

지역적 유사도를 이용한 이미지 색상 정보 추출

손정우[†] 박성배[†] 김상수^{†○} 김구진[‡]

경북대학교 전자전기컴퓨터학부

[†]기계 학습 연구실 [‡]컴퓨터 그래픽스 연구실

{jwson, sbpark, sskim[○]}@sejong.knu.ac.kr, kujin@knu.ac.kr

Extraction of Color Information from Images using Grid Kernel

JeongWoo Son[○] Seong-Bae Park Ku-Jin Kim

Kyungpook National University

School of Electrical Engineering and Computer Science

Machine Learning Lab., Computer Graphics Lab.

요 약

본 논문에서는 이미지 상에 나타난 색상 정보를 추출하기 위한 새로운 커널 메소드(Kernel method)인 Grid kernel을 제안한다. 제안한 Grid kernel은 Convolution kernel의 하나로 이미지 상에 나타나는 자질을 주변 픽셀에서 나타나는 자질로 정의하고 이를 재귀적으로 적용함으로써 두 이미지를 비교한다. 본 논문에서는 제안한 커널을 차량 색상 인식 문제에 적용하여 차량 색상 인식 모델을 제안한다. 이미지 생성시 나타나는 주변 요인으로 인해 차량의 색상을 추출하는 것은 어려운 문제이다. 이미지가 야외에서 촬영되기 때문에 시간, 날씨 등의 주변 요인은 같은 차량이라 하더라도 다른 색상을 보이게 할 수 있다. 이를 해결하기 위해 Grid kernel이 적용된 차량 색상 인식 모델은 이미지를 HSV (Hue-Saturation-Value) 색상 공간으로 사상하여 명도를 배제하였다. 제안한 커널과 색상 인식 모델을 검증하기 위해 5가지 색상을 가진 차량 이미지를 이용하여 실험을 하였으며, 실험 결과 92.4%의 정확율과 92.0%의 재현율을 보였다.

1. 서 론

지능형 교통 시스템(ITS)은 급증하는 차량을 통제하여 운전자에게는 보다 편안한 도로 환경을 보행자에게는 더 안전한 보행 환경을 제공하기 위한 시스템이다. ITS를 구축하기 위해서는 현재 도로 상에 나타난 차량의 정보를 수집하는 것이 필수적인 요소이다. 차량에서 추출할 수 있는 정보 중 색상 정보는 추출하기가 비교적 용이하고 다른 정보의 추출에 사용될 수 있는 low-level 정보이기 때문에[1, 2], ITS를 구축할 때, 차량의 색상 정보 추출은 우선적으로 고려되어야 한다.

차량의 색상 정보 추출은 형태 정보 추출이나 특정 부분의 인식 문제에 비해 쉽다고 인식되지만 여전히 어려운 문제이다. 차량은 한가지 색으로 이루어져 있지 않으며 도로상의 차량을 촬영했을 경우, 주변 풍경의 색상이나, 이미지가 생성될 때의 날씨와 같은 환경적 요소들이 같은 색이라 할지라도 이미지에 따라 다른 색으로 보일 수 있기 때문이다.

현재까지의 연구들은 이미지의 색상 정보로써 평균 색상, 형태 정보 색상 히스토그램 등을 이용하였으나[3], 여러 색상이 함께 나타나는 물체의 경우, 색상 정보를 추출하기 힘들다는 단점이 있어 각 분야에 맞는 자질을 추

가하여 사용하고 있다[4].

차량의 색상 인식 문제를 기계학습의 관점에서 보면, 분류 문제로 볼 수 있다. 임의의 이미지에 나타난 차량을 기존에 정해진 색상 클래스로 분류함으로써 이미지 상에 나타난 차량의 색상을 찾을 수 있다. SVMs(Support Vector Machines)는 이진 분류문제를 풀기 위해 제안된 classifier로써 다양한 문제에 적용되고 있으며 높은 성능을 보여주고 있다. SVMs를 이용하여 분류문제를 풀기 위해서는 이미지를 자질 공간상에 사상하여야 한다. 하지만 이미지를 대표하는 자질을 명시적으로 정하기는 쉬운 일이 아니다.

이와 같이 자질을 명시적으로 정하기 어려운 경우, 기계학습 분야에서는 커널 메소드(kernel method)를 이용하여 명시적으로 자질을 추출하지 않고 두 물체사이의 유사도를 추출하고 있다[5]. 그 중, Convolution kernel[6]은 트리, 그래프, 문자열 등 자질 선택이 어려운 이산 자료 구조들 간의 유사도를 추출하기 위해 사용된다. Convolution kernel은 하나의 구조를 더 작은 구조들의 집합으로 간주하여 이들 부분 구조가 자질 공간을 구성하게 된다. 예로 파스 트리 커널[7]은 부분 트리가 자질 공간을 구성하며, 스트링 커널은 부분 문자열[8], 그래프 커널[9]은 부분 그래프가 자질이 된다.

본 논문에서는 이미지의 차량 색상 정보를 추출하기 위한 새로운 커널인 Grid kernel을 제안한다. 그리고 제안한 커널을 이용한 차량 색상 인식 모델을 제안한다. 제안한 모델에서는 색상 정보 추출을 위해 제안한 커널을 이용하여 SVMs(Support Vector Machines)을 학습시

킨다. 이 때, 이미지를 HSV 색상 공간에 사상하여 명도 값을 배제함으로써 주변 환경 중, 빛에 의해 이미지가 어둡거나 밝아 색상 정보가 변하는 것을 최대한 방지하였다.

본 논문의 구성은 다음과 같다. 먼저 2장에서 커널 메소드와 본 논문에서 제안하는 커널에 대해 설명한다. 3장에서는 차량 색상 인식 문제와 제안한 커널을 적용한 차량 색상 인식 모델에 대해 설명하고, 4장에서는 제안한 커널과 색상 인식 모델을 검증하기 위한 실험과 실험 결과를 보이고 마지막으로 5장에서 결론을 맺는다.

2. Grid Kernel(GK)

커널 메소드는 임의의 고차원 공간에 데이터를 사상하여 데이터간의 유사도를 계산하는 방법으로써 사용자가 자질을 명시적으로 정해주지 않아도 되기 때문에 자질을 정하기 힘든 데이터에 대해 많이 적용되고 있다[10].

Convolution kernel은 이산 자료 구조들 간의 유사도를 얻기 위해 사용되는 커널 함수의 한 종류이다. Convolution kernel에는 파스 트리를 다루는 파스 트리 커널(Parse tree kernel), 문자열을 다루는 스트링 커널(String kernel), 그래프를 다루는 그래프 커널(Graph kernel) 등 다양한 종류의 데이터를 다루는 커널들[11, 12]이 있다.

본 논문에서 제안하는 Grid kernel은 하나의 자질을 더 작은 단위의 근접한 자질로 이루어져 있다고 간주한다. 즉, 하나의 픽셀을 이루는 정보는 그 픽셀에 근접한 주변 픽셀의 정보로 이루어져 있다. 제안한 kernel은 이러한 자질을 명시적으로 나열하지 않고 두 픽셀의 유사도를 내적을 이용하여 계산한다.

가로가 w , 세로가 h 인 임의의 이미지 I 가 있다면 이 이미지 I 는 다음과 같은 자질로 이루어져 있다고 가정한다.

$$\text{Image } I = \{s_i^n, s_i^{n-1}, \dots, s_i^1, s_i^0\}$$

이때, s_i^n 는 이미지 상에 나타나는 깊이가 n 인 l 번째 자질을 의미한다. 자질 s_i^n 가 (i, j) 위치에서 나타났다면, 자질의 값은 아래 수식에서 보여 지는 것과 같이 깊이 $n-1$ 인 자질 중 s_i^{n-1} 의 주변에 위치한 자질들로 이루어진다고 정의한다.

$$s_i^n = \{s_i^{n-1}(i-1, j), s_i^{n-1}(i+1, j), \\ s_i^{n-1}(i, j-1), s_i^{n-1}(i, j+1), \\ s_i^{n-1}(i-1, j-1), s_i^{n-1}(i+1, j+1), \\ s_i^{n-1}(i-1, j+1), s_i^{n-1}(i+1, j-1)\}$$

이때, 함수 $s_i^{n-1}(i-1, j)$ 는 $(i-1, j)$ 위치에서 나타나는 깊이 $n-1$ 인 자질을 반환한다. 자질의 깊이는 0이 될 때까지 재귀적으로 감소하여 0일 경우 현재 위치의 픽셀 값을 의미한다. 이러한 자질을 이용하여 이미지 I 를 나타내는 벡터 V_I 는 아래와 같이 정의 한다.

$$V_I = \{\#s_i^n, \#s_i^{n-1}, \dots, \#s_i^1, \#s_i^0\}$$

위의 수식에서 $\#s_i^n$ 은 자질 s_i^n 가 나타난 빈도수를 뜻한다. 이와 같은 자질 정의는 특정 위치의 픽셀 값이 상위 자질과 하위 자질에서 재귀적으로 나타나게 되어 이미지의 중심 픽셀에 비중을 둔 이미지 벡터가 생성 되도록 한다.

두 이미지 벡터 V_1, V_2 의 내적은 아래 수식과 같이 정의 가능하다.

$$\begin{aligned} \langle V_{I_1}, V_{I_2} \rangle &= \sum_{s \in S} V_{I_1}^s \times V_{I_2}^s \\ &= \sum_j \sum_i \sum_{s \in S} Id_{I_1}^s(j, k) Id_{I_2}^s(j, k) \\ &= \sum_j \sum_i K(j, k, I_1, I_2, n) \end{aligned}$$

위의 수식에서 S 는 두 이미지 상에 나타나는 자질들의 집합이며, 함수 $Id_{I_1}^s(j, k)$ 는 (j, k) 위치에서 나타나는 픽셀이 자질 s 의 중심 픽셀과 같으면 1, 아니면 0을 반환하는 함수이다. 함수 $K(j, k, I_1, I_2, n)$ 는 아래와 같이 정의된다.

$$\sum_{s \in S} Id_{I_1}^s(j, k) Id_{I_2}^s(j, k)$$

이 때, 함수 $K(j, k, I_1, I_2, n)$ 는 아래와 같은 재귀 규칙에 의해 다항 시간(Polynomial time)에 계산 가능하다.

- 이미지 I_1, I_2 에서 (i, j) 위치의 픽셀이 서로 다르면,

$$K(j, k, I_1, I_2, n) = 0$$
- 이미지 I_1, I_2 에서 (i, j) 위치의 픽셀이 서로 같고, n 이 0이라면,

$$K(j, k, I_1, I_2, n) = \lambda$$

- 그 외,

$$K(j, k, I_1, I_2, n) = \lambda \prod_i^{np(j, k)} K(out_i(j, k).height, out_i(j, k).width, I_1, I_2, n-1)$$

이 때, 함수 $np(j, k)$ 는 (j, k) 에 위치한 픽셀의 주변 픽셀 수를 반환하고, 함수 $out_i(j, k)$ 는 (j, k) 에 위치한 픽셀의 주변 픽셀 중, i 번째 픽셀을 반환한다. 위의 수식에서 λ 는 감소 인자(decay factor)로 중심 픽셀에서 멀어질수록 더 작은 비중을 가지게 하기 위해 0에서 1 사이의 값을 가진다.

그림 2.1은 크기가 5 x 6인 같은 이미지를 비교 했을 때 나타나는 빈도수를 보여준다. 그림에서는 제안한 커널이 이미지의 중심에 비중을 두고 비교함을 보여준다.

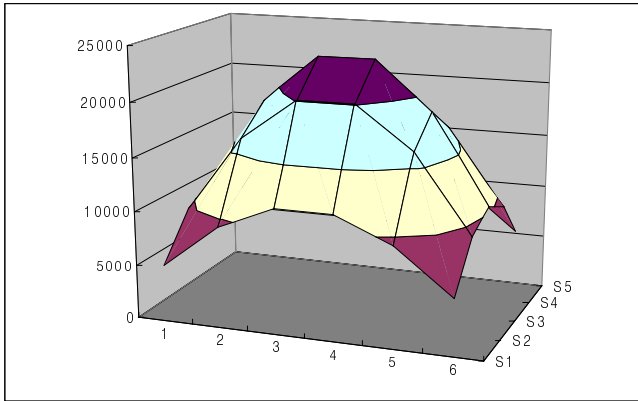


그림 2.1. 같은 이미지를 비교했을 때의 빈도수

3. 색상 인식 모델

본 논문에서 제안하는 색상 인식 모델은 제안한 커널과 SVMs[13]를 이용하여 모델을 생성하고 이미지를 분류하고 있다. 그림 3.1는 제안한 색상 인식 모델의 전체적인 흐름을 보여 준다. 먼저 이미지를 HSV공간으로 사상한다. 사상된 이미지와 제안한 커널을 이용하여 SVMs에서 모델을 생성한다. 그림 3.2는 생성된 모델을 이용하여 차량의 색상을 인식하는 과정을 보여준다.

RGB(red-green-blue) 색상 공간에서 HSV(hue-saturation-value) 색상 공간으로 사상한다. HSV공간으로의 사상은 아래 수식을 이용하여 쉽게 가능하다[3].

$$V = \max(r, g, b)$$

$$S = \frac{V - \min(r, g, b)}{V}$$

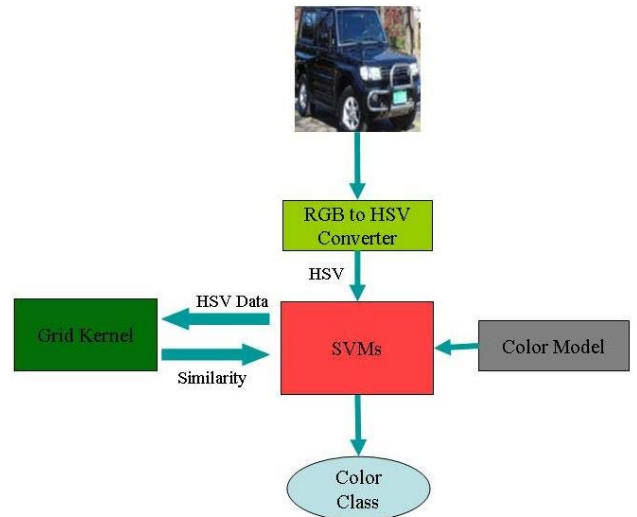


그림 3.2. 색상 인식 모델을 이용한 분류 과정

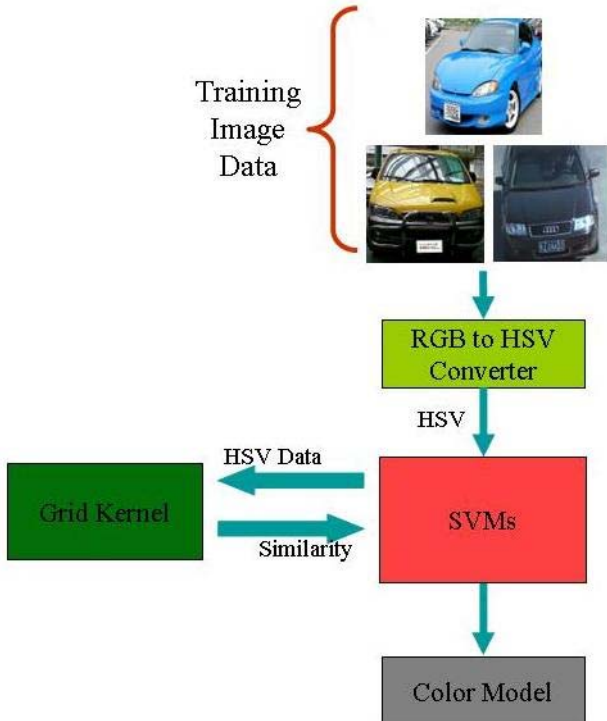


그림 3.2. 색상 인식 모델 생성 과정

차량 색상 인식은 야외에서 생성 되는 이미지의 특성상 차량의 색상이 시간, 날씨 등에 의해 변화할 수 있음을 감안해야 한다. 이를 해결하기 위해 본 논문에서 제안하는 색상 인식 모델은 차량 이미지를 먼저

$$H = \begin{cases} \text{if } r = \max(r, g, b) \text{ and } g = \min(r, g, b) \\ 5 + \hat{b} \\ \text{if } r = \max(r, g, b) \text{ and } g \neq \min(r, g, b) \\ 1 - \hat{g} \\ \text{if } g = \max(r, g, b) \text{ and } b = \min(r, g, b) \\ 1 + \hat{r} \\ \text{if } g = \max(r, g, b) \text{ and } b \neq \min(r, g, b) \\ 3 - \hat{b} \\ \text{otherwise} \\ 5 - \hat{r} \end{cases}$$

HSV 색상 공간은 밝기 값을 V 공간에 할당하기 때문에 이를 배제함으로써 빛에 의해 생기는 변화를 감소시켰다. 사상된 모든 픽셀의 색상 값 중 V 값을 배제하고 계산하기 위해 제안한 커널 메소드의 재귀 규칙을 아래와 같이 수정하였다.

- $S(I_1, I_2, i, j) = 0$
- $K(j, k, I_1, I_2, n) = 0$
- $S(I_1, I_2, i, j) = 1$ 이고, n 이 0이라면,
- $K(j, k, I_1, I_2, n) = \lambda$

- 그 외,

$$K(j, k, I_1, I_2, n) = \lambda \prod_i^{n_p(j, k)} K(out_i(j, k).height, out_i(j, k).width, I_1, I_2, n - 1)$$

이때, 함수 S 는 (i, j) 위치의 두 픽셀이 서로 같으면 1을 다르면 0을 반환하는 함수이다. 함수 S 를 아래와 같이 정의 하여 V 값을 배제하고 각 픽셀의 H 와 S 값이 10 이상 차이 나지 않는다면 같은 색상 값을 가지는 것으로 간주한다.

- $|pixel(i, j, I_1).h - pixel(i, j, I_2).h| > 10$
 $S(I_1, I_2, i, j) = 0$
- $|pixel(i, j, I_1).s - pixel(i, j, I_2).s| > 10$
 $S(I_1, I_2, i, j) = 0$
- 그 외,
 $S(I_1, I_2, i, j) = 1$

이 때, 함수 $pixel(i, j, I_1)$ 은 (i, j) 위치의 픽셀을 반환한다.

4. 실험

제안한 커널 메소드와 색상 인식 모델을 평가하기 위해 자동차 이미지를 색상에 따라 분류하는 실험을 하였다. 실험에 사용된 데이터는 학습 데이터 450개와 실험 데이터 50개로 구성된다. 이미지 데이터의 사이즈는 100×100 으로 단일화하였다. 실험에서는 SVM light package[13]를 사용하여 파란색, 빨간색, 노란색, 흰색, 검정색의 차량을 분류하도록 하였다. SVMs는 기본적으로 이진 분류를 하기 위한 메소드이기 때문에 5가지 색상으로 분류하기 위해 실험에서는 one versus all 기법을 사용하였다. 제안한 커널 메소드의 상대적인 성능을 평가하기 위해 linear 커널(LK)과 비교하였다. 표 4.1은 각 class 별 실험 결과를 보여준다.

| | 파란색 | 빨간색 | 노란색 | 흰색 | 검정색 |
|-----|-------|------|------|-------|-------|
| 정확율 | 90.9% | 100% | 100% | 81.8% | 88.8% |
| 재현율 | 100% | 90% | 100% | 90% | 80% |

표 4.1. 제안한 커널을 이용한 실험 결과

실험 결과, 파란색, 빨간색, 노란색의 유채색 계열 class가 높은 정확율과 재현율을 보였다. 하지만 검정색 차량을 빛의 반사로 인해 흰색으로 분류하는 경우, 반대로 흰색 차량을 범퍼 등의 부분적으로 나타난 색상으로 인해 검정색으로 분류하는 경우 등, 무채색 계열의 class의 경우 상대적으로 낮은 성능을 보였다. 검정색 차량의 경우 2개의 데이터가 잘못 분류되었다. 잘못 분류된 이미지의 경우 과도한 빛으로 인해 검정색이 파란색 혹은 흰색으로 보이는 경우였다. 그림 4.1은 잘못 분류된 이미지를 보여준다. 그림에서 첫번째 이미지의 차량은 파란색으로 두번째 이미지는 흰색으로 분류되었다.



그림 4.1. 잘못 분류된 검정색 차량 이미지

흰색 차량 이미지의 경우 1개의 데이터가 검정색으로 분류되었다. 이 이미지는 전체적으로 어두운 주변 환경과 차량 범퍼의 짙은 검정색이 잘못된 정보를 생성하여 검정색으로 분류된 것으로 보인다. 그림 4.2는 검정색으로 분류된 흰색 차량 이미지를 보여준다.



그림 4.2. 검정색으로 분류된 흰색 차량 이미지

빨간색 차량의 경우 보여 지는 차량의 붉은 색상 부분이 적고 다른 색상이 많이 섞여 있으며, 이미지의 30%를 차지하는 운전자의 모습이 노이즈로 작용한 것으로 보인다. 그림 4.3은 잘못 분류된 빨간색 차량의 이미지를 보여준다. 그림에서 차량 이외의 물체가 차량에 비해 크게 부각됨을 알 수 있다.

Linear 커널(LK)과의 비교 실험에서 제안한 커널은 더 나은 성능을 보여준다. 특히, 파란색의 경우 LK의 경우, 모두 검정색으로 분류하여 주변 환경에 의해 유사한 색상을 나타낼 경우 분류가 불가능함을 보였으나 제안한 커널(GK)은 91%의 정확율과 100%의 재현율을 보여 더 나은 성능을 보임을 알 수 있다.



그림 4.3. 흰색으로 분류된 빨간색 차량 이미지

빨간색 차량 이미지의 경우도 제안한 커널이 50% 더 높은 재현율을 보였다. LK가 가장 잘 분류한 노란색 차량 이미지의 경우도 제안한 커널보다 못한 성능을 보였

다. 제안한 커널이 가장 낮은 성능을 보이는 흰색, 검정색 클래스의 경우에도 LK보다 각각 50%이상 나은 정확율을 보였다. 그림 4.4는 두 커널의 비교 실험 결과를 보여 준다.

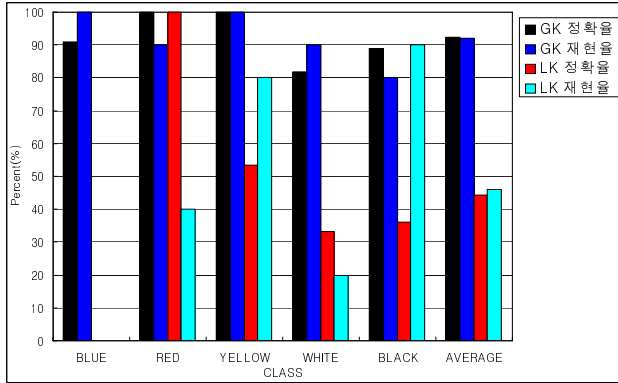


그림 4.3. 비교 실험 결과

5. 결 론

본 논문에서는 이미지의 색상 정보를 추출하기 위해 새로운 커널 메소드인 Grid kernel을 제안하고 이를 색상 인식 모델에 적용하였다. 제안한 Grid Kernel은 특정 픽셀에서 나타나는 자질을 주변에서 나타나는 자질로 표현하고 이를 재귀적으로 적용함으로써 이미지들 간의 유사도를 계산한다.

색상 인식 모델은 차량 이미지를 대상으로 차량의 색상을 인식하는 모델으로써, HSV 색상 공간을 사용하여 이미지를 표현한다. 커널을 이용하여 이미지의 유사도를 계산할 때 주변 환경 요인에 의한 색상 변화를 최대한 배제하기 위해 명도인 V 값을 배제하고 계산하였다. 제안한 모델에서는 먼저 학습 데이터와 커널을 이용하여 모델을 생성한 후, 생성된 모델을 이용하여 이미지 데이터를 분류한다.

제안한 Grid kernel과 색상 인식 모델을 평가하기 위해 간단한 실험을 하였다. 실험에서는 5가지 색상의 차량 이미지를 분류하도록 하였으며, Support Vector Machine을 사용하였다. 학습을 위한 데이터로 각 색상 당 90개씩 450개의 데이터를 사용하였으며 테스트 데이터로 각 색상 당 10개씩 50개의 데이터를 분류하였다.

실험 결과 5가지 색상에 대해 평균 92.4%의 정확율과 92.0%의 재현율을 보였다. 이와 같은 결과는 비교 실험한 linear 커널에 비해 정확율에서 48.0%, 재현율에서 46.0% 향상된 결과로 제안한 커널이 이미지 상에 나타난 색상정보를 잘 반영하고 있음을 보여준다.

제안한 커널은 비록 더 나은 성능을 보였으나 계산량의 차이로 인해 더 많은 시간을 필요로 한다. 이는 파스 트리 커널이나 스트링 커널 등, 다른 Convolution kernel에서도 나타나는 문제점[6,7,8]으로 제안한 커널의 경우 이미지 사이즈가 커질수록 지수적으로 시간이 늘어난다. 이를 해결하기 위해 커널 내부의 계산을 더 빠르게 수행

하기 위한 연구가 필요할 것으로 보인다.

참 고 문 헌

- [1] Y. Rui, T. Huang, and S. Chang, Image retrieval: current techniques, promising directions and open issues, *Journal of Visual Communication and Image Representation*, Volume 10, number 4, pages 39-62, 1999.
- [2] A. Frome, Y. Singer, and J. Malik, Image Retrieval and Classification Using Local Distance Functions, *Advances in Neural Information Processing Systems 20*, 2006.
- [3] J. Smith, S. Chang, Tools and techniques for color image retrieval, In: I. Sethi, R. Jain, *Storage & Retrieval for Image and Video Databases IV*, volume 2670 of IS&T/SPIE Proceedings, pages 426-437, 1996.
- [4] D. Finlayson, B. Schiele, and J. Crowley, Comprehensive Colour Image Normalization, In *Proceedings of 5th European Conference on Computer Vision*, pages 475-490, 1998
- [5] C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
- [6] D. Haussler, Convolution kernels on discrete structures, Technical report, UCS-CRL-99-10, UC Santa Cruz, 1999.
- [7] M. Collins and N. Duffy, Convolution kernels for natural language, In *Proceedings of the 15th Neural Information Processing Systems*, pages 625-632, 2001.
- [8] H. Lodhi, C. Saunders, J. Shawe-Taylor, N. Cristianini, and C. Watkins, Text classification using String Kernels, *Journal of Machine Learning Research*, 2:419-444, 2002.
- [9] J. Suzuki, T. Hirao, Y. Sasaki, and E. Maeda, Hierarchical Directed Acyclic Graph Kernel: Methods for Structured Natural Language Data, In *Proceedings of the 41th Annual Meeting of Association for Computational Linguistics(ACL 2003)*, pages 32-39, 2003
- [10] J. Shaw-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*, Cambridge University Press, 2004.
- [11] A. Moschitti, Efficient Convolution Kernels for Dependency and Constituent Syntactic Trees. In *Proceedings of the 17th European Conference on Machine Learning*, pages 318-329, MIT Press, Cambridge, MA, 2006.
- [12] S. Vishwanathan and A. Smola, Fast kernels on strings and trees, *Advances in Neural Information Processing Systems 15*, pages 569 - 576, MIT

Press, Cambridge, MA, 2003.

- [13] V. Vapnik, S. Golowich, and A. Smola, Support Vector Method for Function Approximation, Regression Estimation and Signal Processing, *Advances in Neural Information Processing Systems 9*, pages 281-287, 1996.
- [14] T. Joachims, *Making Large-Scale SVM Learning Practical*, LS8-Report, 24, University Dortmund, LS VIII-Report, 1998.