ISSN 1598-0170 (Print)
ISSN 2287-1136 (Online)
http://www.jksii.or.kr

사진 콘텐츠 분류를 위한 태그 클러스터링 기법 및 태그 추천☆

A Tag Clustering and Recommendation Method for Photo Categorization

원 지 현[1] 이 종 우[1] 박 희 민[2*]
Ji-Hyeon Won Jongwoo Lee Heemin Park

요 약

개인이 사용할 수 있는 스마트 기기가 다양해지면서 여러 기기로 생산된 사진 콘텐츠가 어떤 기준이나 규칙 없이 분산되어 있어 콘텐츠를 관리하고 원하는 콘텐츠를 검색하는 것이 어려워졌다. 따라서 본 논문에서는 개인 사진 콘텐츠를 효과적으로 분류하기 위하여 의미적 유사도를 기반으로 한 태그 클러스터링 기법과 개인이 사진에 태그를 넣을 때 초기 클러스터를 기반으로 태그를 추천하는 방법을 제안한다. 태그들 사이의 유사도를 계산하여 서로 관련이 있다고 판단되는 태그들을 클러스터링 하는데, 태그가 같은 클러스터에 포함되어 있으면 그 태그를 가진 사진들도 유사성을 가진다고 볼 수 있으므로 개인 사진들을 의미에 따라 분류하는데 이용할 수 있다. 또한 분류된 초기 클러스터로 태그를 추천하여 개인 사용자가 태그를 분류에 맞게 추가할 수 있어 사진 분류 관리가 용이해진다.

주제어 : 태그 분류, 태그 추천, 클러스터링, 사진 콘텐츠, 사진 관리

ABSTRACT

Recent advance and popularization of smart devices and web application services based on cloud computing have made end-users to directly produce and, at the same time, consume the image contents. This leads to demands of unified contents management services. Thus, this paper proposestag clustering method based on semantic similarity for effective image categorization. We calculate the cost of semantic similarity between tags and cluster tags that are closely related. If tags are in a cluster, we suppose that images with them are also in a same cluster. Furthermore, we could recommend tags for new images on the basis of initial clusters.

☞ keyword : Tag Categorization, Tag Recommendation, Clustering, Photo Contents, Photo Management

1. 서 론

스마트 미디어 환경에 따라 스마트 폰, 태블릿, 스마트 TV 등 스마트 기기가 대중적으로 보급됨에 따라 개인 사진이 다수의 기기와 저장장치에 분산되어 있으므로 사진 콘텐츠 모음의 파편화, 중복성이 발생하여 효율적인 검색, 열람이 어려워지게 되었다. 사람들은 매일의 일상을 기록해 두고 싶어 하며 쉽게 찾아서 꺼내보거나 인화 혹은 인터넷 상 업로드를 원하기 때문에 산재해있는 사진 콘텐츠를 통합하여 관리할 수 있도록 하는 통합 관리 서비스에 대한 요구사항이 증가하였다.

이러한 문제를 해결하고 효율적으로 사진을 관리하기 위해서는 사용자가 일일이 수작업으로 분류하고 저장하는 방법보다는 자동으로 사진을 분류하고 관리하는 방법을 필요로 한다. 현재 클라우드 컴퓨팅 기반의 서비스가 증가하여 여러 종류의 서비스를 제공하고 있지만 사진을 통합하여 저장하고 열람하는 기능만을 제공하고 있다. 하지만 스마트 기기 사용자는 점점 더 까다로워지고 있으며 사용자에게 맞춤화된 서비스나 조금 더 편리하고 사용성이 높은 서비스를 원한다. 따라서 사용자가 수작업을 하지 않아도 주제별, 이벤트별로 사진을 분류하고 그에 따라 검색할 수 있게 하는 것이 필요하다.

이에 따라 이 논문에서는 사진을 주제별, 이벤트별로 효과적으로 분류할 수 있기 위한 태그 클러스터링 기법을 제안하고 시스템을 구축하였다. 개인 사용자가 웹 서버에 사진을 업로드 할 때 사진을 자동으로 분류하여 저장하고, 업로드 시 사진에 맞는 태그를 추천해주어 사진 콘텐츠의 관리를 용이하게 하였다.

1 Dept. of Multimedia Science, Sookmyung Women's University, Seoul, 140-742, Korea
2 Dept. of Computer Software Engineering, Sangmyung University, Cheonan, 330-720, Korea
* Corresponding Author (heemin@smu.ac.kr)
[Received 3 December 2012, Reviewed 5 December 2012, Accepted 20 March 2013]
☆ 본 논문은 상명대학교 2013학년도 교내연구비에 의하여 수행되었음

2. 관련 연구

2.1 메타데이터 및 이미지 시각정보를 이용한 사진 분류

먼저 메타데이터 중 위치 데이터의 연관성을 이용하여 콘텐츠를 추천하는 시스템의 예로 GeoLife [1]가 있다. 이 연구는 GPS (Global Positioning System) 정보를 이용하여 여행자의 위치 패턴을 저장하고 여러 사람의 위치 패턴을 분석하여 연관성을 찾아내 다른 여행자에게 여행 코스를 추천해준다. 사용자의 GPS 로그로 사용자의 궤적을 저장하고 위치와 사용자 그래프, 각 사용자간의 관계 그래프, 위치와 위치간의 그래프를 통해 추천 여행코스를 추론하게 된다.

또한 Direct Share [2]에서는 사진을 분류할 때 사용자 설문을 통해 사용 습관과 인터페이스 조사 후 적용하는 방법을 제안한다. 사진을 공유할 때 사용자들의 선호도에 따라 공유할 사진에 점수를 주어 업로드 할 사진을 결정하고 공유하는 사진관리 시스템 제안한다. 라운드-로빈 방식 (round-robin concept)을 이용하여 여러 사용자에게 사진을 보여주고 점수를 받아서 공유할 만한 사진인지 결정한다.

사진 콘텐츠를 웹에 업로드 하여 관리하는 대표적인 서비스로는 구글의 Picasa [3]와 야후의 Flickr [4] 서비스가 있다. 두 서비스 모두 사진을 수정할 수 있는 사진 편집 기능을 제공하고 업로드 한 사진을 메타데이터 기반으로 분류하고 검색할 수 있게 해준다. 그림 과 같이 업로드 한 개인 사진 데이터를 열람할 때 날짜 별로 열람이 가능하다. 사진을 생성한 날짜와 업로드 한 날짜를 구분하여 열람하여 사진을 찾을 수 있다.

하지만 이는 단순히 날짜만을 이용한 분류 방법으로, 사용자가 예전 사진을 검색할 때 날짜를 기억하지 못하는 경우가 많아 의미있는 분류 방법이라고 할 수 없다. 이에 따라 사용성이 떨어지는데 다른 분류 기준을 제시하지 않고 있어서 방대한 양의 사진 중 원하는 사진을 쉽게 찾을 수 있는 사용자 인터페이스나 분류 방법이 없다고 할 수 있다.

시각 정보 (visual information)를 이용한 사진 분류에 관한 기존 연구는 지도를 타일로 나누어 위치 태그 (geo-tag)에 따라 비슷한 위치에서 찍힌 사진을 분석하고 시각적 유사도 (visual similarity)에 따라 같은 물체나 사물이 있는 사진을 모아주는 연구 [5]가 있다. 이 연구에서는 영상처리 기법을 통해 이미지의 유사성을 판단하여 분류기준을 정한다. 사진 간의 시각적 유사도를 계산하여 물체 (object)를 추론하고 추론된 이미지 정보와 위치 정보를 종합하여 이벤트를 추론하여 분류하게 된다.

또한 SheepDog [6]에서는 연구에서 미리 정해놓은 컨셉 (태그) 중에 하나를 키워드로 검색하여 검색 결과로 나온 사진과 그룹을 클러스터링하여 그 시각적 특징을 학습시킨다. 그리고 새로 들어오는 사진에서 특징을 추출하고 학습된 데이터를 이용하여 비교한 후 각 사진에 맞는 상위 n개의 그룹 (top-n)과 태그를 추천한다. 추천된 그룹에 따라 사진을 분류할 수 있고 원래 정해진 컨셉 별로 분류 기준이 정해지게 된다. 이 연구는 미리 정해서 저장해놓은 컨셉에 관련된 사진인 경우에만 적용이 가능한 단점이 있어서 일반화하기에는 적합하지 않다.

2.2 태그 클러스터링 및 태그 추천 관련 연구

태그를 이용한 사진 분류에 초점을 맞추어 살펴보면 태그 클러스터링 관련 연구들을 볼 수 있다. 먼저 의미적 태그 클러스터링 방법을 제안하고 태그를 클러스터링하여 의미에 따라 계층적으로 트리를 만든 연구가 있다 [7]. 이 연구에서는 주어진 조건에 맞춰서 모든 태그에 대해 각 클러스터에 포함될 확률을 계산하고 유효한 클러스터 (effective cluster)가 될 때까지 클러스터를 분할한다. 실제적 클러스터가 되면 그 클러스터를 대표하는 태그를 리딩 태그 (leading tag)로 지정하고 의미적 클러스터링을 하게 된다.

또한 태그들의 동시 발생 빈도수 (co-occurrence)에 따라 유사성을 측정하는 연구도 있는데 [8] 어떤 웹 페이지에서 태그들 사이의 동시 발생 빈도수를 세어서 높게 나타나는 태그들을 강한 연관성이 있는 태그 (strongly related tag)라고 정하여 태그의 의미 사이의 연관성을 알아보는 연구이다. 이를 통해 태그들간의 유사성이나 관련성을 알아내어 의미 있는 계층적 태그 구조를 만드는데 활용할 수 있다. 예를 들면 RSS라는 태그가 주어졌을 때 한 페이지에 동시에 다른 단어가 나타나는 횟수를 저장하게 된다. 횟수에 따라 가중치를 부여하여 태그 간의 유사성을 파악하게 된다. 태그간의 유사성을 파악한 후에 그 유사도에 따라 각 태그와 연관된 태그를 추천할 수 있다.

시각 정보와 메타데이터를 함께 이용하여 태그를 추천할 수도 있다. Tag Suggestr [9]에서 제안한 방법으로 사용자가 이미지를 업로드 할 때 초기 태그를 붙이게 되고 이것과 기존 태그 정보를 이용하여 유사성을 판단하

는 방법과 시각 정보인 컬러 히스토그램과 중요 포인트를 이용하여 이미지 유사성 판단하는 방법 두 가지를 사용하여 태그 랭킹을 정하여 정확도를 높이는 방식이다. 이미 존재하는 초기 태그를 가지고 관련 이미지를 찾는데 이용하고 찾은 이미지에 있는 다른 태그들은 모두 추천태그 후보가 된다. 그 후 이미지 사이의 유사성을 이용하여 태그에 가중치를 부여하여 추천할 태그를 최종적으로 결정한다. 이렇게 시각정보를 이용하면 분류가 더 명확해질 수 있으나 그만큼 시각 정보를 분석하는데 노력과 시간이 소요되는 단점이 있다.

이러한 연구들에서 볼 수 있듯이 사진 콘텐츠를 메타데이터와 시각 정보를 이용하여 분류하는 연구는 다양하지만 사진의 기본적인 메타데이터에 의존하여 일반화된 분류기준을 따르고 있거나 시각 정보를 분석하기 위해 시간과 비용이 더 드는 경우가 있어 개인화된 데이터를 다루는 데에 적합하다고 할 수 없다. 따라서 개인 데이터를 관리하는 특화된 방법이 필요하다. 또한 태그를 추천할 때도 일반적 단어들에 대한 태그 추천 시스템에서 더 나아가서 개인 사용자의 개인 데이터를 관리하기 위한 태그 추천 방법이 필요하다. 따라서 본 논문에서는 개인 사용자의 사진 콘텐츠를 쉽게 분류하고 검색하며, 사용자가 자주 사용하는 단어로 태그를 추천해주는 시스템을 제안하여 사용성을 높이도록 하였다.

3. 의미적 유사도 기반 구조적 태그 클러스터링 기법

3.1 개요

태그는 사진을 가장 잘 설명할 수 있는 새로운 데이터이다. 태그는 이미지 분석의 노력 없이 내용 기반의 분류를 가능하게 한다. 이러한 태그를 이용하여 사진을 분류하면 이전의 메타데이터 기반의 분류보다 더 나은 분류를 할 수 있어서 향상된 검색 결과를 나타낼 것이다. 또한 태그는 사진이 담고 있는 물체보다 더 많은 정보를 포함할 수 있다. 태그를 통해 사진을 찍은 사람, 계절 등의 상황 정보를 포함할 수 있게 되면서 개인 사진 데이터의 분류 기준을 태그의 내용 기반으로 정할 수 있어서 검색하는 것이 용이해진다 [10].

태그는 이렇게 사진의 내용을 대표하는 단어이기 때문에 이 데이터를 이용하면 사진의 의미적 분류도 가능해진다. 또한 태그를 구조적으로 클러스터링하게 되면 1차원적인 나열이 아닌 계층적인 관계가 만들어지기 때문에 의미적인 포함관계를 알 수 있게 되어 좀 더 명확한 의미적 클러스터링이 가능하다. 태그의 의미적 클러스터링에는 상위 태그의 개념이 생기기 때문에 한 클러스터를 대표하는 하나 혹은 몇 개의 태그가 생기고 그에 따라 사진을 의미적으로 분류하는 것이 용이해진다. 태그를 클러스터링 하기 위하여 태그 간 유사도를 이용한다. 태그 간 유사도를 태그 사이의 의미상 거리로 간주하여 계산하고 각 태그가 서로 얼마나 유사한 의미를 지니고 있는지 측정하는 척도가 된다.

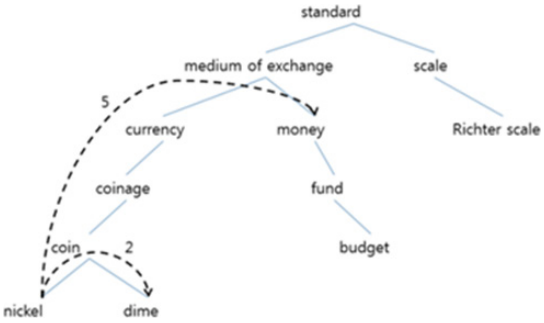
3.2 기법 설명

태그를 구조적으로 클러스터링 하기 위하여 태그 간 의미적 유사성을 측정하고 그것을 이용하여 전체 클러스터를 분할해 나가는 방법을 제시한다. 전체 클러스터를 2개로 분할하고 분할된 클러스터를 다시 2개로 분할하는 것을 반복한다. 유사도를 이용하여 각 클러스터의 비용(cost)를 계산하여 2분할을 진행하는데 더 이상 분할하는 것이 의미가 없는 클러스터는 멈추고 그 외의 클러스터는 최적의 클러스터가 나올 때까지 재분할한다. 그렇게 분할된 마지막 노드들이 최종 클러스터가 된다.

3.2.1 태그 간 유사도 측정 방법

태그 간 유사도 측정 방법은 WordNet-Similarity [11, 12]를 이용한다. WordNet- Similarity는 WordNet [13]을 이용하여 자연어 사이의 의미적 거리를 계산하는 알고리즘이다. WordNet은 자연어 처리를 위한 영어 어휘의 대형 데이터베이스로 명사, 동사, 형용사와 부사를 각각 동의어 집합으로 분류하고 별개의 개념을 표현한다. 동의어 집합은 의미와 어휘 관계로 내부 링크로 연결되어 있는 네트워크 이다. 이 네트워크를 이용하여 WordNet::Similarity에서는 트리 구조로 이루어진 단어들의 관계에서 단어간의 거리를 측정하여 알고리즘을 만들었다. (그림 1)과 같이 단어 사이의 nickel과 money라는 단어가 있을 때 WordNet의 네트워크 상에서 나타나는 종속 관계에 따라 계층적 트리구조를 만들고 거리를 계산한 방법이다. 노드간의 거리를 계산하여 nickel과 money의 거리는 5로 측정된다.

WordNet::Similarity에서 제시하는 방법은 다양한데, 본 논문에서는 사용 가능한 태그들로 실험한 결과 가장 적합하다고 판단한 The Jiang and Conrath 방법 [12]을 이용하였다.



(그림 1) WordNet의 트리 구조를 이용한 거리 계산 예제
(Figure 1) An Example of Similarity Distance Calculation using WordNet Tree Structure

3.2.2 유전 알고리즘

클러스터링 과정에서 각 단계에서의 최적의 비용을 구해서 2분할 하기 위해 유전 알고리즘을 이용하였다. 유전 알고리즘은 병렬적이고 전역적인탐색 알고리즘으로서, 다윈의 적자생존 이론을 기본 개념으로 한다. 유전 알고리즘은 풀고자 하는 문제에 대한 가능한 해들을 정해진 형태의 자료구조 (모집단)로 표현한 다음, 이들을 점차적으로 변형함으로써 점점 더 좋은 해들을 만들어 낸다. n개의 해에서 선택 (selection), 교차 (crossover), 변이 (mutation)의 단계를 거쳐서 새로운 k개의 해를 생성하게 되고 이렇게 만들어진 k개의 해는 기존에 있던 모집단에서 어떤 판단 기준에 의해 k개의 해를 대체해서 새로운 해가 된다. 여기에서 해들을 나타내는 자료구조는 유전자, 이들을 변형함으로써 점점 더 좋은 해를 만들어 내는 과정은 진화로 표현할 수 있다 [14].

3.2.3 유사도 기반 태그 클러스터링 기법

태그 간 의미적 유사도를 이용하여 태그를 의미에 따라 클러스터링 하기 위해서 (그림 2)와 같은 기법을 사용한다. 먼저 N개의 태그를 가진 하나의 클러스터 C가 있을 때 처음 클러스터 C를 queue Q에 넣는다. 그 후 Q에서 C를 pop하여 for문을 시작한다. 먼저 종료조건에 만족하는지 검사하고 더 이상 나눌 필요가 없으면 최종 클러스터 집합인 queue F에 넣는다. 그렇지 않은 경우 클러스터의 모든 태그에 대해 기준 태그와의 거리 (유사도)를 계산하여 최적의 cost를 가진 솔루션을 찾아 클러스터를 2분할 한다. 처음 한번 솔루션을만들고 원소 재조합으로

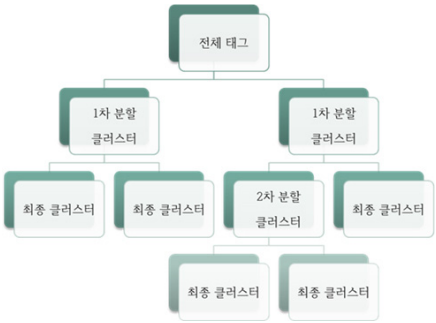
```
Initialize Q.
// Q는 분류하기 전 cluster들의 queue.
Initialize F.
// F는 클러스터링이 완료된 final cluster들의 queue.
N개의 태그를 가진 클러스터 C = {t1, t2, … tn}
while Q is not empty begin
  Pop C from Q
  for each cluster C begin
    Initialize cluster A, B = {}
    if ((태그의 개수 < n(처음 태그 개수)/10) or
        (이전 단계의 cost와 차이가 없을 때))
    begin
      Push C into F
      break
    end
  else
    // 유전 알고리즘을 이용하여 클러스터링
    for each solution S[30] begin // 초기 솔루션
      random put tag A[] from C
      random put tag B[] from C
      cost f(A) = (Σ_{i=1}^{n-1} Σ_{j=i+1}^{n} similarity(ti, tj)) / (n(n-1))/2
      cost f(B) = (Σ_{i=1}^{n-1} Σ_{j=i+1}^{n} similarity(ti, tj)) / (n(n-1))/2
    end
    S[0~14] ← f(A)+f(B)에 따라 성능 우수 솔루션 저장
    S[15~29]← 우수 솔루션들을 이용한 재 조합한 솔루션
    for i from 0 to 100 begin // 두 번째 솔루션부터
      for each solution S[30] begin
        random put tag A[] from C
        random put tag B[] from C
        cost f(A) = (Σ_{i=1}^{n-1} Σ_{j=i+1}^{n} similarity(ti, tj)) / (n(n-1))/2
        cost f(B) = (Σ_{i=1}^{n-1} Σ_{j=i+1}^{n} similarity(ti, tj)) / (n(n-1))/2
      end
      if (cost의 max > 0.12) begin
        최적의 cost(max)가 나온 솔루션의 A, B
          Push A, B into Q
          break
      end
      S[0~14] ← f(A)+f(B)에 따라 높은 솔루션 저장
      S[15~29]← 우수 솔루션들을 이용한 재조합한 솔루션
  end
end
```

(그림 2) 태그 클러스터링 알고리즘
(Figure 2) Tag Clustering Algorithm

두 번째 솔루션을 만드는 것은 무조건 한번 수행하게 되고 그 후에는 cost의 최대값의 조건에 따라 종료하게된다.종료 조건을 만족하여 분할이 완료된 클러스터들은 다시 Q에 넣는다. 2분할 과정을 반복할 때 클러스터 내의 태그 수는 랜덤이기 때문에 각 클러스터의 태그 개수는 다양해진다. 클러스터링 과정에서 최적의 cost를 알아

$$\frac{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} similarity(t_i, t_j)}{\frac{n(n-1)}{2}}$$

(그림 3) 각 클러스터의 cost 계산
(Figure 3) Calculation of Cost of Clusters



(그림 4) 클러스터링 결과 예시
(Figure 4) An Example of Clustering Result

내기위하여 유전 알고리즘을 사용하여 솔루션 30개를 생성하는데 태그의 개수에 상관없이 랜덤으로 2개의 클러스터에 넣는다. 나누어진 솔루션의 비용을 계산하여 솔루션 30개 전체의 중앙값을 구하고 그보다 낮은 솔루션은버린다. 남은 솔루션으로 랜덤으로 양쪽의 태그를 교환하여 (mutation) 새로운 솔루션을 만들어 30개를 채운다. 이 과정을 반복하다가 솔루션의 최대 cost가 0.12보다 크면 종료하고 최대 cost를 가진 솔루션을 최적 클러스터로 정한다.

태그간의 유사도를 계산하기 위해 [11]에서 제안된 API를 이용한다. 먼저 각 클러스터에 있는 모든 태그에 대해 각 태그 쌍의 유사도를 받아와서 평균을 내서 cost를 계산한다. 각 태그 쌍의 유사도를 계산하는 방법은 (그림 3)과 같다.

특정 태그를 기준으로 계산하지 않고 모든 태그에 대해 각 태그 쌍의 유사도를 전부 계산하여 클러스터링의 정확도를 높였다. 이 방법으로 모든 솔루션에대해 각 클러스터의 cost를 계산하여 가장 높은 cost를 가진 솔루션을 선택하고 각 솔루션에 나누어진 클러스터가 종료조건에 맞으면 F에 push한다. 종료조건에 부합하지 않는 클러스터는 다시 Q에 넣고 앞의 과정을 반복한다. 한 클러스터의 태그 개수가 (전체태그)/10개 이하일 때를 종료조건으로 두었다.

제시된 알고리즘에 따라 클러스터링을 완료하면 (그림 4)와 같이 이진 트리의 형태를 띠고 단말 노드들이 최종 클러스터가 된다.

friend	birthday	family	food
restaurant	graduation	trip	coffee
university	airport	vacation	flight
cake	mom	school	interview

(그림 5) 테스트 태그 집합
(Figure 5) Test Tag Set

3.2 태그 클러스터링 결과

태그 클러스터링을 진행하면 다음과 같은 결과를 얻을 수 있다. 개인 사진에 있는 임의의 태그 16개를 분할하는 과정이다. 태그 집합은 (그림 5)와 같다. 테스트를 위해 16개의 적은 개수의 태그를 분할하기 때문에 종료조건을 변경하여 한 클러스터에 4개 이상의 태그가 들어가도록 하였다.

앞의 알고리즘에 따라 전체 태그의 분할을 시작하면 먼저 queue Q에 전체 태그를 넣고 pop하여 솔루션 S를 만든다. 그리고 한번의 재조합을 거치고 cost의 최대값이 0.12보다 큰지 확인하게 된다. 크다는 것이 확인되면 표 2와 같이 선택된 솔루션을 출력하고 각 클러스터의 태그 개수를 확인하여 3보다 크면 다시 Q에 넣는다. 3보다 작으면 최종 queue F에 넣어서 최종 클러스터가 된다.

(그림 6)에서 보이는 것처럼 전체 클러스터 과정은 queue로 관리하게 된다. 처음 2분할 후 재분할하기 위해 Q에 저장된 클러스터 A1을 꺼내서 두 번째 분할을 시작한다. (그림 7)과 같이 최적의 cost를 가진 15번째 솔루션을 선택하여 분할하고 각 클러스터의 태그 개수를 비교하여 4보다 작으면 F에 넣어서 최종 클러스터로 확정한다. B1을 꺼내 과정을 반복한다. (그림 8) 9번 솔루션을선택한 하여 태그 개수를 비교하여 queue에 넣는다. 분할된 두 클러스터모두 4개 이하의 태그를 가지고 있기 때문에 F에 들어가 최종 클러스터가 된다.

그리고 마지막으로 A2를 분할하여 두 클러스터 모두 4개 이하가 되는 솔루션을 F에 넣으면 Q에 남아있는 클러스터가 없기 때문에 클러스터링이 종료된다. (그림 9) 클러스터링이 종료되면 (그림 10)과 같이 최종 클러스터를 출력한다. 이 때 최종 클러스터는 유사도가 높은 태그들끼리 모이게 된다.

<표 1>는 유사도가0.1보다 높은 태그 쌍들을 보여준다. 제공하는 API에 따라 정해진 것이며 개인이 직관적으로 생각할 수 있는 유사도와는 조금 차이를 보이지만 <표 1>의 유사도를 기반으로 클러스터링 결과를 살펴보

Selected solution	
A1: friend restaurant university birthday graduation mom family trip flight interview	B1: cake airport vacation school food coffee
sizeof(A1) > 4 push Q	sizeof(B1) > 4 push Q
first queue Q	**final queue F**
B1	
A1	

(그림 6) 첫 번째 전체 클러스터 분할 과정
(Figure 6) Fist Step: Partition of Entire Cluster

Selected solution	
A2: restaurant university birthday mom family friend	B2: interview graduation trip flight
sizeof(A2) > 4 push Q	sizeof(B2) <= 4 push F
first queue Q	**final queue F**
A2	
B1	B2

(그림 7) 두 번째 A1 분할 과정
(Figure 7) Second Step: Partition of Cluster A1

Selected solution	
A3: cake food coffee	B3: airport vacation school
sizeof(A3)<= 4 / push F	sizeof(B3) <= 4 / push F
first queue Q	**final queue F**
	B3
	A3
A2	B2

(그림 8) 세 번째 B1 분할 과정
(Figure 8) Third Step: Partition of Cluster B1

Selected solution	
A4: restaurant university birthday friend	B4: mom family
sizeof(A4)<= 4 / push F	sizeof(B4) <= 4 / push F
first queue Q	**final queue F**
	A4
	B4
	B3
	A3
	B2

(그림 9) 마지막 A2 분할 과정
(Figure 9) Last Step: Partition of Cluster A2

final cluster: interview graduation trip flight
final cluster: cake food coffee
final cluster: airport vacation school
final cluster: restaurant university birthday friend
final cluster: mom family

(그림 10) 최종 클러스터
(Figure 10) Final Clusters

(표 1) 유사도가 높은 태그 쌍
(Table 1) Tag Pairs that have High Similarities

tag1	tag2	similarity
Restaurant	university	0.1095
Restaurant	school	0.1598
University	family	0.1091
University	school	0.1561
Cake	food	0.2446
Graduation	trip	0.1119
Mom	family	0.1067
Family	school	0.1646
Family	flight	0.117
Trip	flight	0.2282
School	flight	0.1076
Food	coffee	0.4078

면 유사도가 가장 높은 (coffee, food) / (cake, food) 가 한 클러스터에 있고 (restaurant, university) / (trip, flight) / (graduation, trip) / (mom, family) 도 한 클러스터에 묶인 것을 볼 수 있다.

이렇게 분할한 클러스터링 과정을 전체적으로 보면 (그림 11)과 같이 계층적 형태를 가진다. 이 과정에서트리의 밸런스는 상관없이 분할을 진행하며 가장 마지막에 있는 단말 노드 클러스터들이 최종 클러스터가 된다.

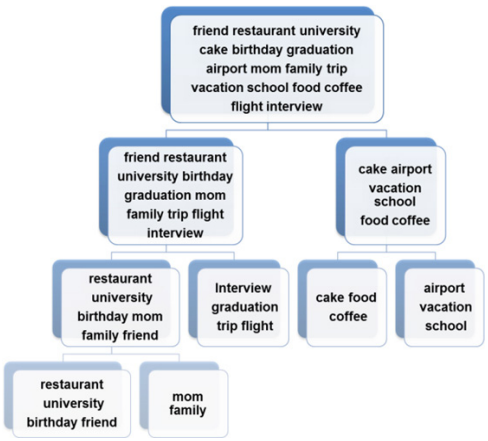
4. 태그 추천

4.1 개요

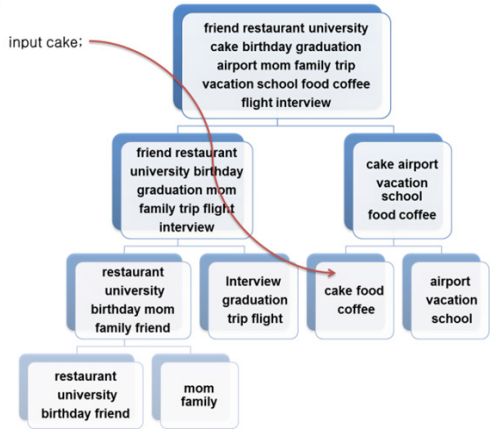
태그를 의미적 유사도에 따라 클러스터링하여 개인 사용자의 초기 태그 클러스터를 만들면, 이에 따라 개인 사용자가 사진을 업로드 할 때 태그를 추천해줄 수 있다. 사용자가 새로운 태그를 입력하면 기존 태그 클러스터에서 초기 태그와의 연관 관계를 파악하여 적합한 태그를 추천하는 방법을 제안한다.

4.2 태그 추천 기법

기존 사진과 태그의 클러스터링을 1차적으로 진행하여 클러스터링 결과를 데이터베이스에 저장한 후에 사용자가 사진을 새로 업로드하면서 초기 태그를 입력하면 태그 클러스터에서 일치하는 태그를 찾아 그 태그가 있는 클러스터에 포함된 다른 태그 중에 유사도가 가장 높



(그림 11) 전체 클러스터링 과정과 결과
(Figure 11) Overall Clustering Process



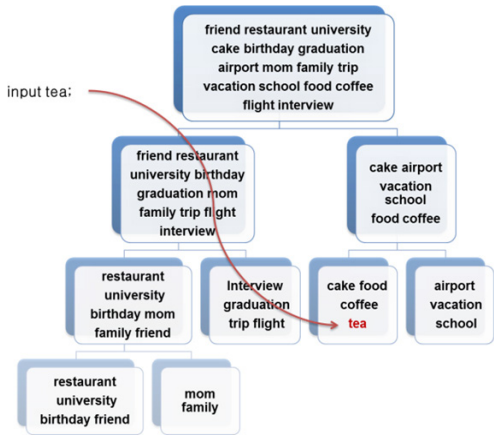
(그림 12) 태그 추천 과정 1
(Figure 12) Tag Recommendation 1

은 태그 5개를 임의로 추천한다.

(그림 12)와 같이 cake이라는 태그가 입력되면 클러스터의 태그와 비교하여 일치하는 태그가 포함된 클러스터를 찾아 같은 클러스터에 있는 food, coffee를 사용자에게 추천한다.

입력 받은 태그가 기존 데이터베이스에 없는 단어라면 기존 태그들과의 유사도에 따라 새로 클러스터링하여 클러스터를 만들고, 그 결과에 따라 입력 받은 태그가 포함된 클러스터에 있는 다른 태그들을 추천해준다.

(그림 13)과 같이 기존 클러스터에 없는 tea라는 태그가 입력되면 새로 클러스터링을 수행한다. <표 2>에서 보이는 것과 같이 cake, food, coffee와의 유사도가 높기



(그림 13) 태그 추천 과정 2
(Figure 13) Tag Recommendation 2

(표 2) 'tea'와 다른 태그와의 유사도
(Table 2) Similarity Other Tags with Tag 'tea'

tag1	tag2	similarity
tea	cake	0.0718
tea	food	0.2131
tea	coffee	0.2035

때문에 그림에서 보이는 클러스터에 들어가게 되고 이에 따라 cake, food, coffee를 추천 태그로 제시하게 된다.

이렇게 기존 태그의 클러스터를 기반으로 태그를 추천하게 되면 개인 사용자가 자주 사용하는 익숙한 단어를 추천하게 되어 새로운 태그나 사진이 업데이트되어도 분류 기준이 흐트러지지 않게 된다. 또한 비슷한 뜻의 단어들은 어느 정도 정해진 단어로 사용하게 되어 유의어의 혼란을 막을 수 있다.

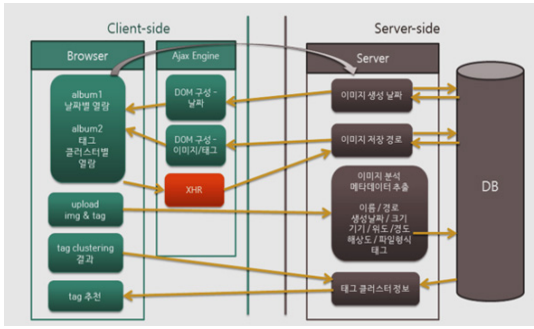
5. 구 현

5.1 구현 환경

개인 사진 콘텐츠 관리 시스템을 구축하기 위하여 웹 서버 (LAMP)를 만들고 Mysql로 데이터베이스를 구축하였다. 또한 PHP, Javascript, HTML5를 이용하여 사용자용 웹 어플리케이션을 구현하였다.

5.2 시스템 개요

본 논문에서는 사용자가 PC나 스마트 기기 (안드로이



(그림 14) 시스템 전체 구성
(Figure 14) Overall System Structure

드OS)에 있는 개인의 사진을 웹 서버에 업로드하여 저장하고, 업로드 된 사진을 자동으로 분류하여 사용자가 좀 더 편리하게 사진 콘텐츠를 관리할 수 있도록 하는 개인 사진 관리 시스템을 구축하였다. 웹 서버에 사진의 정보와 태그를 저장할 수 있는 데이터베이스를 구축하였고, 웹 어플리케이션을 만들어 PC와 스마트 기기에서 시스템을 이용할 수 있게 하였다.

사용자는 개인의 사진을 효율적으로 관리할 수 있으며 주제별 자동 분류를 통해 좀 더 빨리 원하는 사진을 검색할 수 있다. 본 시스템은 (1) 사진 업로드 및 저장 (2) 태그 추천 (3) 태그 클러스터링 및 사진 분류 (4) 분류별 사진 열람 네 파트로 구분할 수 있다.

개인 사진의 효율적인 관리를 위해 사용자가 사진과 태그를 저장하면 데이터베이스에 저장되고 제안한 태그 클러스터링 기법을 적용하여 태그를 1차 분류한다. 분류된 태그에 따라 사진을 2차 분류하여 그 분류에 따라 사용자에게 보여준다. 또한 사진을 업로드할 때 태그를 함께 지정하여 업로드 하고 사진 미리보기와 함께 비슷한 태그를 추천 받아 한번 더 태그를 입력할 수 있다.

사용자 웹 어플리케이션에서 사진을 서버에 업로드하고 사진의 정보를 데이터베이스에 저장하며 이 정보에 따라 사진을 날짜별/태그클러스터별로 열람할 수 있다. Ajax (Asynchronous JavaScript and XML) 기술을 이용하여 DOM (Document Object Model)을 구성하여 메뉴와 결과를 볼 수 있게 하였다. 태그를 클러스터링 할 때는 클러스터링 과정을 보여주고 클러스터의 정보를 테이블에 저장하고 사진 열람 시 불러온다. 전체 시스템의 클라이언트-서버 구조는 (그림 14)와 같다.

5.3 시스템 설계

5.3.1 사진 업로드 및 저장

사용자의 이미지를 웹 서버에 업로드 하여 저장하고 사진에서 필요한 EXIF 정보를 추출하여 데이터베이스에 저장한다. 또한 사진을 업로드 할 때 태그를 입력 받아 따로 저장한다. HTML5를 이용하여 최대 10장까지 다중 업로드가 가능하며 업로드 한 이미지는 지정된 폴더에 저장되고 한 폴더에 1000개 이상의 이미지가 있으면 새로 폴더가 생성된다. 파일이름이 중복되는 경우를 대비해 업로드 될 때 랜덤 이름을 생성하여 파일이름을 변경해 저장한다. 또한 사용자가 사진을 업로드 할 때 태그를 함께 입력하여 태그 정보도 저장하게 된다. 사진 하나에 최대 열 개의 태그를 지정할 수 있다. EXIF 정보와 태그 정보를 한 테이블에 저장하여 참조할 수 있게 하였다

사진을 업로드 하면 <표 3>의 테이블에 사진의 EXIF와 태그 정보가 저장된다. 태그를 입력할 때는 콤마를 구분자로 사용하여 입력해야 하며 입력한 그대로 사진 정보 테이블에 저장된다.

(표 3) 저장한 사진 정보 테이블 필드
(Table 3) Database Table for Photo Information

Field	설명	Type	비고
name	파일이름	char(50)	
fullpath	파일경로	char(100)	
date	파일생성날짜	char(30)	yyyy-mm-dd
size	파일크기	float(8,2)	
device	기기	char(50)	
location1	위도	Float	
location2	경도	Float	
resolution	해상도	Float	
format	파일형식	char(50)	확장자
tag	태그	char(100)	콤마로 구분

5.3.2 태그 추천

태그 클러스터링 결과에 따라 사용자가 사진을 업로드 할 때 태그를 추천해준다. 입력된 태그를 쿼리의 조건으로 주어 일치하는 태그를 찾고 그 태그의 클러스터 아이디를 받아온 후 클러스터 아이디를 조건으로 select하여 태그를 출력하여 사용자에게 보여주어 태그를 추천한다.

일치하는 태그가 없으면 클러스터링 버튼을 주어 클

릭하면 새로운 태그를 포함한 후 다시 태그 클러스터링을 하게 된다. 이렇게 만들어진 클러스터를 토대로 앞의 과정을 반복하여, 새로운 태그와 같은 클러스터에 있는 다른 태그를 추천 태그로 사용자에게 보여준다.

5.3.3 태그 클러스터링

전체 태그를 반복 2분할 하여 최종 분할된 태그 클러스터 정보를 클러스터 테이블에 저장한다. 클러스터 테이블은 <표 4>와 같이 태그와 클러스터 정보를 포함하고 있다. 클러스터링을 시작하면 사진 정보 테이블에서 한 번에 태그 정보를 가져와 클러스터 테이블에 저장한다. 그리고 알고리즘에 따라 클러스터링을 수행한 후 각 태그가 포함된 클러스터의 번호를 insert한다.

(표 4) 클러스터 테이블 필드
(Table 4) Database Table for Cluster

Field	설명	Type	비고
tag	태그이름	char(50)	
cluster_id	클러스터번호	int(10)	

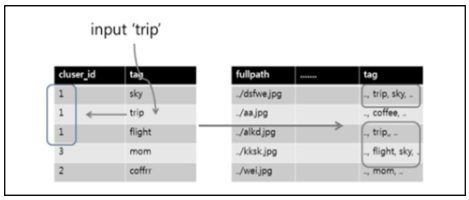
클러스터링 과정은 2분할을 반복하여 계층적 형태를 띄게 되어 레벨이 정해진다. 각 2분할 과정의 솔루션에 따라 선택되는 클러스터의 상황을 사용자에게 보여주어 클러스터링 과정에서 레벨 별 진행상황을 모니터 할 수 있다. 클러스터링이 완료되면 최종 클러스터와 queue의 상황을 볼 수 있다.

5.3.4 유사도 적용

태그간의 의미적 유사도를 위해 WordNet::Similarity [13]에서 제공하는 API를 이용하여 데이터베이스에서 태그 2개를 가져와서 입력하여 유사도를 계산한 후, 계산된 유사도를 반환 받아 테이블에 저장한다. 각 태그간 유사도 테이블은 <표 5>와 같다.

(표 5) 유사도 저장 테이블 필드
(Table 5) Database Table for Similarity

Field	설명	Type
tag1	태그1	char(50)
tag2	태그2	char(50)
similarity	유사도	float(50)



(그림 15) 태그 클러스터 별 열람 프로세스
(Figure 15) Retreival of Photos according to Tag Clusters

5.3.5 사진 열람

본 시스템에서는 사진을 분류된 상태로 열람하는 방법을 두 가지 제공한다. 먼저 데이터베이스에 저장된 날짜 정보를 이용하여 서버에 있는 이미지를 날짜순으로 정렬하여 열람한다. Ajax 기술을 이용해 날짜를 선택하면 페이지 전환 없이 이미지의 태그와 함께 이미지를 열람할 수 있다. 두 번째로, 제안한 태그 클러스터링 기법을 이용하여 분류된 사진들을 열람하는 방법이 있다. 각 클러스터를 대표하는 태그가 임의로 타이틀 (메뉴)로 지정되고, 날짜 별 열람과 마찬가지로 Ajax 기술을 이용하여 페이지 전환 없이 대표 태그를 선택하여 열람할 수 있다. 이에 따라 각 클러스터의 분류기준을 어느 정도 파악할 수 있으며, 사용자가 직관적으로 어떤 분류인지 알 수 있다.

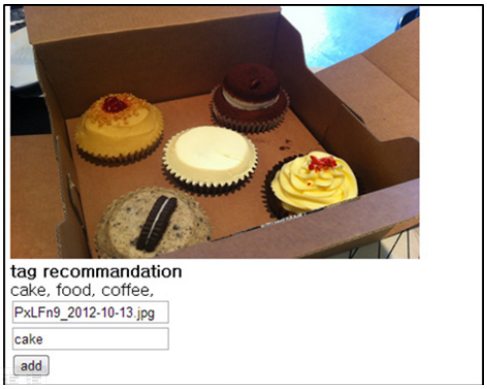
첫번째 열람 방법은 사진 정보 테이블에서 날짜를 참조하여 해당 날짜의 사진 주소와 태그를 리턴받고 페이지에 보여준다.

두 번째 열람 방법은 태그 클러스터 정보에 따라 사진을 분류하여 보여주는데 사진에 여러 개의 태그가 있기 때문에 사진은 한가지 분류에만 속하지 않고 여러 분류에 다양하게 속하게 된다. 또한 각 클러스터 별 대표 태그를 추출하여 클러스터의 대표 단어로 두어 분류 별 주제를 파악하는데 도움을 준다.

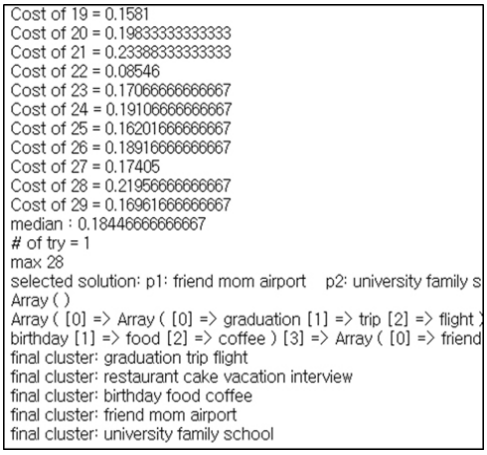
(그림 15)와 같이 클러스터 테이블에서 클러스터 번호를 참조하여 해당하는 태그를 받아온 후 출력하고, 해당 태그를 사진 정보 테이블 <표 3>에서 매치하여 찾은 후 그 컬럼의 전체 주소 값 (fullpath)를 받아서 처리하여 사용자에게 보여준다.

5.4 실행 결과

각 기능별 실행 화면은 다음과 같다.



(그림 16) 태그 추천 화면
(Figure 16) Screen of Tag Recommendation



(그림 17) 태그 클러스터링 과정과 결과 화면
(Figure 17) Screen of Tag Clustering and Results

사진을 업로드 하면 (그림 16)에서처럼 태그 추천을 한다. 사용자가 업로드 한 사진을 썸네일로 보여주고 추천 태그를 그 밑에 출력하고 태그 입력 칸을 두어 사용자가 추천 태그를 보고 입력할 수 있도록 한다.

(그림 17)과 같이 태그 클러스터링의 결과를 볼 수 있다. 각 솔루션의 cost를 확인하고 어떤 솔루션이 선택되어 클러스터가 됐는지 알 수 있다. 또한 마지막에는 최종 분할된 클러스터 결과를 출력하여 보여준다.

업로드 되어 분류된 사진 날짜별, 태그별로, 그리고 클러스터별로 열람이 가능하다. (그림 18)에 날짜별로 사진을 열람하는 경우 (a)와 클러스트별로 열람한 예 (b)를 보인다. (그림 18) (b)에서와 같이 음식과 관련된 사진들이 함께 분류되어 보여지는 것을 알 수 있다. 사진날짜별 열

(a) 날짜별 사진 열람
(a) Photo Retreival by Data

(b) 클러스터별 사진 열람
(b) Photo Retreival by Clusters

(그림 18) 사진 열람 화면
(Figure 18) Screen of Photo Retreival

람 화면에서는 날짜를 메뉴로 제공하고 사진과 그 사진의 태그를 함께 보여준다. 태그 클러스터별 열람에서는 클러스터에 있는 태그를 출력하고 각 태그를 포함하고 있는 모든 사진을 함께 보여줌으로써 클러스터내의 모든 사진을 한꺼번에 열람하게 되는 것이다.

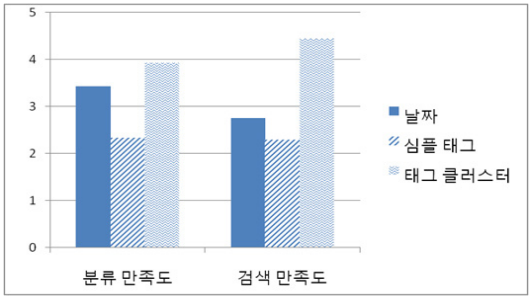
5.5 사용성 평가

5.5.1 개요

사진 콘텐츠 분류 방법에 따라 사용성 평가를 진행하였다. 사진의 열람, 분류, 검색 기능을 평가하기 위해 기존에 존재하는 (1) 메타데이터(날짜)를 이용한 사진 분류와 (2) 심플 태그를 이용한 분류 그리고 본 논문에서 제시한 (3) 태그 클러스터링를 이용한 사진 분류 방법을 비교하여 사용자의 만족도를 5점 척도로 측정한다. 또한 태그 추천 기능을 평가하여 추천된 태그가 적절한지 조사하고 태그 추천의 유용성 5점 척도로 측정한다. 설문조사 개요는 다음 <표 6>과 같다.

(표 6) 사용성 평가를 위한 설문조사 개요
(Table 6) Summary of Survey for Usability Evaluation

대상	스마트 기기를 이용하는 20명의 개인 사용자	
기간	2012.11.30	
방법	Flickr를 이용하여 날짜 별 열람과 심플 태그 별 열람을 수행하며, 본 논문에서 구현한 시스템을 이용하여 태그 클러스터 별 열람을 수행 1. 각 시스템에 개인의 사진 콘텐츠를 태그와 함께 100장씩 업로드. 사진의 종류는 일상생활 등을 담은 개인 사진으로 업로드. 2. 업로드 된 사진을 각 열람 방법으로 열람 3. 원하는 사진 한 장을 검색 4. 새로운 사진을 업로드 하고 태그를 추천 받음	
항목	**사진 분류, 검색 기능 평가**	검색
		1. 원하는 사진을 검색하기 쉬운가? 2. 검색하는 데 소요되는 시간은 만족스러운가?
		전체 만족도
		1. 분류의 기준이 만족스러운가? 2. 사진을 분류, 열람한 결과에 대한 만족도는 어떠한가? 3. 이 방법이 나에게 유용한가?
	태그 추천 기능 평가	1. 추천된 태그가 적절한가? 2. 추천된 태그가 익숙하고 이해하기 쉬운가? 3. 태그 추천이 유용하다고 생각하는가?



(그림 19) 각 방법 별 분류, 검색 만족도
(Figure 19) Satisfaction Results with respect to Categorization Method

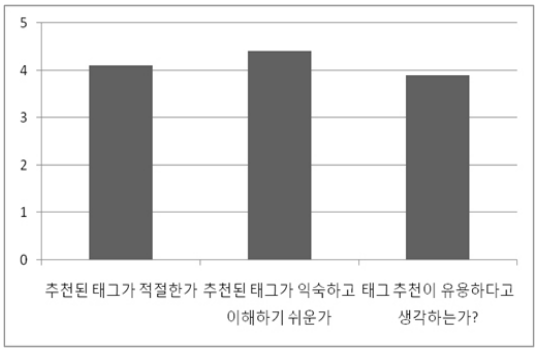
5.5.2 결과 분석

사진 분류와 검색 결과의 만족도 평가의 결과를 평균 내어서 그래프로 보면 다음 (그림 19)와 같다. 각 시스템 방법 별로 만족도의 차이가 크게 나타났다. 사진 분류 방법에 대한 만족도는 날짜와 본 논문의 태그 클러스터 방법이 만족도가 높았다.

날짜 별로 분류하는 방법도 사용자가 유용하게 사용할 수 있다는 것이다. 하지만 태그 클러스터 방법이 더 높은 만족도를 보여 사진을 태그를 이용하여 주제별로 분류하는 것을 유용하게 사용할 것이라는 결과를 내었다.

다음으로 검색 기능에 대한 만족도는 분류보다 더 큰 차이를 보였다. (그림 19)에서 보이는 것처럼 태그 클러스터 방법을 이용한 검색이 다른 방법들의 두 배의 만족도 점수를 얻었다. 사용자가 특정 사진을 검색할 때에 날짜를 기억하기 힘들기 때문에 주제별로 분류된 태그 클러스터 방법이 찾는데 시간이 덜 걸린다는 의견이었다. 또한 심플 태그를 이용하면 타겟 사진과 비슷한 사진이 너무 많이 나와서 오히려 눈으로 스캔해서 찾기가 더 어렵다. 이에 따라 원하는 사진을 쉽고 빠르게 하기 위해서는 주제별 분류가 반드시 필요하다.

태그 추천에 관한 연구가 늘고 있지만 실제 상용화된 서비스에서 태그 추천 기능을 제공하거나 많은 사람들이 이용하는 것은 아직 거의 없다고 볼 수 있다. 그래서 사용자가 덜 익숙해하는 경향이 있었다. 하지만 (그림 20)과 같이 태그 추천 기능에 대한 만족도와 유용성 점수는 모두 4점 정도로 높게 나타났다. 특히 본인이 입력한 태그를 추천해주는 기능이라 익숙하고 이해하기 쉬우며 비슷한 사진에 같은 단어를 넣을 수 있어서 좋다는 의견이 있었다. 또한 이 기능의 유용성에 대한 점수도 4점으로 높게 나타났다. 앞으로 태그 추천 기능을 개인 사진 관리 서비스에 추가하면 좋은 반응을 얻을 것이라 예상된다.



(그림 20) 태그 추천 기능의 만족도와 유용성
(Figure 20) Satisfction and Usability of Tag Recommendation Capability

6. 결론 및 향후 연구

개인 사진 콘텐츠의 분류를 위해 태그의 의미적 유사도을 이용하여 계층적 및 구조적으로 태그를 클러스터링 하는 기법을 제안하였다. 이는 태그를 가진 사진 콘텐츠를 이미지 분석 과정 없이 메타데이터만으로 의미적 유사성에 따라 분류할 수 있게 도와준다. 또한 사진의 의미적인 분류를 통해 사용자가 원하는 사진을 검색하기가 용이해졌다. 현재 연구에서는 다른 메타데이터를 배제하고 태그만을 이용하였는데 사진의 기본 메타데이터를 함께 이용하여 사진을 분류하게 되면 더 정확한 결과를 얻을 수 있을 것이라 예상한다. 개인의 데이터를 관리하는 것이기 때문에 각 개인 사용자의 특성을 따르는 사용자 맞춤형 분류에 대한 연구도 필요하다. 이에 따라 의미적 유사도를 계산하는 방법도 일반적인 단어와 다르게 적용이 되어야 하기 때문에 각 개인의 데이터를 이용한 학습 기능으로 유사도를 계산하여 적용하면 더 좋은 결과를 얻을 수 있을 것이다.

또한 본 논문에서는 초기 태그를 기반으로 개인 사용자에게 태그를 추천해주어 데이터가 많아질수록 개인 사용자에게 맞는 더욱 정확한 분류 기준을 만들 수 있게 된다. 하지만 사진과 태그를 넣을 때 사용자가 반드시 초기 태그를 주어야 하는 불편함이 있기 때문에 이에 따라 내용 기반의 태그 분류와 태그 추천 연구를 진행하면 초기 태그 없이도 태그를 추천 받고 사진을 자동으로 분류할 수 있을 것이라 예상한다. 평가 방법에 대해서는 현재는 설문조사에 기반하여 사용성 평가를 진행하고 있지만 좀 더 계량적이고 분석적인 평가방법의 도입이 필요하다. 예를 들면 분류된 태그들이 제대로 클러스터 되었는지를 평가하는 정확률 (precision)과 얼마나 클러스터에 포함되었는지를 판단하는 재현률 (recall)등의 계산법(metric)을 도입하는 것을 향후 연구로 계획하고 있다.

참 고 문 헌(Reference)

[1] Yu Zheng, Longhao Wang, Ruochi Zhang, Xing Xie and Wei-Ying Ma, "GeoLife: Managing and Understanding Your Past Life over Maps", Mobile Data Management, pp. 211-212, 2008.

[2] Song, Tae Houn, Jeong, Soon Mook, Kim, Hyung Min, Kwon, Key Ho, Jeon, Jae Wook, "Direct Share: Photo Management System Based on Round-robin Concept-driven User Preference Feedback", KSII, the Korean Society for Internet Information, pp. 1346-1367, 2011.

[3] Picasa Web Album [Online]. Available: https://picasaweb.google.com/

[4] Flickr [Online]. Available: http://www.flickr.com/

[5] Till Quack , Bastian Leibe , Luc Van Gool, "World-scale Mining of Objects and Events from. Community Photo Collections", Proceedings 7th ACM international conference on image and video retrieval - CIVR '08, pp. 47-56, 2008.

[6] Hong M. Chen, Ming H. Chang, Ping C. Chang, Ming C. Tien, Winston H. Hsu, Ja L. Wu, "SheepDog – Group and Tag Recommendation for Flickr Photos by Automatic Search-based Learning", In MM '08: Proceeding of the 16th ACM international conference on Multimedia, pp. 737-740, 2008

[7] Mianwei Zhou, Shenghua Bao, Xian Wu and Yong Yu, "An Unsupervised Model for Exploring Hierarchical Semantics from Social Annotations", ISWC'07/ASWC'07, pp. 680-693, 2007.

[8] Grigory Begelman et. Al, "Automated Tag Clustering: Improving search and exploration in the tag space", In Proc. of the Collaborative Web Tagging Workshop at WWW'06, 2006.

[9] Onur Kucuktunc, Sare G. Sevil, A. Burak Tosun, Hilal Zitouni, Pinar Duygulu, and Fazli Can, "Tag Suggestr: Automatic Photo Tag Expansion Using Visual Information for Photo Sharing Websites", SEMANTIC

MULTIMEDIA, pp. 61-73, 2008.
[10] Hyunwoo Kim, Kangpyo Lee, Hyung-Joo Kim, "Tag Recommendation Algorithms in Tagging System", Journal of KIISE: Computing Practices and Letters, Vol. 16, No. 9, pp. 927-935, 2010.
[11] WordNet::Similarity [Online]. Available:http://marimba.d.umn.edu/cgi-bin/similarity/similarity.cgi
[12] Siddharth Patwardhan, Satanjeev Banerjee, Ted Pedersen, "Using Measures of Semantic Relatedness for Word Sense Disambiguation", Springer, Computational Linguistics and Intelligent Text Processing, pp 241-257, 2003.
[13] Christiane Fellbaum, "WordNet", Springer, Theory and Applications of Ontology: Computer Applications, pp 231-243, 2010.
[14] Genetic algorithm [Online]. Available: http://en.wikipedia.org/wiki/Genetic_algorithm.

◐ 저 자 소 개 ◑

원 지 현(Ji-Hyeon Won)
2011년 숙명여자대학교 멀티미디어과학과 졸업(학사)
2013년 숙명여자대학교 대학원 멀티미디어과학과 졸업(석사)
관심분야 : 웹 미디어
E-mail : blessing_of@naver.com

이 종 우(Jongwoo Lee)
1990년 서울대학교 컴퓨터공학과(학사)
1992년 서울대학교 컴퓨터공학과(석사)
1996년 서울대학교 컴퓨터공학과(박사)
1996년~1998년 현대전자㈜ 정보시스템 사업본부 과장
1999년~1999년 현대정보기술㈜ 책임연구원
1999년~2002년 한림대학교 정보통신공학부 조교수
2002년~2003년 광운대학교 컴퓨터공학부 조교수
2003년~2004년 아이닉스소프트㈜ 개발이사
2004년~현재 숙명여자대학교 정보과학부 멀티미디어과학전공 조교수
2008년 뉴욕주립대 스토니브룩 Research Scholar
2012년~현재 숙명여자대학교 관리정보처장
관심분야 : Mobile System Software, Storage Systems, Computational Finance, Cluster Computing, Parallel and Distributed Operating Systems, and Embedded System Software.
E-mail : bigrain@sm.ac.kr

박 희 민(Heemin Park)
1993년 서강대학교 전자계산학과 졸업(학사)
1995년 서강대학교 대학원 컴퓨터공학과 졸업(석사)
2006년 University of California, Los Angeles 대학원 전자공학과 졸업(박사)
1995년~2010년 삼성전자 수석
2010년~2012년 숙명여자대학교 멀티미디어과학과 조교수
2012년~2013년 연세대학교 컴퓨터과학과 연구교수
2013년~현재 상명대학교 컴퓨터소프트웨어공학과 조교수
관심분야 : 웹기반정보시스템, 유비쿼터스 컴퓨팅, 멀티미디어, 사이버물리 시스템
E-mail : heemin@smu.ac.kr