(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0008699
(43) 공개일자 2019년01월25일

(51) 국제특허분류(Int. Cl.)
G06F 16/00 (2019.01) *G06F 17/18* (2006.01)
(52) CPC특허분류
G06F 16/5846 (2019.01)
G06F 17/18 (2013.01)
(21) 출원번호 10-2017-0090390
(22) 출원일자 2017년07월17일
심사청구일자 2017년07월17일

(71) 출원인
경희대학교 산학협력단
경기도 용인시 기흥구 덕영대로 1732 (서천동, 경희대학교 국제캠퍼스내)
(72) 발명자
이영구
대전광역시 서구 청사서로 11, 106동 302호 (월평동, 무지개아파트)
키파야트 울아 칸
경기 용인시 기흥구 덕영대로 1732 경희대학교 국제캠퍼스
(뒷면에 계속)
(74) 대리인
특허법인도담

전체 청구항 수 : 총 15 항

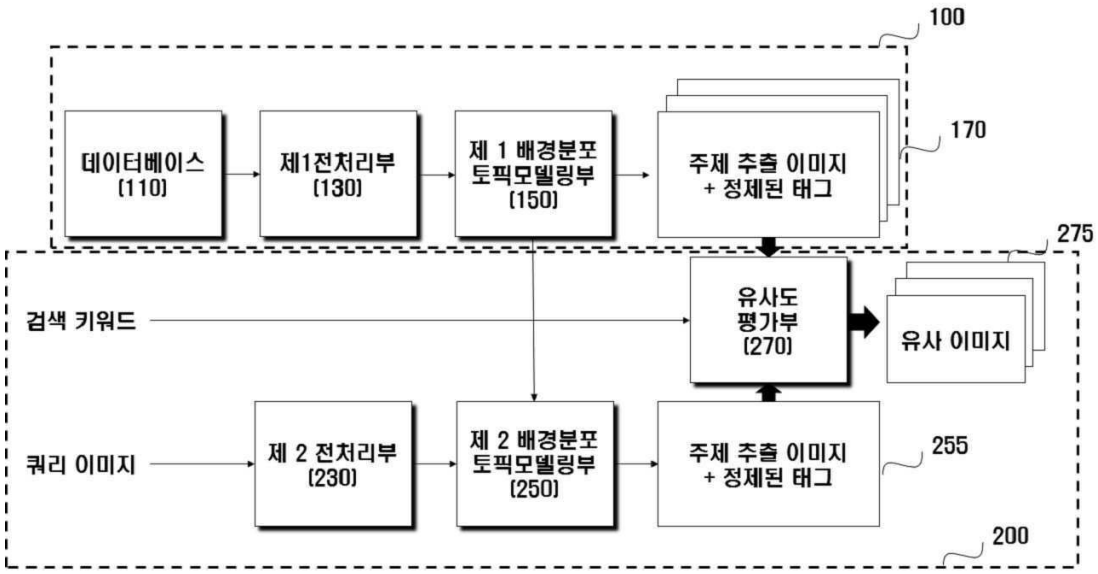
(54) 발명의 명칭 **토픽 모델링 기반 시맨틱 이미지 검색 방법, 시스템 및 컴퓨터 프로그램**

(57) 요 약

본 발명은 시맨틱 이미지 검색 방법, 시스템 및 컴퓨터 프로그램에 관한 것이다. 본 발명은 데이터베이스에 저장된 하나 이상의 데이터베이스 이미지에서 쿼리이미지 또는 검색 키워드에 대응되는 이미지를 검색하는 방법에 있어서, 하나 이상의 태그가 맵핑된 상기 데이터베이스 이미지로부터 하나 이상의 시각 단어 및 하나 이상의 텍스

(뒷면에 계속)

대 표 도 - 도1



트 단어를 획득하는 전처리 단계, 상기 시각 단어와 상기 텍스트 단어를 이용하여 주제별 시각 단어 분포($\Phi_{v,t}$), 주제별 텍스트 단어 분포($\Phi_{w,t}$) 및 배경의 시각 단어 분포(Ω)를 포함하는 배경분포 토픽모델의 모델 파라미터를 추정하는 단계, 상기 추정된 모델 파라미터를 포함하는 상기 배경분포 토픽모델을 이용하여 상기 데이터베이스 이미지의 주제 분포(θ_d)를 모델링하는 단계, 상기 주제별 텍스트 단어 분포($\Phi_{w,t}$)와 이미지 내 주제 비율(θ_{td})을 이용하여 상기 데이터베이스 이미지에 대한 상기 태그의 상응도를 산출하고, 그에 따라 상기 태그를 정제하는 단계, 상기 쿼리 이미지를 포함하는 검색 요청이 입력되면 상기 쿼리 이미지로부터 하나 이상의 시각 단어를 획득하고, 상기 추정된 모델 파라미터를 포함하는 상기 배경분포 토픽모델을 이용하여 상기 쿼리 이미지의 주제 분포(θ_q)를 모델링하는 단계, 상기 데이터베이스 이미지의 주제 분포(θ_d)와 상기 데이터베이스에서 상기 쿼리 이미지의 주제 분포(θ_q)를 이용하여 상기 데이터베이스 이미지의 유사도를 평가하는 단계를 포함하는 것을 일 특징으로 한다. 이러한 본 발명에 따르면, 수백만 개의 이미지를 텍스트 또는 이미지를 이용하여 효율적으로 검색할 수 있다.

(72) 발명자

무하마드 알리

경기도 용인시 기흥구 덕영대로 173-2(서천동) 전자정보대학 351호

구엔 안 투

경기도 용인시 기흥구 덕영대로 1732, 경희대학교 국제캠퍼스 전자정보대학 349호

이 발명을 지원한 국가연구개발사업

과제고유번호	2015-0-00742
부처명	미래창조과학부
연구관리전문기관	정보통신기술진흥센터
연구사업명	Grand ICT 연구센터 지원사업
연구과제명	라이프 컴패니온쉽 경험을 위한 지능형 인터랙션 융합연구
기 여 율	1/1
주관기관	경희대학교 산학협력단
연구기간	2017.01.01 ~ 2017.12.31

명 세 서

청구범위

청구항 1

데이터베이스에 저장된 하나 이상의 데이터베이스 이미지에서 쿼리이미지 또는 검색 키워드에 대응되는 이미지를 검색하는 방법에 있어서,

하나 이상의 태그가 맵핑된 상기 데이터베이스 이미지로부터 하나 이상의 시각 단어 및 하나 이상의 텍스트 단어를 획득하는 전처리 단계;

상기 시각 단어와 상기 텍스트 단어를 이용하여 주제별 시각 단어 분포($\Phi_{v,t}$), 주제별 텍스트 단어 분포($\Phi_{w,t}$) 및 배경의 시각 단어 분포(Ω)를 포함하는 배경분포 토픽모델의 모델 파라미터를 추정하는 단계;

상기 추정된 모델 파라미터를 포함하는 상기 배경분포 토픽모델을 이용하여 상기 데이터베이스 이미지의 주제 분포(θ_d)를 모델링하는 단계;

상기 주제별 텍스트 단어 분포($\Phi_{w,t}$)와 이미지 내 주제 비율(θ_{td})을 이용하여 상기 데이터베이스 이미지에 대한 상기 태그의 상응도를 산출하고, 그에 따라 상기 태그를 정제하는 단계;

상기 쿼리 이미지를 포함하는 검색 요청이 입력되면 상기 쿼리 이미지로부터 하나 이상의 시각 단어를 획득하고, 상기 추정된 모델 파라미터를 포함하는 상기 배경분포 토픽모델을 이용하여 상기 쿼리 이미지의 주제 분포(θ_q)를 모델링하는 단계;

상기 데이터베이스 이미지의 주제 분포(θ_d)와 상기 데이터베이스에서 상기 쿼리 이미지의 주제 분포(θ_q)를 이용하여 상기 데이터베이스 이미지의 유사도를 평가하는 단계를 포함하는 이미지 검색 방법.

청구항 2

제1항에 있어서,

상기 전처리 단계는

상기 데이터베이스 이미지에서 이미지 특징을 추출하는 단계;

상기 추출한 이미지 특징을 클러스터링하여 클러스터의 대표 값을 추출하는 단계;

상기 대표 값을 양자화하여 시각 단어를 획득하는 단계를 포함하는 이미지 검색 방법.

청구항 3

제1항에 있어서,

상기 모델 파라미터 추정 단계는

주제 분포에 할당된 시각 단어의 비율, 특정 주제(t)에서 특정 시각 단어(v_{di})가 나올 확률, 특정 이미지(I_d)에서 특정 주제(t)가 나올 확률을 이용하여 주제 분포에 할당된 시각 단어의 비율을 도출하는 a 단계;

배경 분포에 할당된 시각 단어의 비율, 배경 단어의 확률을 이용하여 배경 분포에 할당된 시각 단어의 비율을 도출하는 b 단계;

특정 주제(t)에서 텍스트 단어(w_{di})가 나올 확률, 특정 이미지(I_d)에서 특정 주제(t)가 나올 확률을 이용하여 상기 텍스트 단어(w_{di})가 상기 특정 주제(t)에 할당될 확률을 도출하는 c 단계;

상기 하나 이상의 데이터베이스 이미지에 대하여 상기 a 단계 내지 상기 c 단계를 반복적으로 수행하여, 정규화된 디리클레 파라미터인 상기 주제별 시각 단어 분포($\Phi_{v,t}$), 상기 주제별 텍스트 단어 분포($\Phi_{w,t}$) 및 상기 배경의 시각 단어 분포(Ω)를 획득하는 모델 파라미터 획득 단계를 포함하는 이미지 검색 방법.

청구항 4

제3항에 있어서,

상기 모델 파라미터 획득 단계는

상기 주제별 시각 단어 분포($\Phi_{v,t}$), 상기 주제별 텍스트 단어 분포($\Phi_{w,t}$) 및 상기 배경의 시각 단어 분포(Ω)는 하기 수식에 따라 추정되며,

상기 수식에서 β_b, β_v, β_w는 기 설정된 디리클레 파라미터, n_{vt}^{VT}는 시각 단어 v가 주제 t에 할당된 횟수, n_{wt}^{WT}는 텍스트 단어 w가 주제 t에 할당된 횟수, n_v^V는 시각 단어 v가 배경 분포에 할당된 횟수인 이미지 검색 방법.

$$\Phi_{v,t} = \frac{n_{vt}^{VT} + \beta_v}{\sum_{v'} n_{v't}^{VT} + V\beta_v}$$

$$\Phi_{w,t} = \frac{n_{wt}^{WT} + \beta_w}{\sum_{w'} n_{w't}^{WT} + W\beta_w}$$

$$\Omega = \frac{n_v^V + \beta_b}{\sum_{v'} n_{v'}^V + V\beta_b}$$

청구항 5

제1항에 있어서,

상기 유사도 평가 단계는

상기 데이터베이스 이미지(I_d) 각각에 대하여 쿼리 이미지(I_q)와의 유사도($S(I_d, I_q)$)를 하기 수식에 따라 계산하는 단계를 포함하고,

상기 수식에서 θ_d 및 θ_q는 상기 데이터베이스 이미지(I_d) 및 상기 쿼리 이미지(I_q)의 시각 정보에 각각 대응되며, r_d 및 r_q는 상기 데이터베이스 이미지(I_d) 및 상기 쿼리 이미지(I_q)의 텍스트 정보에 각각 대응되고, $\mu \in [0,1]$는 미리 설정된 텍스트 유사성의 가중치를 제어하는 파라미터인 것을 특징으로 하는 이미지 검색 방

법.

$$S(I_d, I_q) = (1-\mu)\theta_d\theta_q + \mu r_d r_q$$

청구항 6

제1항에 있어서,

상기 유사도 평가 단계는

상기 검색 키워드를 포함하는 검색 요청이 입력되면 상기 정제된 태그와 상기 검색 키워드의 유사도를 평가하는 단계를 더 포함하는 이미지 검색 방법.

청구항 7

제1항에 있어서,

상기 유사도 평가 결과에 따라, 상기 쿼리 이미지 또는 상기 검색 키워드 중 적어도 하나와 유사한 데이터베이스 이미지를 유사도가 높은 순서대로 정렬하는 단계를 더 포함하는 이미지 검색 방법.

청구항 8

하나 이상의 데이터베이스 이미지를 저장하는 데이터베이스;

하나 이상의 태그가 맵핑된 상기 데이터베이스 이미지로부터 하나 이상의 시각 단어 및 하나 이상의 텍스트 단어를 획득하는 제 1 전처리부;

상기 시각 단어와 상기 텍스트 단어를 이용하여 주제별 시각 단어 분포($\Phi_{v,t}$), 주제별 텍스트 단어 분포($\Phi_{w,t}$) 및 배경의 시각 단어 분포(Ω)를 포함하는 배경분포 토픽모델의 모델 파라미터를 추정하고, 상기 추정된 모델 파라미터를 포함하는 상기 배경분포 토픽모델을 이용하여 상기 데이터베이스 이미지의 주제 분포(θ_d)를 모델링하고, 상기 주제별 텍스트 단어 분포($\Phi_{w,t}$)와 이미지 내 특정 주제 비율(θ_{td})을 이용하여 상기 데이터베이스 이미지에 대한 상기 태그의 상응도를 산출하고, 그에 따라 상기 태그를 정제하는 제 1 배경분포 토픽모델링부;

쿼리 이미지 또 는검색 키워드가 입력되면 상기 쿼리 이미지로부터 하나 이상의 시각 단어를 획득하고, 상기 검색 키워드로부터 하나 이상의 텍스트 단어를 획득하는 제 2 전처리부;

상기 추정된 모델 파라미터를 포함하는 상기 배경분포 토픽모델을 이용하여 쿼리 이미지의 주제 분포(θ_q)를 모델링하는 제 2 배경분포 토픽모델링부;

상기 데이터베이스 이미지의 주제 분포(θ_d)와 상기 데이터베이스에서 상기 쿼리 이미지의 주제 분포(θ_q)를 이용하여 상기 데이터베이스 이미지의 유사도를 평가하는 유사도 평가부를 포함하는 이미지 검색 시스템.

청구항 9

제8항에 있어서,

상기 제 1 전처리부는

상기 데이터베이스 이미지에서 이미지 특징을 추출하고, 상기 추출한 이미지 특징을 클러스터링하여 클러스터의

대표 값을 추출하고, 상기 대표 값을 양자화하여 시각 단어를 획득하는 이미지 검색 시스템.

청구항 10

제8항에 있어서,

상기 제 1 배경분포 토픽모델링부는

주제 분포에 할당된 시각 단어의 비율, 특정 주제(t)에서 특정 시각 단어(v_{di})가 나올 확률, 특정 이미지(I_d)에서 특정 주제(t)가 나올 확률을 이용하여 주제 분포에 할당된 시각 단어의 비율을 도출하고, 배경 분포에 할당된 시각 단어의 비율, 배경 단어의 확률을 이용하여 배경 분포에 할당된 시각 단어의 비율을 도출하고, 특정 주제(t)에서 텍스트 단어(w_{di})가 나올 확률, 특정 이미지(I_d)에서 특정 주제(t)가 나올 확률을 이용하여 상기 텍스트 단어(w_{di})가 상기 특정 주제(t)에 할당될 확률을 도출하는 계산부;

상기 하나 이상의 데이터베이스 이미지에 대하여 상기 계산부에서의 계산을 반복적으로 수행하여, 정규화된 디리클레 파라미터인 상기 주제별 시각 단어 분포($\Phi_{v,t}$), 상기 주제별 텍스트 단어 분포($\Phi_{w,t}$) 및 상기 배경의 시각 단어 분포(Ω)를 획득하는 모델 파라미터 추정부를 포함하는 이미지 검색 시스템.

청구항 11

제8항에 있어서,

상기 모델 파라미터 추정부는

상기 주제별 시각 단어 분포($\Phi_{v,t}$), 상기 주제별 텍스트 단어 분포($\Phi_{w,t}$) 및 상기 배경의 시각 단어 분포(Ω)를 하기 수식에 따라 추정하며,

상기 수식에서 β_b, β_v, β_w는 기 설정된 디리클레 파라미터, n_{vt}^{VT} 는 시각 단어 v가 주제 t에 할당된 횟수, n_{wt}^{WT} 는 텍스트 단어 w가 주제 t에 할당된 횟수, n_v^V 는 시각 단어 v가 배경 분포에 할당된 횟수인 이미지 검색 시스템.

$$\Phi_{v,t} = \frac{n_{vt}^{VT} + \beta_v}{\sum_{v'} n_{v't}^{VT} + V\beta_v}$$

$$\Phi_{w,t} = \frac{n_{wt}^{WT} + \beta_w}{\sum_{w'} n_{w't}^{WT} + W\beta_w}$$

$$\Omega = \frac{n_v^V + \beta_b}{\sum_{v'} n_{v'}^V + V\beta_b}$$

청구항 12

제8항에 있어서,

상기 유사도 평가부는

상기 데이터베이스 이미지(I_d) 각각에 대하여 쿼리 이미지(I_q)와의 유사도($S(I_d,I_q)$)를 하기 수식에 따라 계산하고,

상기 수식에서 θ_d 및 θ_q는 상기 데이터베이스 이미지(I_d) 및 상기 쿼리 이미지(I_q)의 시각 정보에 각각 대응되며, r_d 및 r_q는 상기 데이터베이스 이미지(I_d) 및 상기 쿼리 이미지(I_q)의 텍스트 정보에 각각 대응되고, $\mu \in [0,1]$는 미리 설정된 텍스트 유사성의 가중치를 제어하는 파라미터인 이미지 검색 시스템.

$$S(I_d,I_q) = (1-\mu)\theta_d\theta_q + \mu r_d r_q$$

청구항 13

제8항에 있어서,

상기 유사도 평가부는

상기 검색 키워드를 포함하는 검색 요청이 입력되면

상기 정제된 태그와 상기 검색 키워드의 유사도를 평가하는 이미지 검색 시스템.

청구항 14

제8항에 있어서,

상기 유사도 평가부에서의 유사도 평가 결과에 따라, 상기 쿼리 이미지 또는 상기 검색 키워드 중 적어도 하나와 유사한 데이터베이스 이미지를 유사도가 높은 순서대로 정렬하는 리스트 제공부를 더 포함하는 이미지 검색 시스템.

청구항 15

제1항 내지 제7항의 방법 중 어느 하나의 방법을 실행시키기 위하여 컴퓨터 판독 가능 매체에 저장된 이미지 검색 응용 프로그램.

발명의 설명

기 술 분 야

[0001] 본 발명은 시맨틱 이미지 검색 방법 및 시스템에 관한 것으로, 보다 자세하게는 배경분포 토픽 모델링을 이용한 시맨틱 이미지 검색 방법 및 시스템에 관한 것이다.

배 경 기 술

[0003] 최근 인터넷의 발달로 소셜 미디어 네트워크 서비스가 급격히 성장하고 있다. 그 결과, 멀티미디어의 양이 폭발적으로 증가하면서 효과적인 이미지 검색 시스템이 요구되며, 이미지 어노테이션은 폭발적으로 증가하는 웹이미지에 따른 효율적인 이미지 검색의 필요성으로 인해 그 중요도가 점점 높아지고 있다.

[0004] 대부분의 이미지 검색 연구는 주로 이미지의 내용을 분석하는 내용 기반 이미지 검색(CBIR: Content-based

Image Retrieval) 방법이 많이 진행되어 왔다. 내용 기반 이미지 검색은 색상, 텍스처 및 형태와 같은 시각적 특징을 이용하여 이미지의 내용을 분석한다. 이러한 방법은 정의하는 태그의 개수가 적을 경우에는 잘 작동하지만, 데이터셋이 커지고 태그의 종류가 다양해짐에 따라 성능이 떨어지게 된다.

[0005] 텍스트 기반 이미지 검색(TBIR: Text-based Image Retrieval)은 텍스트를 쿼리로 하여 텍스트에 대응되는 이미지를 검색하는 방식이다. 이 방식은 이미지의 시각적 내용이 수동으로 태깅된 텍스트 디스크립터에 의해 표현되며, 데이터베이스 관리 시스템에서 이미지 검색을 수행하는데 사용된다.

[0006] 내용 기반 이미지 검색은 대규모의 데이터베이스를 다루는데 효율적이나, 저차원 이미지 특성과 고차원 이미지 특성 간의 시맨틱 갭 문제를 갖는다. 텍스트 기반 이미지 검색은 높은 수준의 개념을 갖는 쿼리를 지원할 수 있지만, 대규모 데이터베이스에서 개별 이미지를 수동으로 일일이 태깅해야 하므로 시간이 과도하게 소모된다. 뿐만 아니라, 소셜 네트워크 서비스 이미지의 경우에는 이미지와 관련성이 낮은 태그가 태깅되는 경우도 빈번하다.

[0007] 따라서 폭발적으로 증가하는 웹 이미지를 효과적으로 검색하기 위해서는 상술한 시맨틱 갭을 줄이고 이미지 태깅 성능을 향상시킬 수 있는 시맨틱 검색 시스템이 필요하다.

발명의 내용

해결하려는 과제

[0009] 본 발명은 전술한 문제점을 해결하기 위한 것으로서, 시각 단어(visual word), 텍스트 단어(text word), 배경(background)의 상관관계를 공식화하고, 이를 이미지 검색에 활용하여 이미지 검색 정확도를 높일 수 있는 이미지 검색 방법 및 시스템을 제공하는 것을 일 목적으로 한다.

[0010] 또한, 본 발명은 태그의 누락을 예측하고 노이즈를 제거할 수 있는 이미지 검색 방법 및 시스템을 제공하는 것을 다른 목적으로 한다.

[0011] 또한, 본 발명은 효율적이고 효과적으로 쿼리 및 데이터베이스 이미지 사이의 유사성을 계산할 수 있는 이미지 검색 방법 및 시스템을 제공하는 것을 다른 목적으로 한다.

과제의 해결 수단

[0013] 이러한 목적을 달성하기 위한 본 발명은 데이터베이스에 저장된 하나 이상의 데이터베이스 이미지에서 쿼리이미지 또는 검색 키워드에 대응되는 이미지를 검색하는 방법에 있어서, 하나 이상의 태그가 맵핑된 상기 데이터베이스 이미지로부터 하나 이상의 시각 단어 및 하나 이상의 텍스트 단어를 획득하는 전처리 단계, 상기 시각 단어와 상기 텍스트 단어를 이용하여 주제별 시각 단어 분포($\Phi_{v,t}$), 주제별 텍스트 단어 분포($\Phi_{w,t}$) 및 배경의 시각 단어 분포(Ω)를 포함하는 배경분포 토픽모델의 모델 파라미터를 추정하는 단계, 상기 추정된 모델 파라미터를 포함하는 상기 배경분포 토픽모델을 이용하여 상기 데이터베이스 이미지의 주제 분포(θ_d)를 모델링하는 단계, 상기 주제별 텍스트 단어 분포($\Phi_{w,t}$)와 이미지 내 주제 비율(θ_{td})을 이용하여 상기 데이터베이스 이미지에 대한 상기 태그의 상응도를 산출하고, 그에 따라 상기 태그를 정제하는 단계, 상기 쿼리 이미지를 포함하는 검색 요청이 입력되면 상기 쿼리 이미지로부터 하나 이상의 시각 단어를 획득하고, 상기 추정된 모델 파라미터를 포함하는 상기 배경분포 토픽모델을 이용하여 상기 쿼리 이미지의 주제 분포(θ_q)를 모델링하는 단계, 상기 데이터베이스 이미지의 주제 분포(θ_d)와 상기 데이터베이스에서 상기 쿼리 이미지의 주제 분포(θ_q)를 이용하여 상기 데이터베이스 이미지의 유사도를 평가하는 단계를 포함하는 것을 일 특징으로 한다.

[0014] 또한, 본 발명은 이미지 검색 시스템에 있어서, 하나 이상의 데이터베이스 이미지를 저장하는 데이터베이스, 하나 이상의 태그가 맵핑된 상기 데이터베이스 이미지로부터 하나 이상의 시각 단어 및 하나 이상의 텍스트 단어를 획득하는 제 1 전처리부, 상기 시각 단어와 상기 텍스트 단어를 이용하여 주제별 시각 단어 분포($\Phi_{v,t}$), 주

제별 텍스트 단어 분포($\Phi_{w,t}$) 및 배경의 시각 단어 분포(Ω)를 포함하는 배경분포 토픽모델의 모델 파라미터를 추정하고, 상기 추정된 모델 파라미터를 포함하는 상기 배경분포 토픽모델을 이용하여 상기 데이터베이스 이미지의 주제 분포(θ_d)를 모델링하고, 상기 주제별 텍스트 단어 분포($\Phi_{w,t}$)와 이미지 내 특정 주제 비율(θ_{td})을 이용하여 상기 데이터베이스 이미지에 대한 상기 태그의 상응도를 산출하고, 그에 따라 상기 태그를 정제하는 제 1 배경분포 토픽모델링부, 상기 쿼리 이미지 또는 상기 검색 키워드가 입력되면 상기 쿼리 이미지로부터 하나 이상의 시각 단어를 획득하고, 상기 검색 키워드로부터 하나 이상의 텍스트 단어를 획득하는 제 2 전처리부, 상기 추정된 모델 파라미터를 포함하는 상기 배경분포 토픽모델을 이용하여 쿼리 이미지의 주제 분포(θ_q)를 모델링하는 제 2 배경분포 토픽모델링부, 상기 데이터베이스 이미지의 주제 분포(θ_d)와 상기 데이터베이스에서 상기 쿼리 이미지의 주제 분포(θ_q)를 이용하여 상기 데이터베이스 이미지의 유사도를 평가하는 유사도 평가부를 포함하는 것을 일 특징으로 한다.

발명의 효과

[0016] 전술한 바와 같은 본 발명에 의하면, 시각 단어(visual word), 텍스트 단어(text word), 배경(background)의 상관관계를 공식화하고, 이를 이미지 검색에 활용하여 이미지 검색 정확도를 높일 수 있다.

[0017] 또한, 본 발명에 의하면 태그의 누락을 예측하고 노이즈를 제거할 수 있다.

[0018] 또한, 본 발명에 의하면 효율적이고 효과적으로 쿼리 및 데이터베이스 이미지 사이의 유사성을 계산할 수 있다.

도면의 간단한 설명

[0020] 도 1은 본 발명의 일 실시 예에 따른 이미지 검색 시스템의 구성을 설명하기 위한 도면,

도 2는 본 발명의 일 실시 예에 따른 배경분포 토픽 모델을 설명하기 위한 도면,

도 3은 본 발명의 일 실시 예에 따른 이미지 검색 방법을 설명하기 위한 순서도,

도 4는 본 발명의 일 실시 예에 따른 전처리 단계를 보다 자세하게 설명하기 위한 순서도,

도 5는 본 발명의 일 실시 예에 따른 모델 파라미터 추정 단계를 보다 자세하게 설명하기 위한 순서도이다.

발명을 실시하기 위한 구체적인 내용

[0021] 전술한 목적, 특징 및 장점은 첨부된 도면을 참조하여 상세하게 후술되며, 이에 따라 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다. 본 발명을 설명함에 있어서 본 발명과 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 상세한 설명을 생략한다. 이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시 예를 상세히 설명하기로 한다. 도면에서 동일한 참조부호는 동일 또는 유사한 구성요소를 가리키는 것으로 사용되며, 명세서 및 특허청구의 범위에 기재된 모든 조합은 임의의 방식으로 조합될 수 있다. 그리고 다른 식으로 규정하지 않는 한, 단수에 대한 언급은 하나 이상을 포함할 수 있고, 단수 표현에 대한 언급은 또한 복수 표현을 포함할 수 있음이 이해되어야 한다.

[0022] 수동 이미지 태깅 문제를 해결하기 위해, 태그를 자동으로 추론하고 수정할 수 있는 자동 어노테이션 시스템이 제안된 바 있다. 최근의 가장 주목할만한 연구들에서 이미지 어노테이션은 일종의 분류 작업으로 간주되었으며, 분류기는 시각적 특징을 태그에 맵핑하는 훈련을 수행하였다. 그러나 이러한 접근법의 가장 큰 단점은 어노테이션이 잘 분류된 훈련 데이터에 의한 적은 수의 태그 어휘들로 제한된다는 점이다. 인터넷 상의 이미지에 달린 태그들은 실제로는 매우 높은 수준으로 분화된 잡음을 포함하기 때문에, 상술한 제한 사항은 실제 애플리케이션에서의 사용을 어렵게 만든다.

[0023] 또한 분류 기반 접근 방식은 잘 동작하기는 하나, 실행이 쉽지 않다. 분류 클래스의 개수는 단어의 수 만큼 많아지며, 정확도를 높이기 위해서는 이는 많은 수의 트레이닝 이미지 셋을 필요로 하기 때문이다. 또한 시각적

특성과 언어적 개념 간 상관관계가 그리 단순하지만은 않다는 것도 실행을 어렵게 만드는 문제 중 하나이다.

[0024] 시맨틱 갭 문제를 돌파하고, 이미지 관련 키워드를 예측하기 위한 방법으로 Correspondece Latent Dirichlet Allocation(이하, "CorrLDA"라 함)라고 불리는 생성 확률 모델(Generative Probabilistic Model)이 있다. 생성 확률 모델은 어떤 확률 분포와 그 파라미터가 있다고 할 때, 그로부터 랜덤 프로세스에 따라 데이터를 생성하는 관점의 모델이다. CorrLDA는 이미지 영역과 텍스트(워드) 집합의 잠재적 변수 표현 간 관계를 찾는 모델로, 이미지의 관련 키워드를 예측하기 위하여 이미지 내용과 이에 대응되는 텍스트의 동시 발생 패턴으로부터 잠재적인 시맨틱 주제를 발견할 수 있다. CorrLDA 모델은 이미지 특징과 텍스트 단어로부터 잠재적인 주제를 학습하는 방법을 제공한다. 이 모델은 문서-주제 비율을 통해 시각적 특징과 텍스트 단어 사이의 상응 정도를 이용하여 시각적 주제와 텍스트적 주제 사이의 직접적 관계를 도출한다.

[0025] CorrLDA 모델은 시각적, 텍스트적 주제 간 상관관계가 잠재적 주제를 통해 명시적으로 활용되며, 그 결과 효과적으로 이미지의 의미적 내용을 추출할 수 있다는 점에서 장점을 가지며, 다중 라벨링 문제를 직접 해결할 수 있다는 점에서 데이터 셋이 동적으로 업데이트되는 실제 애플리케이션에서 효과적으로 활용될 수 있다.

[0026] 본 발명은 CorrLDA가 검색 작업에 충분히 활용되지 못하고 있는 문제를 해결하기 위해 고안된 것으로, 배경 단어에 관한 새로운 개념을 이용하여 CorrLDA 모델을 확장하는 것을 그 특징으로 한다.

[0027] 상술한 CorrLDA 모델의 원형이 되는 잠재 디리클레 할당(Latent Dirichlet Allocation, LDA)는 주어진 문서에 대하여 각 문서에 어떤 주제들이 존재하는지에 대한 확률 모형이다. 미리 알고 있는 주제별 단어수 분포를 바탕으로 주어진 문서에서 발견된 단어수 분포를 분석함으로서 해당 문서가 어떤 주제들을 함께 다루고 있을지를 예측할 수 있다.

[0028] 이와 마찬가지로, 본 발명의 일 실시 예에서 사용되는 데이터베이스 이미지 및 쿼리 이미지는 여러 개의 주제를 가지며, 주제들은 디리클레 분포를 따른다고 가정할 수 있다. 즉, 잠재 디리클레 할당이 문서를 단어의 집합으로 보는 것처럼, 이미지는 하나 이상의 시각 단어(visual word)의 집합으로 볼 수 있는 것이다.

[0029] 다시 말해서, 주제는 이미지에 포함된 시각 단어의 확률분포이며, 의미적으로 유사하거나 연관성 높은 단어들로 이루어진다. 본 발명에서 각 이미지는 주제의 확률적 혼합체로 표현되며, 각 주제는 시각 단어 및 텍스트 단어(태그)의 분포로 표현된다.

[0030] 본 발명의 일 실시 예에서 이미지 특징(feature)는 시각 단어(visaul word)의 형태로 정의될 수 있으며, 이는 데이터베이스 이미지 또는 쿼리 이미지에서 특징 정보를 추출하여, 이산화하는 과정을 통해 획득 가능하다. 이렇게 생성된 본 발명의 일 실시 예에 따른 이미지 코퍼스는 시각 단어집(Bag of visual word) 및 텍스트 단어집(Bag of text word)으로 표현할 수 있다. 시각 단어집은 이미지로부터 SIFT 등으로 이미지 특징 정보를 추출하고, 추출된 이미지 특징 정보들에 대해 클러스터링을 수행하여 시각 단어를 찾아낸 후, 각각의 이미지들을 시각 단어들의 히스토그램으로 표현한 것으로 이해될 수 있으며, 텍스트 단어집은 각 이미지에 수동으로 달린 하나이상의 태그들의 집합으로 이해될 수 있다.

[0031] 본 명세서에서 '하나 이상의 시각 단어'는 시각 단어집을 의미하며, '하나 이상의 텍스트 단어'는 텍스트 단어집을 의미하는 것으로 이해될 수 있다. 또한 '텍스트 단어'와 '태그'는 지칭하는 대상은 동일하나, '태그'는 웹, 소셜 네트워크 서비스 등 수집처에서 사용자들이 이미지에 수동으로 맵핑시킨 단어를 의미하며, '텍스트 단어'는 '시각 단어'와 대비되는 보다 포괄적인 개념으로 이해될 수 있다.

[0032] 한편, 본 발명의 일 실시 예에 의한 이미지 검색 방법은 이미지의 배경에 일반적으로 포함되어 있는 의미가 좁거나 없는 시각 단어인 '배경 단어'를 제거하는 개념을 사용한다. 배경 단어는 거의 모든 이미지에 포함되어 있으나 유용하지 않고, 이미지 검색 시 유사도를 측정함에 있어서 노이즈가 될 수 있으므로, 배경 단어를 고려한 토픽 모델링을 통해 이미지 검색의 정확도를 향상시킬 수 있다.

[0033] 나아가, 본 발명에서는 주제를 추출하고 태그를 정제함에 있어서, 쿼리와 데이터베이스에 저장된 이미지 사이의 유사성을 평가하는 점수 체계를 제안한다.

[0034] 본 발명의 일 실시 예에 따르면, 시각 단어는 주제 단어와 배경 단어로 분류될 수 있다. 따라서 본 발명에서 각각의 이미지는 시각 주제 분포, 배경 분포, 텍스트 주제 분포로 나타낼 수 있다.

[0035] 이하에서는, 도면을 참조하여 본 발명의 일 실시 예에 따른 토픽 모델링 기반 시맨틱 이미지 검색 방법을 설명한다.

[0036] 도 1은 본 발명의 일 실시 예에 따른 이미지 검색 시스템을 도시한 도면이다.

[0037] 도 1을 참조하면, 본 발명의 일 실시 예에 따른 이미지 검색 시스템은 검색 데이터 생성 모듈(100) 및 검색 모듈(200)을 포함할 수 있으며, 검색 데이터 생성 모듈(100)은 데이터베이스(110), 제 1 전처리부(130), 제 1 배경분포 토픽모델링부(150)를 포함할 수 있고, 검색 모듈(200)은 제 2 전처리부(230), 제 2 배경분포 토픽모델링부(250), 유사도 평가부(270)를 포함할 수 있다.

[0038] 실시 예에 따라, 검색 데이터 생성 모듈(100)은 오프라인에서 실행될 수 있으며, 검색 모듈(200)은 온라인에서 실행될 수 있다.

[0039] 검색 데이터 생성 모듈(100)의 데이터베이스(110)는 하나 이상의 데이터베이스 이미지를 저장한다. 데이터베이스 이미지는 웹에서 수집된 이미지 또는 사용자에 의하여 저장된 이미지로, 하나 이상의 태그가 맵핑된 이미지일 수 있다.

[0040] 최근 널리 사용되고 있는 소셜 네트워크 서비스의 경우 해시태그(#태그)를 이미지에 기재하여 이미지의 특징을 설명할 수 있도록 하고, 태그를 이용하여 태그와 관련된 이미지 검색이 용이하게 이루어질 수 있도록 한다. 이러한 웹 수집 이미지의 특징에 따라서, 데이터베이스에는 수집된 이미지가 해당 이미지에 맵핑된 태그와 함께 저장될 수 있다.

[0041] 데이터베이스에 저장된 하나 이상의 데이터베이스 이미지는 제 1 전처리부(130)로 전달될 수 있다. 제 1 전처리부(130)는 데이터베이스 이미지를 전처리하여 데이터베이스 이미지로부터 하나 이상의 시각 단어 및 하나 이상의 텍스트 단어를 획득한다. 예를 들어, 제 1 전처리부(130)는 아핀 불변 검출기(Affine invariant detector)를 이용하여 데이터베이스 이미지에서 이미지 특징(image feature)를 추출하고, 이미지 특징은 128-DSIFT(Scale-InvariantFeature Transform) 디스크립터(discriptor)를 사용하여 표현될 수 있다. 제 1 전처리부(130)는 추출한 이미지 특징을 클러스터링하여 클러스터의 대표 값을 추출하고, 대표 값을 양자화하여 시각 단어를 획득할 수 있다. 일 예로, 제 1 전처리부(130)는 k-means 클러스터링에 의해 학습된 코드북을 사용하여 디스크립터를 시각 단어로 양자화할 수 있다. 나아가 제 1 전처리부(130)는 데이터베이스 이미지에 태그가 맵핑되어 있지 않은 경우, 시각 단어를 이용하여 상기 데이터베이스 이미지에 대한 태그(텍스트 단어)를 생성할 수 있다.

[0042] 제 1 배경분포 토픽모델링부(150)는 제 1 전처리부(130)에서 획득한 하나 이상의 시각 단어와 태그로부터 획득한 하나 이상의 텍스트 단어를 본 발명의 일 실시 예에 의한 배경분포 토픽모델의 모델 파라미터를 학습 및 추정하는데 사용할 수 있다.

[0043] 보다 구체적으로, 제 1 배경분포 토픽모델링부(150)는 계산부(153, 미도시) 및 모델 파라미터 추정부(155, 미도시)를 포함할 수 있다.

[0044] 제 1 배경분포 토픽모델링부(150)는 상기 시각 단어와 상기 텍스트 단어를 이용하여 주제별 시각 단어 분포($\Phi_{v,t}$), 주제별 텍스트 단어 분포($\Phi_{w,t}$) 및 배경의 시각 단어 분포(Ω)를 포함하는 배경분포 토픽모델의 모델 파라미터를 추정할 수 있는데, 계산부(153)는 주제 분포에 할당된 시각 단어의 비율, 특정 주제(t)에서 특정 시각 단어(v_{di})가 나올 확률, 특정 이미지(I_d)에서 특정 주제(t)가 나올 확률을 이용하여 주제 분포에 할당된 시각 단어의 비율을 도출하고, 배경 분포에 할당된 시각 단어의 비율, 배경 단어의 확률을 이용하여 배경 분포에 할당된 시각 단어의 비율을 도출하고, 특정 주제(t)에서 텍스트 단어(w_{di})가 나올 확률, 특정 이미지(I_d)에서 특정 주제(t)가 나올 확률을 이용하여 상기 텍스트 단어(w_{di})가 상기 특정 주제(t)에 할당될 확률을 도출할 수 있다.

[0045] 모델 파라미터 추정부(155)는 모델 파라미터가 수렴될 때 까지 데이터베이스에 저장된 데이터베이스 이미지 셋에 대하여 계산부(153)에서의 동작을 반복적으로 수행함으로써, 정규화된 디리클레 파라미터인 상기 주제별 시각 단어 분포($\Phi_{v,t}$), 상기 주제별 텍스트 단어 분포($\Phi_{w,t}$) 및 상기 배경의 시각 단어 분포(Ω)를 획득할 수 있다.

[0046] 도 2를 참조하여 본 발명의 일 실시 예에 의한 배경분포 토픽모델을 설명한다. 도 2의 (a)는 종래의 상응 잠재 디리클레 할당(이하, CorrLDA 이라 함)의 확률 그래프를 도시한 것이고, (b)는 본 발명의 일 실시 예에 따른 배

경분포 토픽모델(CTMB)의 확률 그래프를 도시한 것이다.

[0047] 도 2(b)에 도시된 본 발명의 일 실시 예에 따른 CTMB의 파라미터에 관한 설명은 표 1과 같다.

표 1

[0049]

파라미터	설명
$\alpha,\beta_v,\beta_w,\beta_b,\gamma$	기 설정된 디리클레 하이퍼 파라미터
$\lambda_d \in R^2$	이미지 내 단어 종류(시각 단어/배경)의 베르누이 분포
$\Omega \in R^V$	배경의 시각 단어 멀티노미얼 분포
$\theta_d \in R^T$	이미지 d의 주제 멀티노미얼 분포
$\phi_{v,t} \in R^V$	주제 t의 시각 단어 멀티노미얼 분포
$\phi_{w,t} \in R^W$	주제 t의 텍스트 단어 멀티노미얼 분포

[0051] 도 2(b)에서, 사각형은 복제물을 나타낸다. D개의 데이터베이스 이미지와 개별 데이터베이스 이미지(I_d)는 N_v 개의 시각 단어와 N_w개의 텍스트 단어로 표현할 수 있다. 도 2(b)에서 음영처리된 노드는 관찰된 변수를 나타내고, 흰색 노드는 잠재 변수를 나타낸다. CTMB에서 각 데이터베이스 이미지는 주제 분포로 모델링되며, 각 주제는 크기 v의 시각 단어와 크기 w의 텍스트 단어에 대한 분포로 모델링 된다. 잠재 변수 z는 시각 주제를 나타내고, y는 텍스트 주제를 나타낸다. CTMB에서 이미지 특징, 즉, 이미지의 시각적 특징(visual feature)은 텍스트 단어 생성에 사용되며, CTMB의 확률 모델은 다음과 같은 단계를 거친다.

[0053] 1. 각 토픽 t에 대하여,

[0054] a. 시각 단어 분포: $\phi_{v,t} \sim Dirichlet(\beta_v)$

[0055] b. 텍스트 단어 분포: $\phi_{w,t} \sim Dirichlet(\beta_w)$

[0056] 2. 배경 분포: $\Omega \sim Dirichlet(\beta_b)$

[0057] 3. 데이터베이스 이미지 각각에 대하여(I_d, d=1,...,D)

[0058] a. 단어 타입 분포: $\lambda_d \sim$ Beta(γ)

[0059] b. 주제 분포: $\theta_d \sim Dirichlet(\alpha)$

[0060] 4. 각 시각 단어 분포에 있어서, θ_{di}, $i \in 1,2,...,N_v$

[0061] a. 스위치 샘플 $s_{di} \sim Bernoulli(\lambda_d)$

[0062] b. $s_{di} = 1$ 이면,

[0063] i. 시각 주제 변수 z_{di} ~ Multinomial(θ_d)

[0064] ii. 시각 주제 단어 v_{di} ~ Multinomial($\phi_{v,z_{di}}$)

[0065] c. $s_{di}=2$ 이면,

[0066] i. 시각 배경 단어 v_{di} ~ Multinomial(Ω)

[0067] 5. 각 텍스트 단어 w_{dj}, $j\in 1,2,...,N_w$

[0068] a. 텍스트 주제 변수 y_{dj} ~ Unif($z_1,\cdots,z_{N_v}$)

[0069] b. 텍스트 단어 w_{dj} ~ Multinomial($\phi_{w,y_{dj}}$)

[0071] 위 프로세스에서, Dirichlet, Multinomial은 각각 디리클레 분포, 멀티노미얼 분포를 의미한다. 멀티노미얼 분포는 단어 분포에 대해 디리클레 분포와 공역을 맺도록 선택되어 계산의 단순화와 효율적인 추론을 보장한다. 스위치 변수 s는 시각 단어의 생성을 제어한다. 이미지는 두 가지 유형의 시각 단어를 포함하는데, 하나는 주제 분포 Multinomial(ϕ_v)으로부터 생성되고, 다른 하나는 배경 분포 Multinomial(Ω)에서 생성되어 시각 단어와 배경 간의 직접적인 상관관계를 유도한다. 그렇지 않은 경우, 텍스트 주제 y는 균일한 분포를 통해 시각 주제 z 중 하나에 대응되며, 텍스트 단어는 주제 분포 Multinomial(ϕ_w)에서 생성된다. 따라서 본 발명의 일 실시 예에 따른 CTMB를 사용하면 시각 단어와 텍스트 단어의 상관 관계가 크게 강화된다.

[0072] CTMB 모델의 목표는 주어진 이미지 코퍼스 $\{V_d, W_d\}_{d=1}^{D}$로부터 하기 수학식 1을 최대화하여 모델 파라미터 (Φ_v,Φ_w,Ω)를 추정하는 것이다.

수학식 1

[0073] $$L(\Phi_v,\Phi_w,\Omega)=\sum_{d=1}^{D}\log(p(V_d,W_d,Z_d,Y_d,S_d|\Phi_v,\Phi_w,\Omega,\Pi)$$

[0074] 여기서, Π는 하이퍼 파라미터 {$\alpha,\beta_v,\beta_w,\beta_b,\gamma$}를 나타낸다. 파라미터 ϕ_v와 $\Omega\in R^{W\times T}$, ϕ_w는 각각 시각 주제, 배경, 텍스트 주제 분포를 나타낸다.

[0075] 위의 분포를 정확하게 추정하는 것은 매우 어려운 문제이므로, 근사 추정 알고리즘을 사용할 수 있다. CTMB의 파라미터 추정의 일 실시 예로, 아래와 같은 Monte Carlo EM 알고리즘이 사용될 수 있다. 이 알고리즘에서 잠재 변수 z, s, y의 샘플링에는 뭉개진 깁스 샘플링 알고리즘이 사용될 수 있으며, 이는 수학식 2 내지 4와 같이 계산될 수 있다.

[0076] 보다 구체적으로, Monte Carlo EM 알고리즘에 의하면, 하나 이상의 시각 단어(시각 단어집) 및 하나 이상의 텍스트 단어(텍스트 단어집)으로 형성된 이미지 코퍼스가 입력이 되고, 출력은 추정된 파라미터인 Φ_v,Φ_w,Ω가 된다. 상기 알고리즘에 따르면, 먼저 각 파라미터를 초기화({$\Phi_v^{(0)},\Phi_w^{(0)},\Omega^{(0)}$})하고, 주제의 개수를 K라고 할 때, k=1, ..., K에서, {$\Phi_v^{(k-1)},\Phi_w^{(k-1)},\Omega^{(k-1)}$}가 주어졌을 때, 하기 수학식 2 내지 수학식 4를 이용한 뭉개진 깁스 샘플링으로부터 각각의 이미지(I_d)에 대하여 N깁스 단계에 의한 샘플 잠재 변수를 획득할 수 있다.

다음으로, 수학식 6 내지 8을 이용하여 $\{\Phi_v^{(k)}, \Phi_w^{(k)}, \Omega^{(k)}\}$의 계산을 완료할 수 있다.

[0077]

수학식 2

$$p(z_{di}=t, s_{di}=1|\mathbf{v_d}, \mathbf{z_{-di}}, \mathbf{s_{-di}}, \Pi) \propto \frac{N_{v1,-di}+\gamma}{N_{v,-di}+2\gamma} \times \frac{n_{vt,-di}^{VT}+\beta_v}{\sum_{v'} n_{v't,-di}^{VT}+V\beta_v} \times \frac{n_{td,-di}^{TD}+\alpha}{\sum_{t'} n_{t'd,-di}^{TD}+T\alpha}$$

[0078]

수학식 3

$$p(s_{di}=2|\mathbf{v_d}, \mathbf{s_{-di}}, \Pi) \propto \frac{N_{v2,-di}+\gamma}{N_{v,-di}+2\gamma} \times \frac{n_{v,-di}^{V}+\beta_b}{\sum_{v'} n_{v',-di}^{V}+V\beta_b}$$

[0079]

수학식 4

$$p(y_{dj}=t|\mathbf{w_d}, \mathbf{z}, \mathbf{y_{-di}}, \Pi) \propto \frac{n_{wt,-dj}^{WT}+\beta_w}{\sum_{w'} n_{w't,-di}^{WT}+W\beta_w} \times \frac{n_{td}^{TD}}{N_v}$$

[0080]

[0082] 상기 수학식 2 내지 4를 살펴보면, N_{v1}과 N_{v2}는 각각 연관된 주제 단어와 배경 단어에 할당된 이미지 d의 시각 단어 수를 의미한다. n_{td}^{TD}는 주제 t에 할당된 이미지 d의 시각 단어 수이며, n_{vt}^{VT}, n_{wt}^{WT}는 시각 단어 v, 텍스트 단어 w가 각각 주제 t에 할당된 횟수, n_v^V는 단어 v가 이미지 코퍼스의 배경 단어 분포에 할당된 횟수이다. 그리고 $-d_i$는 현재 할당을 제외한 모든 할당을 나타낸다.

[0083] 위 방정식은 파라미터 $\Phi_v, \Phi_w, \Omega, \theta$를 별도로 주변화(marginalize)시킴으로써 얻어진다. 관찰되는 바와 같이, 수학식 2의 첫 번째 항은 주제 분포 $s_{di}=1$과 배경 분포 $s_{di}=2$dp 할당된 시각 단어의 비율을 나타낸다. 수학식 2의 두 번째 항은 특정 주제(t)에서 특정 시각 단어(v_{di})가 나올 확률을 나타내며, 수학식 3의 두 번째 항은 배경 단어의 확률을 나타낸다. 수학식 2의 마지막 항은 특정 이미지(I_d)에서 특정 주제(t)가 나올 확률을 나타낸다. 나아가 수학식 4는 관찰된 텍스트 단어 w_{di}가 주제(t)에 할당되는 확률을 측정하는데, 이 식의 마지막 항은 하나의 이미지에서 주제 t에 할당된 시각 단어의 비율을 통해 시각적 내용과의 상응도를 나타낸다.

[0084] 모든 잠재 변수가 샘플링 방정식으로부터 계산되므로, 파라미터 Φ_v, Φ_w, Ω는 사후 분포를 검토하여 추정된다. 몇 번의 반복되는 단계에 따라, 파라미터들은 $\Phi_v^*, \Phi_w^*, \Omega^*$로 수렴한다. 주제-시각 단어 다항식의 뒷부분은 다음 수학식 5 내지 8과 같이 계산된다.

수학식 5

[0085] $$p(\Phi_{v,t}|\mathrm{v},\mathrm{z},\mathrm{s}) = \mathrm{Dir}\left\{\beta_v + n_{vt}^{VT}\right\}$$

[0086] $V=\{V_d\}_{d=1}^{D}$, $Z=\{Z_d\}_{d=1}^{D}$, $S=\{S_d\}_{d=1}^{D}$이다. 따라서, Φ_v는 $p(\Phi_{v,t}|V,Z,S)$의 사후 평균으로 추정할 수 있다. 이는 다음과 같이 단순히 정규화된 디리클레 파라미터이다.

수학식 6

[0088] $$\Phi_{v,t} = \frac{n_{vt}^{VT} + \beta_v}{\sum_{v'} n_{v't}^{VT} + V\beta_v}$$

[0089] 수학식 6에서의 추정과 유사하게, 배경 분포 및 텍스트 분포에 해당하는 Ω와 Φ_w도 다음과 같이 추정할 수 있다.

수학식 7

[0090] $$\Omega = \frac{n_v^V + \beta_b}{\sum_{v'} n_{v'}^V + V\beta_b}$$

수학식 8

[0091] $$\Phi_{w,t} = \frac{n_{wt}^{WT} + \beta_w}{\sum_{w'} n_{w't}^{WT} + W\beta_w}$$

[0093] 모델 파라미터가 추정되면, 본 발명의 일 실시 예에 따른 배경분포 토픽모델링부(130, 230)는 쿼리이미지가 입력되거나, 새 이미지가 데이터베이스에 추가될 때, 이미지를 추론할 준비를 한다. 추론 단계에서, 수학식 2 및 수학식 3의 두 번째 항과 수학식 4의 첫 번째 항을 이미 추정된 파라미터 $\Phi_{vt},\Phi_{wt},\Omega$로 대체하여, z_{di},y_{di}와 같은 보이지 않는 이미지들의 잠재 파라미터를 유도할 수 있다.

[0094] 본 발명의 일 실시 예에 따른 이미지검색 프레임워크는 각 이미지를 독립적으로 고려한다. 따라서 처리할 여러 시스템 간에 이미지 분배가 가능하므로, 프레임워크의 확장이 가능하며, 데이터베이스에 있는 수십억개의 이미지가 실제 응용 프로그램에서 작동할 수 있다.

[0096] 제 1 배경분포 토픽모델링부(150)는 추정된 모델 파라미터를 포함하는 배경분포 토픽모델을 이용하여 데이터베

이스 이미지의 주제 분포(θ_d)를 모델링하고, 주제별 텍스트 단어 분포($\Phi_{w,t}$)와 이미지 내 특정 주제 비율(θ_{td})을 이용하여 데이터베이스 이미지에 대한 태그의 상응도를 산출하고, 그에 따라 태그를 정제할 수 있다. 이미지의 주제 분포 모델링은 전술한 바와 같으며, 태그의 정제는 후술하는 바와 같다.

[0098] 본 발명의 일 실시 예에 따른 배경분포 토픽모델(CTMB)에서, 특정 이미지 I_d에 대한 태그 w의 상응도(상응 확률)은 수학식 9에 따라 계산될 수 있다.

수학식 9

[0099]
$$P(w|\mathbf{v_d}) = \sum_t P(w|t)P(t|\mathbf{v_d}) = \sum_t \Phi_{w,wt}\theta_{td}$$

[0101] 따라서 태그를 예측하고, 누락된 태그를 추가하여 향상시키는 태그의 정제는 관련 태그가 계산된 확률에 기초하여 정렬하는 것을 통해 달성된다. 다시 말해서, 관련성 없는 태그의 확률은 작아야 하지만, 누락된 태그의 확률은 추출된 주제를 통해 증가된다. 여기서, Φ_w와 θ는 각각 태그-토픽 분포와 토픽-문서 비율을 나타낸다.

[0102] 검색 모듈(200)에 쿼리 이미지 또는 검색 키워드가 입력되면, 제 2 전처리부(230)는 쿼리 이미지로부터 하나 이상의 시각 단어를 획득하고, 검색 키워드로부터 하나 이상의 텍스트 단어를 획득할 수 있는데, 제 2 전처리부(230)의 동작은 데이터베이스 이미지 및 데이터베이스 이미지에 맵핑된 태그를 전처리하는 제 1 전처리부(130)의 동작과 동일하다.

[0103] 제 2 배경분포 토픽모델링부(250)는 추정된 모델 파라미터를 포함하는 배경분포 토픽모델을 이용하여 쿼리 이미지의 주제 분포(θ_q)를 모델링할 수 있다. 여기서, 제 2 배경분포 토픽모델링부(250)는 제 1 배경분포 토픽모델링부(150)에서 추정된 모델 파라미터를 이용하여 토픽 모델링을 수행한다. 제 2 배경분포 토픽모델링부(250)의 토픽 모델링은 전술한 제 1 배경분포 토픽모델링부(150)에서의 동작과 동일한 방법으로 이루어진다.

[0104] 제 1 배경분포 토픽모델링부(150)가 데이터베이스 이미지의 주제를 추출하고, 데이터베이스 이미지에 맵핑된 태그를 정제하고(170), 제 2 배경분포 토픽모델링부(250)가 쿼리 이미지의 주제를 추출하고 이를 이용하여 태그를 생성하면, 유사도 평가부(270)는 이들을 이용하여 데이터베이스 이미지와 쿼리 이미지의 유사도 또는 검색 키워드와 데이터베이스 이미지의 유사도를 평가한다.

[0105] 보다 구체적으로, 유사도 평가부(270)는 데이터베이스 이미지(I_d) 각각에 대하여 쿼리 이미지(I_q)와의 유사도($S(I_d,I_q)$)를 수학식 10에 따라 계산할 수 있다.

수학식 10

[0106]
$$S(I_d,I_q) = (1-\mu)\,\theta_d\,\theta_q + \mu\; r_d\, r_q$$

[0107] 위 수식에서 θ_d 및 θ_q는 데이터베이스 이미지(I_d) 및 쿼리 이미지(I_q)의 시각 정보에 각각 대응되며, r_d 및 r_q는 데이터베이스 이미지(I_d) 및 쿼리 이미지(I_q) 각각의 텍스트 정보를 나타내는 두 개의 W차원 벡터이다. $\mu \in [0,1]$는 미리 설정된 텍스트 유사성의 가중치를 제어하는 파라미터로, 이 값은 사용자의 기본 설정 또는

쿼리 유형을 기반으로 설정될 수 있다. 예를 들어, 쿼리가 검색 키워드이면, μ는 1로 설정될 수 있다. 이 경우, $r_q \in \{0,1\}^W$이며, 여기서 i 번째 태그가 쿼리에 나타나면 i번째 요소는 1로 설정될 수 있다.

[0108] 반면, 검색 요청이 이미지 또는 이미지와 검색키워드 모두를 포함하는 경우, 시각적 표현 θ와 정제된 태그 r의 텍스 표현은 CTMB 모델 적용 후에 추정될 수 있다. 이 작업에서 텍스트 정보 및 시각 정보는 동일한 중요도를 가지므로, 파라미터 μ는 0.5로 설정될 것이다. 검색 요청(쿼리)에 검색 키워드 없이 쿼리 이미지만 포함된 경우, CTMB의 텍스트 부분은 제외되고, 주제는 온전하게 이미지 특징에서만 추출될 수 있다.

[0109] 최종적으로, 리스트 제공부(280, 미도시)는 유사도 평가부(270)에서의 평가 결과에 따라 데이터베이스 이미지를 점수에 따라 내림차순으로 정렬하여 사용자에게 제공할 수 있다.

[0110]

[0111] 다음으로, 도 3을 참조하여, 본 발명의 일 실시 예에 따른 토픽 모델링을 이용한 시맨틱 이미지 검색 방법을 설명한다. 이하에서의 설명의 편의를 위해 본 발명의 일 실시 예에 따른 시맨틱 이미지 검색이 임의의 검색 시스템에서 이루어진다고 가정하고 설명한다. 실제로 본 발명의 일 실시 예에 따른 이미지 검색 방법은 다수의 서로 다른 주체(모듈)에 의해 수행될 수 있으며, 오프라인/온라인에서 각 단계가 별도로 이루어질 수 있음에 유의한다.

[0112] 도 3을 참조하면, 데이터베이스에 저장된 하나 이상의 데이터베이스 이미지에서 쿼리이미지 또는 검색 키워드에 대응되는 이미지를 검색하는 본 발명의 일 실시 예에 의한 이미지 검색 방법에 의하면, 검색 시스템은 하나 이상의 태그가 맵핑된 데이터베이스 이미지로부터 하나 이상의 시각 단어 및 하나 이상의 텍스트 단어를 획득하고(S100), 시각 단어와 텍스트 단어를 이용하여 주제별 시각 단어 분포($\Phi_{v,t}$), 주제별 텍스트 단어 분포($\Phi_{w,t}$) 및 배경의 시각 단어 분포(Ω)를 포함하는 배경분포 토픽모델의 모델 파라미터를 추정할 수 있다(S200). 다음으로, 검색 시스템 추정된 모델 파라미터를 포함하는 배경분포 토픽모델을 이용하여 데이터베이스 이미지의 주제 분포(θ_d)를 모델링할 수 있다(S300). 다음으로 검색 시스템은 주제별 텍스트 단어 분포($\Phi_{w,t}$)와 이미지 내 주제 비율(θ_{td})을 이용하여 데이터베이스 이미지에 대한 태그의 상응도를 산출하고, 그에 따라 태그를 정제할 수 있다(S400).

[0113] 한편 검색 시스템에 검색 요청이 수신되면, 검색 시스템은 검색 요청에 쿼리 이미지가 포함되어 있는지, 검색 키워드가 포함되어 있는지 여부를 판단할 수 있다(S500). 판단 결과, 쿼리 이미지를 포함하는 검색 요청이 입력되면 쿼리 이미지로부터 하나 이상의 시각 단어를 획득하고, 추정된 모델 파라미터를 포함하는 배경분포 토픽모델을 이용하여 쿼리 이미지의 주제 분포(θ_q)를 모델링하고(S630), 데이터베이스 이미지의 주제 분포(θ_d)와 데이터베이스에서 쿼리 이미지의 주제 분포(θ_q)를 이용하여 데이터베이스 이미지의 유사도를 평가할 수 있다(S700). 만약, 검색 요청에 검색 키워드가 포함되어 있으면, 검색 시스템은 단계 100 내지 단계 300을 통해 정제된 데이터베이스 이미지의 태그 중에서 검색 키워드에 대응되는 태그를 검색할 수 있다(S650). 그리고 단계 700에서 키워드-태그 간 유사도를 평가할 수 있다. 단계 630과 단계 650은 동시에 이루어질 수 있으며, 검색 요청에 쿼리 이미지 또는 검색 키워드 중 어느 하나만 포함되어 있다면 하나의 단계만 수행될 수 있다.

[0114] 단계 700의 유사도 평가 결과에 따라, 쿼리 이미지 또는 검색 키워드 중 적어도 하나와 유사한 데이터베이스 이미지의 유사도 점수가 산출되면, 검색 시스템은 유사도가 높은 순서대로 데이터베이스 이미지를 정렬하여 제공할 수 있다(S800)

[0115] 도 4를 참조하여 단계 100을 보다 자세하게 살펴보면, 단계 100에서 검색 시스템은 데이터베이스 이미지에서 이미지 특징을 추출하고(S130), 추출한 이미지 특징을 클러스터링하여 클러스터의 대표 값을 추출하고(S150), 대표 값을 양자화하여 시각 단어를 획득(S170)함으로써, 데이터베이스 이미지에서 시각 단어를 획득할 수 있다.

[0116] 도 5를 참조하여 단계 200을 보다 자세하게 살펴보면, 단계 210에서 검색 시스템은 주제 분포에 할당된 시각 단어의 비율, 특정 주제(t)에서 특정 시각 단어(v_{di})가 나올 확률, 특정 이미지(I_d)에서 특정 주제(t)가 나올 확

률을 이용하여 주제 분포에 할당된 시각 단어의 비율을 도출할 수 있다. 다음으로, 단계 230에서 검색 시스템은 배경 분포에 할당된 시각 단어의 비율, 배경 단어의 확률을 이용하여 배경 분포에 할당된 시각 단어의 비율을 도출한다. 그리고 단계 250에서 특정 주제(t)에서 텍스트 단어(w_{di})가 나올 확률, 특정 이미지(I_d)에서 특정 주제(t)가 나올 확률을 이용하여 텍스트 단어(w_{di})가 특정 주제(t)에 할당될 확률을 도출한다. 검색 시스템은 파라미터 값이 수렴 될 때까지 하나 이상의 데이터베이스 이미지에 대하여 a 단계 내지 c 단계를 반복적으로 수행하여, 정규화된 디리클레 파라미터인 주제별 시각 단어 분포($\Phi_{v,t}$), 주제별 텍스트 단어 분포($\Phi_{w,t}$) 및 배경의 시각 단어 분포(Ω)를 획득할 수 있다(S290). 그리고 이렇게 획득한 모델 파라미터는 이후 단계 630에서 쿼리 이미지의 주제 분포를 모델링할 때 사용된다.

[0117] 단계 290에서 주제별 시각 단어 분포($\Phi_{v,t}$), 상기 주제별 텍스트 단어 분포($\Phi_{w,t}$) 및 상기 배경의 시각 단어 분포(Ω)는 전술한 수학식 6내지 수학식 8에 따라 추정되며, 이 수식에서 β_b, β_v, β_w는 기 설정된 디리클레 파라미터, n_{vt}^{VT}는 시각 단어 v가 주제 t에 할당된 횟수, n_{wt}^{WT}는 텍스트 단어 w가 주제 t에 할당된 횟수, n_v^V는 시각 단어 v가 배경 분포에 할당된 횟수를 의미한다.

[0118] 한편, 단계 700의 유사도 평가 방법을 보다 자세히 살펴보면, 데이터베이스 이미지(I_d) 각각에 대하여 쿼리 이미지(I_q)와의 유사도($S(I_d,I_q)$)는 전술한 수학식 10에 따라 계산될 수 있으며, 수학식 10에서 θ_d 및 θ_q는 데이터베이스 이미지(I_d) 및 쿼리 이미지(I_q)의 시각 정보에 각각 대응되며, r_d 및 r_q는 데이터베이스 이미지(I_d) 및 쿼리 이미지(I_q)의 텍스트 정보에 각각 대응되고, $\mu \in [0,1]$는 미리 설정된 텍스트 유사성의 가중치를 제어하는 파라미터일 수 있다.

[0119] 본 발명의 일 실시 예에 따르면, 인터넷의 발전으로 인해 매일 인터넷에 업로드 되는 수백만 개의 이미지를 텍스트 또는 이미지를 이용하여 효율적으로 검색할 수 있다. 본 발명은 웹 애플리케이션 및 모바일 애플리케이션에 적용 가능하며, 사용자는 스마트 폰 또는 인터넷 브라우저를 사용하여 텍스트, 이미지 또는 둘 다를 사용하여 관련 이미지를 매우 손쉽고 간단하게 쿼리할 수 있다. 각 이미지의 처리는 독립적으로 수행될 수 있어, 본 발명에 의하면 빠른 이미지 처리가 가능하며, 따라서 애플리케이션에 실제적인 적용이 가능하다. 또한, 각 이미지는 시각 단어 및 텍스트 단어의 작은 벡터로 표시되기 때문에 저장 공간을 줄이고 네트워크 대기 시간을 줄일 수 있다는 장점이 있다.

[0120] 본 명세서에서 생략된 일부 실시 예는 그 실시 주체가 동일한 경우 동일하게 적용 가능하다. 또한, 전술한 본 발명은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 있어 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 여러 가지 치환, 변형 및 변경이 가능하므로 전술한 실시 예 및 첨부된 도면에 의해 한정되는 것이 아니다.

부호의 설명

[0122] 100: 검색 데이터 생성 모듈

110: 데이터베이스

130: 제 1 전처리부

150: 제 1 배경분포 토픽모델링부

200: 검색 모듈

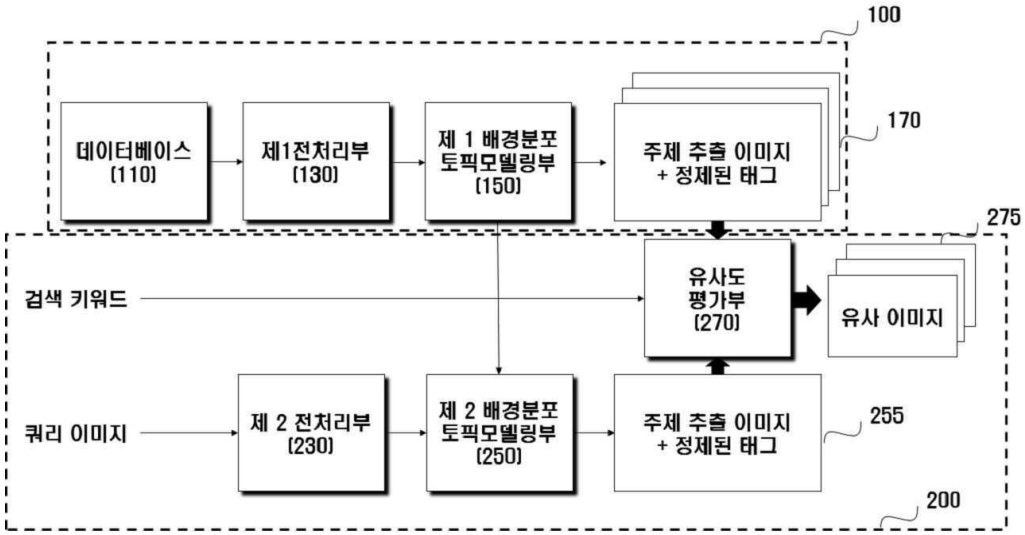
230: 제 2 전처리부

250: 제 2 배경분포 토픽모델링부

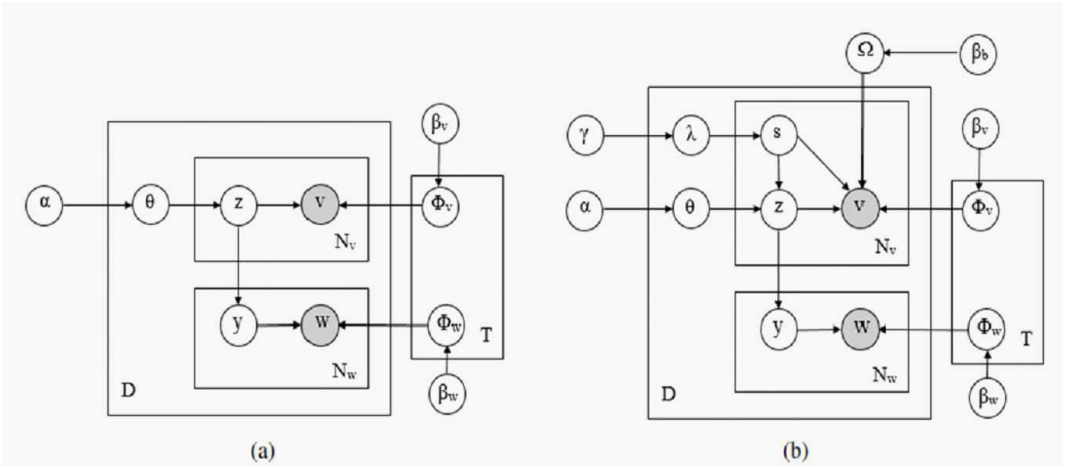
270: 유사도 평가부

도면

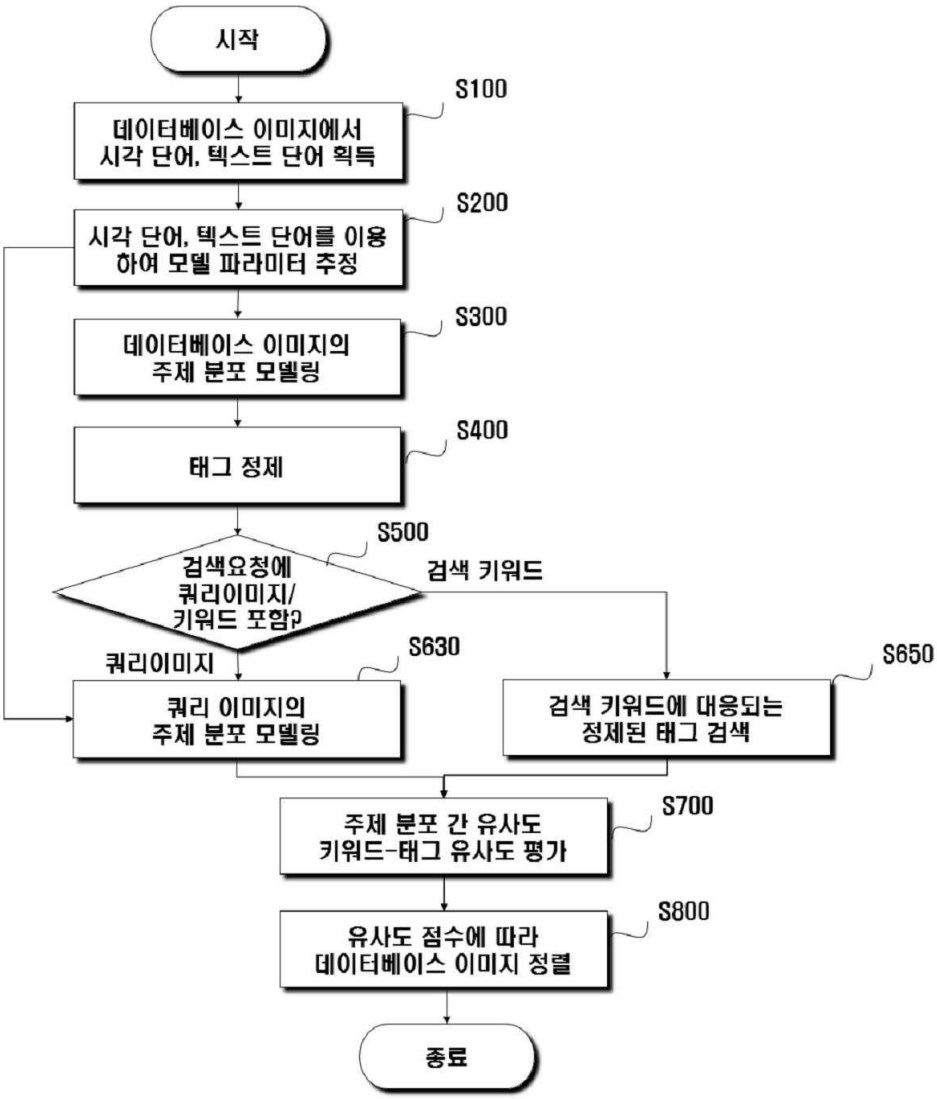
도면1



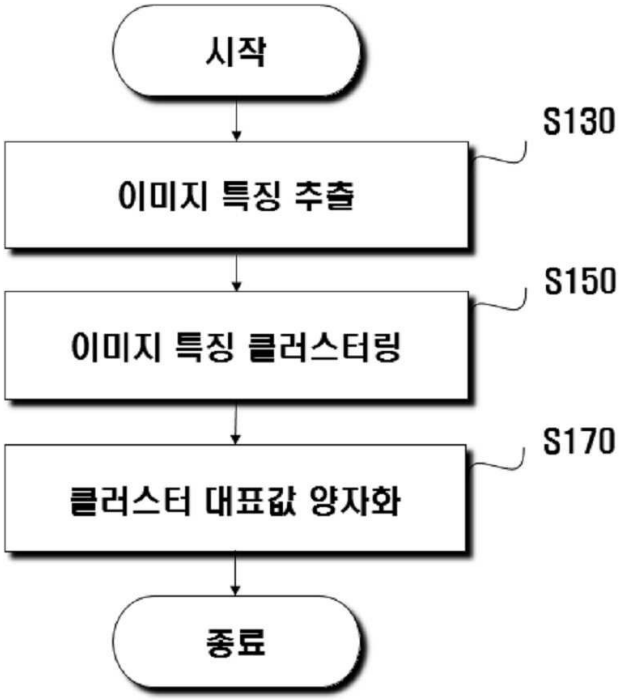
도면2



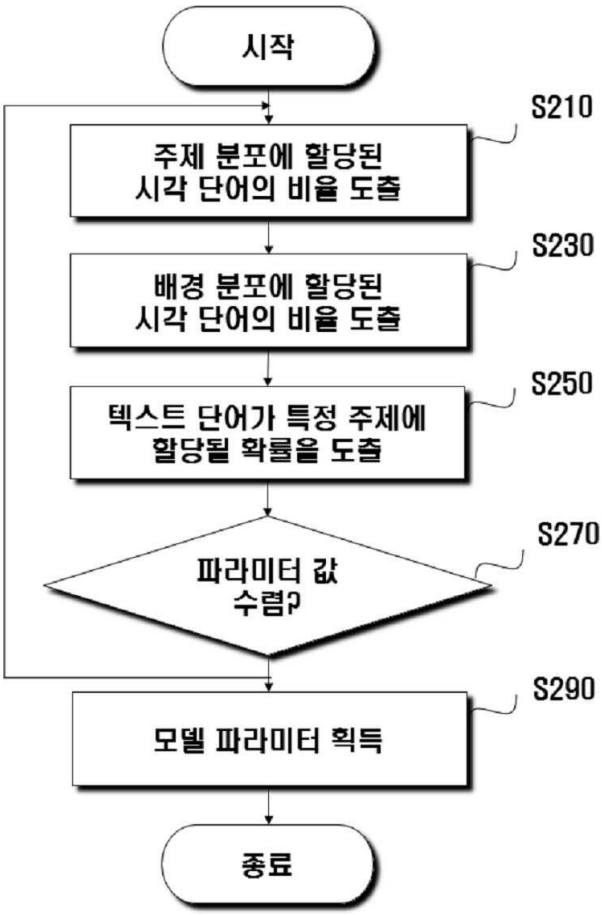
도면3


시작
S100
데이터베이스 이미지에서 시각 단어, 텍스트 단어 획득
S200
시각 단어, 텍스트 단어를 이용하여 모델 파라미터 추정
S300
데이터베이스 이미지의 주제 분포 모델링
S400
태그 정제
S500
검색요청에 쿼리이미지/키워드 포함?
검색 키워드
쿼리이미지
S630
쿼리 이미지의 주제 분포 모델링
S650
검색 키워드에 대응되는 정제된 태그 검색
S700
주제 분포 간 유사도 키워드-태그 유사도 평가
S800
유사도 점수에 따라 데이터베이스 이미지 정렬
종료

도면4


시작
S130
이미지 특징 추출
S150
이미지 특징 클러스터링
S170
클러스터 대표값 양자화
종료

도면5


시작
S210
주제 분포에 할당된
시각 단어의 비율 도출
S230
배경 분포에 할당된
시각 단어의 비율 도출
S250
텍스트 단어가 특정 주제에
할당될 확률을 도출
S270
파라미터 값
수렴?
S290
모델 파라미터 획득
종료