_d \cdot e_i) + w_2(u \cdot e_i) + w_3M(e_i)$$

여기서 첫 번째 term은 대표 클러스터의 평균 벡터와 영상들의 감성벡터간의 코사인 유사도 이고, 두 번째 term은 사용자 벡터와 영상간의 코사인 유사도 이다. 마지막 term은 마할라노비스 거리를 이용한 유사도이다. 각 term의 가중치는 0.3, 0.5, 0.2로 설정하여 계산하였다.

III. 실험

실험에서 사용되는 영상 데이터는 bridge, car, field, forest, skirts, sunset, temple의 7가지 질의에 대해 Yahoo 영상 검색 엔진을 통해 각 400장씩 얻는다. 이러한 7가지 질의는 다양한 감성의 어휘와 조합되어 자연스럽게 사용될 수 있는 질의이다. Mean shift clustering에서의 대역폭은 감성 공간의 최대 거리인 2.828의 30%값인 0.85로 설정하였고, 400개의 영상들 중에 상위 200개의 이미지에

대해 클러스터링을 수행하였다. 그림 3은 사용자의 feedback이 없을 경우 mean shift clustering에 의

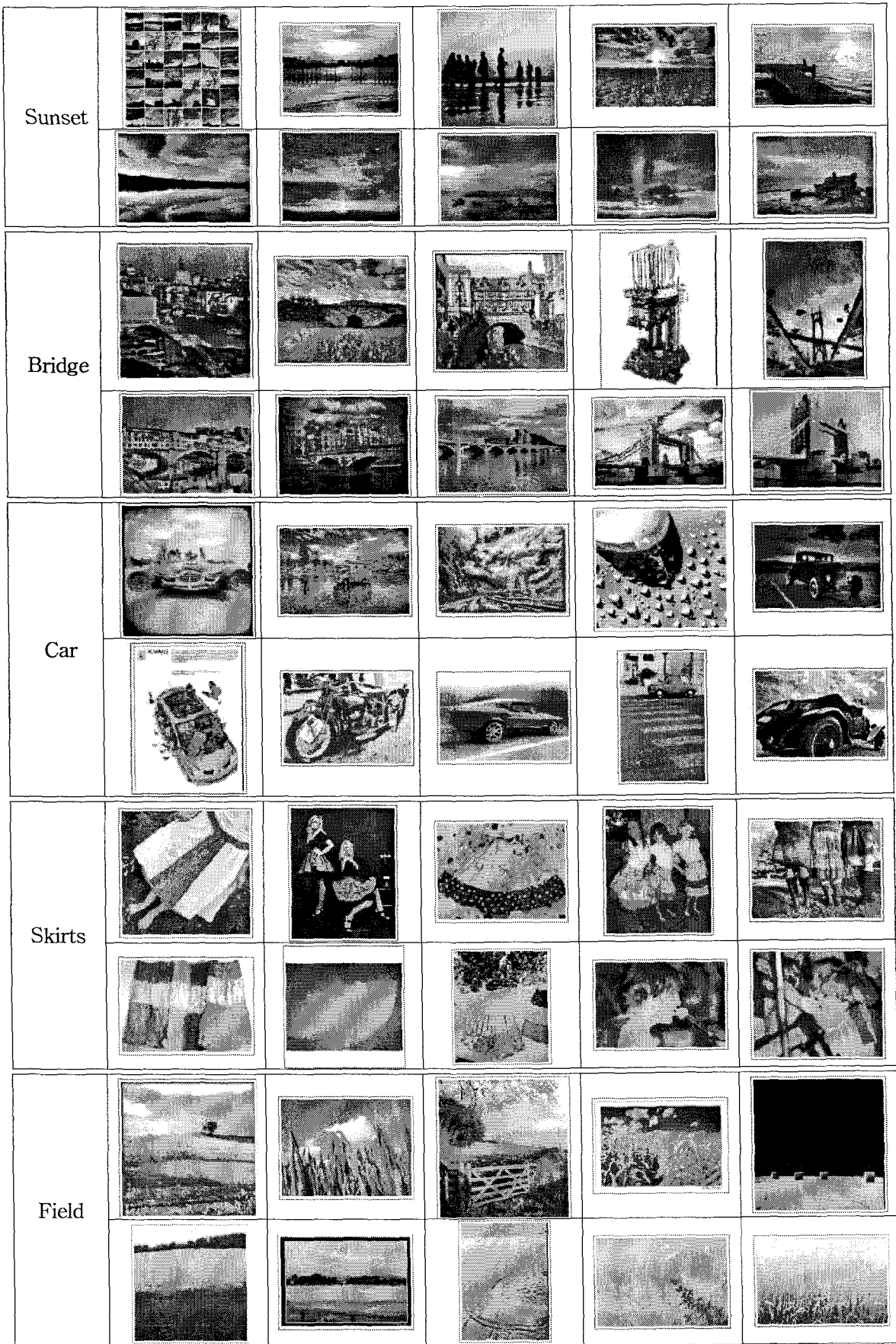


그림 3. 질의에 따른 검색 결과 비교: 각 질의의 첫 행은 Yahoo의 검색 결과, 둘째 행은 제안된 방법의 검색 결과

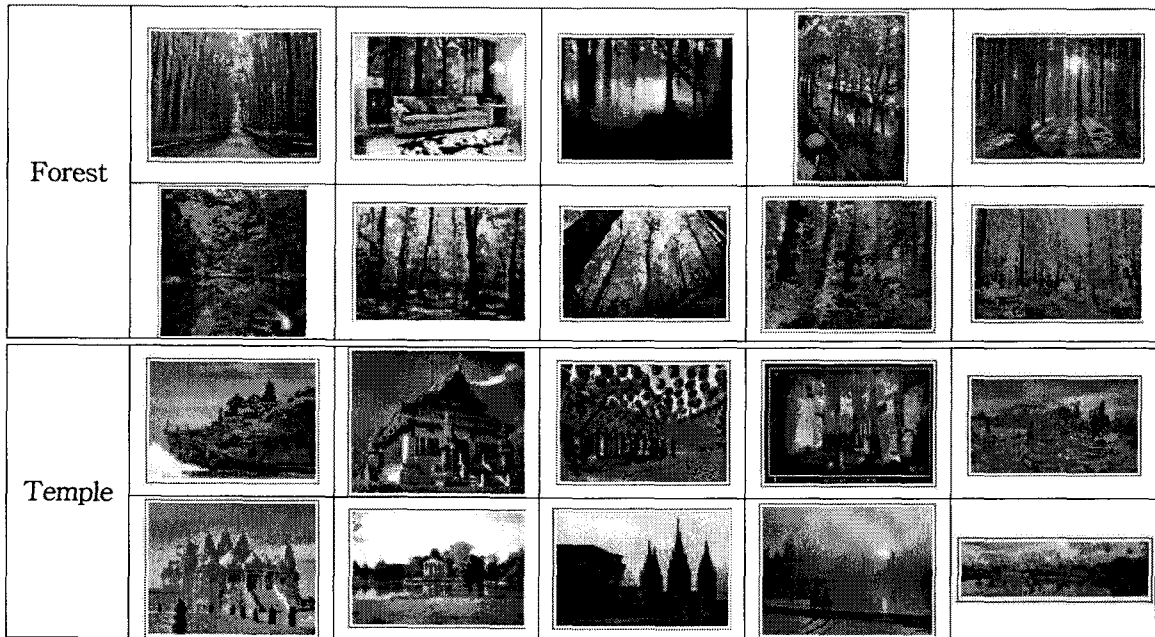


그림 3. 질의에 따른 검색 결과(계속)

한 검색 결과와 Yahoo의 검색결과를 상위 5개의 결과에 대해 비교하여 보여준다. 결과에서 보듯이 제안된 방법의 검색결과가 Yahoo의 검색 결과보다 질의에 더 의미있는 영상들을 결과로 보여줌을 확인할 수 있다.

한 Yahoo 검색 결과와 제안된 시스템을 개선된 검색 결과를 비교하였다. 결과로 Yahoo의 검색결과에 비해 제안된 방법의 검색결과가 더 향상되었음을 확인할 수 있었다.

참 고 문 헌

IV. 결론

본 논문에서는 영상 검색결과를 개선하기 위해 감성 공간에서 mean shift clustering과 검색 수행시 user feedback을 이용하여 영상 검색 시스템을 제안하였다. 제안된 시스템은 사용자 인터페이스, 감성 공간 변환, 검색결과 순위 재지정으로 구성하였다. 감성 공간으로의 변환을 위해 Yahoo 결과 영상들로부터 컬러와 패턴 정보를 컬러 양자화와 웨이블릿 변환을 통해 기술했고, 이들을 8개의 감성 분류기를 통해 영상을 감성 벡터들로 변환했다. 우선 감성 공간 내에 존재하는 영상의 감성 벡터들을 입력으로 하는 mean shift clustering을 통해 찾아진 대표 클러스터의 평균 벡터를 통해 검색 결과가 개선되었고, user feedback이 있을 경우에 대한 순위 재지정을 통해 검색 결과를 재개선 하였다. 제안된 시스템의 성능을 검증하기 위해 7개의 질의의 각 400장, 총 2,800장에 대

- [1] Ritendra Datta, Dhiraj Joshi., Jia Li., James Z., Wang., "Image retrieval: Ideas, influences, and trends of the new age", In: ACM Computing Surveys (CSUR), vol.40 no.2, pp.1-60, 2008
- [2] S. Wang, F. Jing, J. He, Q. Du, and L. Zhang. IGroup, "Presenting Web Image Search Results in Semantic Clusters", CHI, pp. 587-596, 2007
- [3] Mohammad Reza Keyvanpour., Shabnam Asbaghi., "A new approach for interactive semantic image retrieval using the high level semantics" In: Proc. SAC, pp.1175-1179, 2008
- [4] Shigenobu Kobayashi, "The Aim and Method of the Color Image Scale, Color Research & Application", Vol. 6, pp. 93-107, 1981
- [5] Huiping Li, David Doermann, and Omid Kia, "Automatic Text Detection and Tracking in Digital Video", IEEE Transactions on Image Processing, Vol. 9, No. 1, January 2000.
- [6] Simon Haykin, Neural Networks a comprehensive foundation 2nd edition, Prentice Hall, pp.10-13, pp.156-173, 1999.

- [7] H.-U. Bauer and T. Geisel, "Dynamics of signal processing in feedback multilayer perceptrons," in Proc. Int. Joint Conf. Neural Networks, pp. 131-136, 1990.