

다중 해상도 피라미드 기반 영상 인식자

박제호^{*†}

^{*†} 단국대 소프트웨어학과

Multi-resolution Pyramid based Image Identification

Je-Ho Park^{*†}

^{*†} Dankook University, Dept. of Software Science

ABSTRACT

Unlike modern photography technology, in the early days, efforts to physically compose an image with a concept similar to the current photograph have not been popular or commercially successful. The limitation of the use of images as artistic media or recordings has reached the stage of introducing the technology of image analysis to automate the function that humans recognize and judge through vision. In addition, the accuracy of the image has exceeded the human visual ability, enabling the technology that enables the step of recognizing and informing the fact that the human is not aware of it. Based on such a base, the range that can be applied through the image data in the future era can be said to be unpredictable, and the technology that targets large scale image database instead of an image is also expanding the possibilities as a new application technology. In order to identify a particular image from a massive database, different methodologies have been introduced. In this paper, we discuss image identifier production methods based on multi-resolution pyramid.

Key Words : Image, Entropy, Optimization, Identification, Image Pyramid

1. 서 론

현대 기술에서 빛을 이용하여 영상을 착상시키는 카메라 방식과는 달리, 하나의 핀홀(pinhole)을 이용하여 투사면에 영상을 착상시키는 아이디어는 1604년에 사용되었지만[1], 기원전 500년 전에 원천적인 아이디어는 인간의 역사에 나타났다는 기록이 있다[2]. 과거에는 현재 사진과 유사한 개념으로 영상을 물리적으로 착상 시키려는 노력은 대중적이거나 상업적 성공을 거두지는 못했지만, 미래 기술에 대한 가능성을 내포하는 원시적인 접근을 시도를 거듭하면서 현재에 이르게 되었다. 인간은 사진이라는 기계에 매료되어 광학적 화학적 방법을 응용해 영상을 생산하는 아날로그 카메라의 개발에 많은 투자를 하면서

상업적 성공과 실패라는 경험을 축적하면서 영상 기술 획득이라는 목표를 성취하였다. 1970년대 CCD기술 개발을 시점으로, 1991년에 현재 우리가 사용하는 디지털 카메라의 형태가 상업적으로 생산되고, 2002년 이후로 디지털 카메라는 기존의 아날로그 카메라 기술을 앞서면서 새로운 영상 기기의 획기적인 변곡점을 찍었다. 이러한 디지털 영상 기기의 발전은 카메라라는 물리적 한계를 넘어설 수 있는 소형 컴포넌트로 발전하면서, 다양한 응용 프레임으로 발전을 했다. 현재 영상이 가지는 정보의 가치는 물리적 영상의 한계에 머물지 않고 새로운 가능성을 확대할 수 있는 정보로서, 또는 인간 사회를 이해할 수 있는 새로운 지식을 창출할 수 있는 원천 정보로서의 활용 범주를 넓혀가고 있다.

기본적으로 영상을 획득할 수 필요한 기기는 다양한 프레임에 걸쳐 다양하게 성능과 형태를 확대시켜 왔다. 모바일 기기, 노트북, 고화질 카메라, CCTV, 인공위성, IoT

[†]E-mail: dk_jhpark@dankook.ac.kr

영상 모듈 등을 예로 들 수 있다. 획득된 영상을 분석하고 활용하는 응용 프레임은 지속적으로 개선되고 있으며, 결과적으로 수집되는 영상의 분량은 가능할 수 없을 정도가 되었다. 이러한 영상들을 활용하기 위해서는 영상을 가공, 편집, 수집, 저장 등의 과정을 처리해야 할 수 있는 시스템이 절대적으로 필요하다. 또한, 생성된 영상을 대용량 저장할 수 있는 기능에 대한 필요성이 대두되면서 영상 대용량 시스템의 효율성과 신뢰성에 대한 요구는 날로 높아지고 있다[3].

영상 처리 시스템에 저장된 각 영상을 식별하는 방법은 크게 두 가지로 나눌 수 있다. 첫번째 방법은 사용자가 영상에 임의의 식별 기능이 있는 이름을 부여하는 방법이다. 이러한 임의적인 식별자 부여 방법은 특정 방법론을 사용하지 않고 임의적 의미나 개별적 발상을 통해 수행하는 방법이다. 이러한 방법은 시스템에 적용하기 힘든 환경이므로 실제로는 논의에서 제외된다.

특정 방법론을 사용하는 경우에, 장점은 단순한 방법에 의한 편의성을 들 수 있다. 단점으로는 식별자로 활용되는 식별체계의 완전성이 보장될 수 없기에 발생하는 문제를 고려하여야 한다. 예를 들어, 일정 임계값을 가지는 순차적 숫자를 사용하면서, 임계값까지 사용할 수 있는 숫자를 전부 사용하게 되면 이미 사용된 숫자 기반 식별자를 순환하여 사용할 수 밖에 없고, 따라서 식별자의 중복은 불가피하다. 또한, 독립적인 다수의 영상 저장 시스템을 통합하는 경우 동일 식별자가 발생할 수 있다. 이 문제를 해결하기 위해서는 정밀한 방법으로 영상을 비교하는 과정을 피할 수 없기 때문에 대용량 시스템의 경우, 이러한 문제는 시스템의 성능을 저하시킬 뿐 아니라 완벽한 해결책을 고안하기가 힘들다고 할 수 있다.

임의적 식별 기능 부여 방법을 해결하기 위하여, 영상의 내부적인 속성을 이용하여 식별자 생성에 활용하는 방법의 사용이 가능하다. 이 방법론은 특정 속성을 영상으로부터 추출하여 정량적 변환을 통하여 식별자 생성에 속성값을 이용하기 때문에 임의성이 가지는 식별자 중복성을 극복할 수 있다. 여러 가지 방법론 중, 영상 영역에서 추출할 수 있는 직선 성분을 활용하여 계수화 하는 직선성분 계수기반 다중 인덱싱 구성 방법을 예로 들 수 있다[4]. 이러한 방법론은 영상 영역 전체를 식별자 생성에 사용하거나 영상의 일부만을 식별자 생성에 사용할 수 있다. 영상의 일부만을 식별자 생성에 사용할 경우는 식별자의 다양성을 확보할 수 있는 영상의 하부영역을 구별하는 것이 중요하게 되며, 이를 해결하기 위해 엔트로피 개념을 이용하는 방법이 제안되었다[5].

이러한 영상 엔트로피를 영상 식별자에 활용하는 경우, 중요한 결정은 식별자의 다양성을 확보할 수 있는 영상

의 하부영역의 선택 과정이 중요한 역할을 한다. 기존의 방법론은 이러한 중요한 속성을 단일 하부영역에서 추출하는 방법론을 사용하였다. 본 논문에서는 영상 식별자의 다양성 확보를 개선할 수 있도록 다수의 하부영역에서 식별자를 생성할 수 있는 방법론에 대한 논의를 하고자 한다.

2. 본 론

영상 식별자는 앞에서 논의한 것과 같이 영상과 영상의 분간을 영상 자체를 비교하여 결론을 내는 것이 아니라, 영상 비교 알고리즘과는 비교가 되지 않을 만큼 크기가 작고 데이터베이스화에 적절한 대표값을 이용하여 영상의 존재성을 검증하고 응용하고자 하는 것이다. 식별자와 영상 간의 결합에 관한 무결성만 보장이 된다면, 유효성이 인정되는 영상 식별자는 여러 응용 프레임에서의 활용은 영상 시스템의 효율성을 개선하는 데 기여하는 바가 크다.

이 절에서는 먼저 영상 식별자의 유효성을 만족을 위하여 필요한 개념적 속성에 대하여 논의하고, 식별자 생성에 응용되는 배경에 대하여 논의를 하고, 본 논문에서 제안하고자 하는 방법론에 대해 논의한다.

2.1 영상 식별자를 위한 개념적 모델링

영상 분류는 사용하는 식별자 또는 인덱스가 개개의 영상을 분류하는 목적이 아니고 유사성을 표현하는 경우가 있다. 이러한 경우, 해당 영상들은 영상 집합 공간에서 유사성 정도에 따라 다수의 하부그룹 또는 클러스터를 구성하게 된다. 본 논문에서 고려하는 영상 식별자는 이와는 다르게, 식별자의 유효성을 강화하여 데이터베이스에서 데이터의 유일무이한 존재에 대한 기능을 하는 기값에 대응하는 기능을 부여하기 위함이다. 이러한 목적을 명세하기 위하여, 본 논문에서는 논의하는 영상 식별자는 다음과 같은 요구조건을 만족시키기 위한 것으로 한다.

1. 유일성: 특정 영상을 위해 생성되는 식별자는 전체 식별자 집합에서 유일한 값을 가진다. 즉, 해당 영상과 식별자 사이에는 1 대 1의 관계를 부여한다.
2. 효율성 속성: 식별자를 구성하는 여러 인자값들은 최적화되어야 한다. 이 요구 사항은 식별자를 이용하여 물리적 영상 집합을 관리하는 경우, 메타 데이터베이스의 관리 및 유지를 효율적으로 하기 위한 정보표현 측면에서 고려한다.
3. 식별자 생성 효율성 요구: 식별자를 생성하는 방법은

대용량 영상 관리 시스템에 사용할 수 있도록 최적화된 알고리즘을 고려한다.

영상 식별자 관련 방법론들이 식별자 생성 알고리즘이 대상 영상의 전체 영역을 데이터 공간으로 인식하는 경우, 자연스럽게 전체 식별자 생성에 필요한 비용은 높아지게 되며, 경우에 따라서는 알고리즘 복잡도의 증가를 피할 수 없다. 본 논문에서 고려하는 영상 식별자는 고비용 최적화와 효율성이라는 측면에서, 비용의 개선을 최대화하고 효율성의 손실을 최소화하고자 하는 접근을 시도한다.

본 논문에서 고려하는 영상 식별자의 구성은 다수의 인자값을 함께 사용하는 것이다. 이러한 인자값은 식별자 생성 알고리즘에서 생성되는 값 이외에도 변경될 수 없는 확정된 영상 크기, 넓이, 높이 등의 물리적 속성값을 포함한다. 따라서, 하나의 영상 인식자는 벡터 형태로 표현될 수 있으며, 식 1과 같이 정형화할 수 있다. 식 1에서 a_i 는 앞에서 설명한 값들의 정량화된 표현이다.

$$ImgID = (a_0, \dots, a_k) \quad (1)$$

특정 알고리즘을 사용하여 영상으로부터 추출하는 속성값을 응용하여 영상 식별자를 조합하는 경우, 영상을 구성하는 엣지(edge) 또는 라인 세그먼트의 분포도를 기반으로 알고리즘을 구성할 수 있다. 다른 방법으로는, 영상의 화소값 공간의 크기를 최적화된 크기로 변경하고, 새로운 화소 공간에서 인접하는 화소를 화소값에 기반하여 응집하여 클러스터를 결집하여, 클러스터의 분포를 활용하는 방법도 사용될 수 있다[6].

2.2 선형 영상 엔트로피

통계적 접근법을 통해 정보의 양을 측정하는 엔트로피 개념은 기본적으로 다음과 같이 정의된다.

난수적 이벤트 Event의 발생 확률이 $P(Event)$ 일 경우, $Info(Event)$ 로 표현되는 정보의 양은 식 2와 같다[7,8].

$$Info(Event) = \log(1/P(Event)) = -\log(P(Event)) \quad (2)$$

영상 분야에서 엔트로피 개념의 정량화는 화소의 반복성 또는 정렬성을 기반으로 화소들의 변화하는 크기를 측정한다. 즉, 엔트로피 분포를 해석할 때, 엔트로피 값이 작으면 영상의 특징점이 편향되는 경향이 있고, 엔트로피 값이 크면 특징점이 영상 영역 전반에 고르게 분포되어 있는 것으로 해석한다. 영상 엔트로피를 이용한 식별자 생성은 영상의 하부 영역을 결정 시에 응용한다. 높이와 넓이가

Height x Width 인 영상을 $m \times n$ 개의 하부영상으로 분할하면, 전체 하부영역의 수는 다음과 같다.

$$No_{subImages} = \lfloor Height/m \times Width/n \rfloor \quad (3)$$

각 하부영역에 대한 엔트로피를 다음과 같은 식을 이용하여 계산을 한다. 각 하부영역에 대한 엔트로피 $Entropy_{sub}$ 를 계산하기 위해서는 먼저 한 하부영역에서 하나의 그레이스케일 g 의 발생 빈도를 $freq(g)$ 를 해당 하부영역의 픽셀수 $Pixels$ 로 나눈 값 $freq/Pixels$ 으로 표현되면 이 는 식 4에 나타나 있다.

$$Entropy_{sub} = \sum_{grayscale} freq(g) \times \log_{10} freq(g) \quad (4)$$

엔트로피 개념을 이용하여 엔트로피 값으로 표현되는 역동성을 기반으로 역동성이 가장 큰 하부영역에서 식별자를 생성하는 방법을 생각할 수 있다. 위에서 논의한 엔트로피 값을 계산하는 방법에서 보이는 바와 같이, 엔트로피 값 계산은 영상 전체를 대상으로 알고리즘을 수행한다. 결과적으로 대략적인 계산 복잡도는 높이를 H 라고 하고, 넓이를 W 라고 하는 경우, $O(H \times W)$ 이 된다. 앞에서 기술한 바와 같이, 본 논문에서는 고비용의 최소화와 동시에 식별자의 정확도를 확보하는 것을 동기로 하고 있다. 이러한 해결 방법으로 선형적인 계산으로 영상 엔트로피를 대체할 수 있는 방법론을 통해 비용 절감 개선을 할 수 있다(Jeho, 2019)

기본적인 방법으로 고안된 방법은 지정된 선형 영역에 분포되어 있는 화소값의 변화율을 측정하는 것이다. 따라서, 그레이스케일 영상에서 시점 $S(x,y)$ 에서 종점 $E(x,y)$ 까지의 선형 영역 $L(S, E)$ 에 대한 역동성 변화율 E_{linear} 은 다음과 같이 측정할 수 있다.

```

C=0
for p; pixels(p0...pk) in L
  if g(pi) != g(pi+1) then C = C + 1
```

이를 식으로 표현하면 식 5와 6과 같이 된다.

$$\begin{aligned} Pixels_{line} &= \{p(x_i, y_i) : p(x_i, y_i) \in L(S, E) \\ &\text{where } p(x_0, y_0) = S(x, y) \text{ and} \\ &p(x_k, y_k) = E(x, y) \text{ and} \\ &k = \text{the number of pixels in } L\} \quad (5) \end{aligned}$$

$$E_{linear} = \{C \mid \text{if } s(x_i, y_i) \neq s(x_{i+1}, y_{i+1}), C=C+1 \text{ for } i=0..k\} \quad (6)$$

위와 같은 방법을 여러 선형 영역에 적용하여 기초적인 엔트로피 계산과의 역동성 변화율 측정 비교 분석은 선형적 역동성 측정을 통해서 가능하다(Jeho, 2019).

2.3 해상도 피라미드 기반 선형 엔트로피

주어진 영상의 하부영역 중에서 엔트로피가 높은 하부영역에서 영상 식별자를 추출 방법론은 영상 개개의 특성상 화소 집합의 역동성이 저조하여 식별자 생성 방법론의 기본적인 속성을 만족시키지 못하는 하부영역을 피하기 위해서 선택될 수 있는 방법 중 하나이다. 이러한 방법을 선택하는 경우, 식별자의 요구조건을 만족할 수 있는 하부영역의 선택은 하부영역 크기의 최적화에 의해 영향을 받을 수 있다.

본 논문에서 논의하는 방법론은 이러한 특정 하부영역의 선택에 의해 발생할 수 있는 방법론의 단점을 보완하여 생성된 식별자의 기본 요구조건에 대한 충족도를 높이고자 하는 것이다. 논의하고자 하는 방법론은 두 단계로 나누어지며, 먼저 첫번째 단계는 식 7과 같이 주어진 영상을 주어진 크기 $m \times n$ 개의 하부영역으로 구분한다. 분할된 각 하부영역은 하나의 점에 대응하여, 가상의 공간 HyperPlane을 구성하게 되며, 식 8과 같이 각 점은 엔트로피 값 Ent_i 속성을 가진다.

$$Image = \{\text{block}(i, j) \mid \sum \text{block}(i, j) \text{ where } 0 \leq i \leq m, 0 \leq j \leq n\} \quad (7)$$

$$HyperPlane = \{Ent_{ij} \mid Ent_{ij} = Entropy(\text{block}(i, j)) \text{ where } 0 \leq i \leq m, 0 \leq j \leq n\} \quad (8)$$

위와 같이 구성된 HyperPlane 에서 최고의 엔트로피를 가지는 점을 Ent_{max} 라 정의한다. HyperPlane 이 가지는 엔트로피 값을 순차적으로 가장 큰 값에서 작은 값으로 정렬했을 때, 가장 큰 엔트로피 값에서 k 번째로 큰 값을 Ent_k 라고 할 때, 순차적으로 Ent_{max} 값부터 Ent_k 값을 가지는 HyperPlane 상에서 점의 집합을 $Kemel_{points}$ 를 식 9와 같이 정의한다.

$$Kemel_{points} = \{KPoint(x, y) \mid Ent_k \leq Ent_{ij} \leq Ent_{max} \text{ where } 0 \leq i \leq m, 0 \leq j \leq n\} \quad (9)$$

위와 같이 HyperPlane 을 구성하고 $Kemel_{points}$ 가 결정되면 최종적인 식별자는 다음과 같은 두 가지 방법으로 생

성할 수 있다: (1) 하부영역 식별자 배열, (2) 내포 영역 식별자. 하부영역 식별자 방법론은 $Kemel_{points}$ 에 속하는 점에 속하는 각각의 하부영역에 대한 식별자를 구성요소로 다단계 식별자 $ImgID_{Seq}$ 를 구성하며 이는 식 10과 같이 정의된다.

$$ImgID_{Seq} = \{(SubID_0, \dots, SubID_k) \mid SubID_k = \text{img. identifier for } Kernel_{points} \text{ and } Entropy(SubID_{i+1}) \leq Entropy(SubID_i) \text{ where } 0 \leq i \leq k-1\} \quad (10)$$

내포 영역 식별자는 다시 두 가지 방법으로 분류가 된다: (1) 총계 내포 영역 식별자, (2) 가상 내포 영역 식별자. 총계 내포 영역 식별자는 $Kemel_{points}$ 에 속하는 하부영역에 대한 각각의 인식자를 통합하는 방법으로 식 11과 같이 정의된다.

$$ImgID = \{(a_0, \dots, a_k) \mid a_i = \sum s_j \text{ where } ImgID \text{ for block}(l, m) = (s_0, \dots, s_k) \text{ and } KPoint(l, m) \in Kernel_{points}\} \quad (11)$$

가상 내포 영역 식별자는 Ent_{max} 를 가지는 하부영역을 포함하고, $Kemel_{points}$ 에 포함된 하부영역을 최대한 포함하는 새로운 하부영역 $KemelPlane$ 을 식 12과 같이 정의한다.

$$KemelPlane = \{\text{block}(i, j) \mid \text{block for } Ent_{max} = Entropy(\text{block}) \text{ and } \exists \text{block}(i, j) \text{ so that } KPoint(i, j) \in Kernel_{points}\} \quad (12)$$

이렇게 구성된 $KemelPlane$ 은 Ent_{max} 값을 가지는 하부영역과 그 다음으로 큰 엔트로피 값을 가지는 하부영역을 최대한 포함하는 새로운 하부영역으로 정의된 후, 해당 영역에 식별자 생성 알고리즘을 적용하게 된다. 이로 인해 기대하는 것은 엔트로피를 낮은 영역을 제외하고 하부영역에 대한 식별자 생성 시 생기는 단점을 최대한 보완하는 동시에 본래의 알고리즘이 가지는 장점을 부각하여 최대한 비용의 감소를 이루면서, 최대의 효율적인 식별자를 생성하고자 함이다.

3. 결론

본 논문에서 논의하는 영상 식별자는 기본적인 식별자의 요구조건을 최대한 만족시키면서, 그에 따른 계산 비용을 최대한 감소시키는데 있다. 기존의 방법이 비용을

최소하기 위한 알고리즘에 초점을 맞추고 있다면 본 논문에서는 기존의 방법론을 발전시켜 잠재적으로 발생할 수 있는 식별자로서의 불완전성을 개선하는데 있다고 할 수 있다. 이를 위해서 기존에는 정해진 하부영역을 중심으로 영상 식별자를 생성하는 방법을 선택하였다면, 본 논문에서 제안하는 방법론은 하부영역의 한계를 확장하여 안정성이 개선된 식별자를 생성하는데 그 의미가 있다고 할 수 있다.

참고문헌

1. "Introduction to the Camera Obscura", *Science and Media Museum*. 28 Jan. 2011, Retrieved 6 March (2019).
2. Boulger, Demetrius Charles (1969). *The Asiatic Review*.
3. Yongcheol Jeong, "The amount of the global digital data is increasing rapidly so that, 2020, the amount would reach 40ZB... it is 300 times of the amount of 2005", *Digital Times*, Dec. (2012).
4. Je-Ho Park, "Still Image Identifier based over Low-frequency Area", *Journal of Digital Contents Society*, 11(3), pp 393-398, Sep. (2010).
5. Je-Ho Park, "Noble Approach of Linear Entropy based Image Identification", *Journal of the Semiconductor & Display Technology*, 18(3), pp 31-35, Sep. (2019).
6. Je-Ho Park, Taeg Keun Whangbo, Kuinam J. Kim, "A Novel Image Identifier Generation Method Using Luminance and Location", *Wireless Personal Communications*, 94(1), pp 99-115, May (2017).
7. Mohamed A. El-Sayed and Tarek Abd-El Hafeez, "New Edge Detection Technique based on the Shannon Entropy in Gray Level Images", *Int. J. on Comput. Sci.*, 4, pp 186-191, (2008).
8. Mohamed A. El-Sayed, Sayed F. Bahgat and Abdel-Khalek, "Novel Approach of Edges Detection for Digital Images Based On Hybrid Types of Entropy", *Int. J. of Applied Mathematics and Information Science*, pp 1809-1817, (2013).

접수일: 2020년 1월 31일, 심사일: 2020년 3월 11일,
 게재확정일: 2020년 3월 18일