

## Study on OCR Enhancement of Homomorphic Filtering with Adaptive Gamma Value

Heeyeon Jo\*, Jeongwoo Lee\*, Hongrae Lee\*\*

\*Student, Division of Digital Healthcare, Yonsei University, Wonju, Korea

\*Student, Division of Digital Healthcare, Yonsei University, Wonju, Korea

\*\*Professor, Software Division, Yonsei University, Wonju, Korea

### [Abstract]

AI-OCR (Artificial Intelligence Optical Character Recognition) combines OCR technology with Artificial Intelligence to overcome limitations that required human intervention. To enhance the performance of AI-OCR, training on diverse data sets is essential. However, the recognition rate declines when image colors have similar brightness levels. To solve this issue, this study employs Homomorphic filtering as a preprocessing step to clearly differentiate color levels, thereby increasing text recognition rates. While Homomorphic filtering is ideal for text extraction because of its ability to adjust the high and low frequency components of an image separately using a gamma value, it has the downside of requiring manual adjustments to the gamma value. This research proposes a range for gamma threshold values based on tests involving image contrast, brightness, and entropy. Experimental results using the proposed range of gamma values in Homomorphic filtering suggest a high likelihood for effective AI-OCR performance.

▶ **Key words:** AI-OCR, Optical Character Recognition(OCR), Preprocessing, Homomorphic Filtering, Adaptive Threshold Range

### [요약]

AI-OCR은 광학 문자 인식(OCR) 기술과 Artificial intelligence(AI)의 결합으로 사람의 인식이 필요하던 OCR의 단점을 보완하는 기술 향상을 이뤄내고 있다. AI-OCR의 성능을 높이기 위해서는 다양한 학습데이터의 훈련이 필요하다. 하지만 이미지 색상이 비슷한 밝기를 가진 경우에는 인식률이 떨어지기 때문에, Homomorphic filtering(HF)을 이용한 전처리 과정으로 색상 차이를 분명하게 하여 텍스트 인식률을 높이게 된다. HF은 감마값을 이용해 이미지의 고주파와 저주파를 각각 조절한다는 점에서 텍스트 추출에 적합하지만 감마값의 조절이 수동적으로 이뤄지는 단점이 존재한다. 본 연구는 시험적 과정을 거쳐 이미지의 대비, 밝기 및 엔트로피를 근거하는 감마의 임계값 범위를 제안한다. 제안된 감마값 범위를 적용한 HF의 실험 결과는 효율적인 AI-OCR의 높은 등장 가능성을 시사한다.

▶ **주제어:** AI-OCR, 광학 문자 인식 (OCR), 전처리, 호모모픽 필터링, 적응적 임계 범위

• First Author: Heeyeon Jo, Co-Authors: Jeongwoo Lee, Corresponding Author: Hongrae Lee

\*Heeyeon Jo (7531hy@yonsei.ac.kr), Division of Digital Healthcare, Yonsei University

\*Jeongwoo Lee (jwlee87@yonsei.ac.kr), Division of Digital Healthcare, Yonsei University

\*\*Hongrae Lee (vim0614@gmail.com), Software Division, Yonsei University

• Received: 2023. 12. 04, Revised: 2024. 02. 06, Accepted: 2024. 02. 06.

## I. Introduction

4차 산업혁명의 확산과 함께 빅데이터 활용이 중요시되며 문서 전자화 시스템을 도입하는 산업들이 증가하고 있다[1]. 특히 인공지능의 발전으로 OCR 기술과 AI 결합한 AI-OCR 방법이 여러 산업 분야에서 사용되는 추세이다[2]. AI-OCR의 문자인식은 학습을 통해 만든 빅데이터로 학습된 결과물에 대한 목표 인식률이 보장될 때 더 이상 학습하지 않으며, 목표 인식률이 보장되지 않았을 때 재학습을 통해 인식률을 보장할 수 있는 기술이 적용된다[3]. OCR 정확도는 패턴인식에 딥러닝 기술을 접목하면서 높아졌으나 여전히 학습데이터 이미지 내의 조명 변화, 밝기와 대비의 변화와 같은 요인 따라 많은 영향을 받는다[4]. 기업 내 내부 보안 정책이나, 학습데이터를 만드는 하드웨어 교체가 어려워 입력 이미지의 개선이 힘든 경우에는 잡음이 포함된 학습데이터에서 잡음을 제거하는 전처리가 필요하다[5].

전처리는 학습데이터 이미지로부터 텍스트 영역을 컴퓨터가 쉽게 인식할 수 있도록 하는 이미지를 보정의 일환이다. 컴퓨터는 학습데이터 이미지 내에 비슷한 밝기를 가진 픽셀들을 하나의 덩어리처럼 인식하기 때문에 밝기의 차이를 확실히 하기 위하여 Homomorphic filtering(HF)을 사용한다. HF은 학습데이터 이미지 내 고주파와 저주파 성분에 사용자가 설정한 감마값을 곱해 불균형한 밝기와 대비를 조절한다. 이를 위해 이미지를 주파수 도메인으로 변환하여 작업을 수행하여 고주파와 저주파를 조정하는 HF이 자주 사용된다. [6]에서는 저조명 토마토 이미지의 색상 왜곡과 뚜렷하지 않은 엽지를 HF을 결합한 알고리즘을 통해 해결하고 [7]에서는 의료 및 비균일 조명 이미지 데이터셋의 이미지의 조명을 교정하기 위해 adaptive HF을 사용한다. [8]에서는 Si3N4 세라믹 베어링 볼의 표면 결함 개선을 위해 이미지의 고주파 성분을 강화하고, 저주파 성분을 압축하여 컨볼루션 노이즈와 혼합한 신호를 제거한다. [9]에선 저조명 상황에서의 피부와 유사한 다른 물체 이미지를 분리하기 위해 HF이 적용된다. 하지만 HF은 고주파와 저주파를 조정하기 위해 감마값을 사용자가 임의값으로 설정해야 하는 문제점이 있다. 이에 본 논문에서는 학습데이터 이미지의 휘도, 대비, 엔트로피의 값을 이용하여 이미지의 특성을 분석하고 이에 따른 감마값의 임계 범위를 제안한다.

본 논문의 구성은 다음과 같다. II장에선 이미지 내 빛 정보 제거에 사용되는 HF에 관련된 연구를 살펴본다. III장에서는 실험의 전체적인 순서도를 제시하며 감마의 임

계값 범위를 제안한다. IV장에서는 앞선 III장에서 제시한 방법의 실제 인식률을 가시적으로 보여주며 이를 토대로 V장에서 본 논문의 결론을 서술한다.

## II. Related works

이미지를 구성하는 픽셀은 주변광과 물체에 의해 반사되는 빛이라고 불리는 두 가지 빛의 요소에 의해 영향을 받는다. 이를 수식으로 나타내면 아래와 같다[10].

$$I(x,y) = L(x,y)R(x,y) \quad (1)$$

식 (1)에서 볼 수 있듯이 이미지의 픽셀은 주변광(L)과 반사율(R)이 곱해져 이루어진다[11]. 주변광(L)은 물체를 비추는 빛의 양이고 반사율(R)은 표면에서 반사되는 빛의 양을 나타낸다[12]. 물체에 포함된 빛의 양이 불균형할 경우 물체를 인식하는데 어려움이 생기게 된다. 이것을 해결하기 위해 주변광과 반사율을 조정하여 빛 정보를 조절한다[13]. 이미지에서 저주파 성분인 주변광은 이미지가 전체적으로 지나치게 밝은 경우에는 저주파 성분을 감쇠하고 어두운 경우에는 저주파 성분을 강화한다. 반대로 고주파 성분인 반사율은 경계나 텍스처를 약하게 하고 싶을 때 고주파 성분을 감쇠하고 강하게 하고 싶을 때 강화한다[14]. HF은 이미지를 주파수 도메인으로 변환하여 작업을 수행하여 고주파와 저주파를 조정하는 데 사용한다.

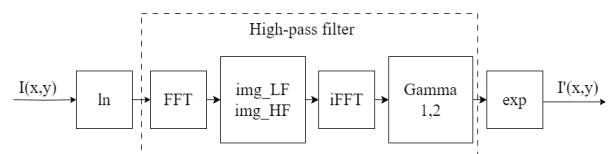


Fig. 1. Homomorphic filtering flowchart

Fig. 1은 HF의 과정을 간략하게 나타낸 순서도이다. 첫 번째로 HF은 이미지 내의 밝기 변화는 일반적으로 곱셈 연산으로 나타낸 이미지 정보  $I(x,y)$ 를 가산 성분으로 변환한다. 이를 위한 변환하는 수식은 아래와 같다.

$$\ln(I(x,y)) = \ln(L(x,y)) + \ln(R(x,y)) \quad (2)$$

식 (2)는 이미지를 로그 도메인으로 이동시키기 위해 식 (1)에 로그를 취하여 곱셈 성분을 가산 성분으로 변환하는 것을 나타낸다. 두 번째로 가산 성분으로 변환된 이미지에서 밝기 정보만을 분리하여 Fast Fourier Transform

(FFT)을 적용한다. 이를 위한 수식은 아래와 같다.

$$img_{FFT} = F\left\{\ln\left(1 + \frac{y}{255}\right)\right\} \quad (3)$$

식 (3)에서  $y$ 는 YUV 이미지에서 조명정보(Y)를 의미하고 변환된 이미지에서 분리한 밝기 정보에 푸리에 변환(F)을 적용한다. 세 번째로 FFT를 통해 주파수 영역으로 변환된 이미지를 고주파 성분과 저주파 성분을 분리하기 위해 대표적으로 Gaussian Mask 또는 Butterworth mask을 사용한다[15].

$$LPF = \exp\left(-\frac{(X-X_c)^2 + (Y-Y_c)^2}{2\sigma^2}\right) \quad (4)$$

식 (4)는 로그 도메인에서 생성한 저주파 통과 마스크로  $X, Y$ 는 Gaussian Mask 생성을 위한 그리드이며  $X_c, Y_c$ 는 그리드의 중심 좌표이다. 네 번째로, 이렇게 생성한 저주파 통과 마스크를 사용하여 만든 저주파 필터(LPF)와 고주파 필터(HPF = 1 - LPF)로 이미지 내의 두 성분을 분리한다.

$$img_{LF} = F^{-1}\{img_{FFT} \times LPF\} \quad (5)$$

$$img_{HF} = F^{-1}\{img_{FFT} \times HPF\} \quad (6)$$

식 (5)와 (6)은 식(3)에서 생성한 푸리에 변환된 이미지 ( $img_{FFT}$ )를 식 (4)에서 만든 저주파 필터(LPF)와 고주파 필터(HPF)를 곱한 후 역 푸리에 변환을 적용한다. 이렇게 이미지 밝기 정보의 저주파와 고주파 성분을 분리해 낸다. 다섯 번째로 이미지 내의 주변광과 반사율 두 가지 요소를 조절하기 위해 감마값을 고주파와 저주파 성분에 곱해준다. 이를 위한 수식은 아래와 같다.

$$img = \gamma_1 \times img_{LF} + \gamma_2 \times img_{HF} \quad (7)$$

식 (7)에  $img_{LF}$  은 식(5)의 필터를 이용한 이미지에서 분리한 저주파 성분을 나타내며,  $img_{HF}$  은 식(6)의 필터를 이용한 이미지에서 분리한 고주파 성분을 나타낸다.  $\gamma_1$  은 저주파 성분을 곱해지는 감마값이며,  $\gamma_2$  은 고주파 성분에 곱해지는 감마값이다. 불균일한 조도로 인한 낮은 목표 검출률을 줄이기 위해 적합한 감마값  $\gamma_1$  과  $\gamma_2$  을 설정하여 분리한 저주파와 고주파 성분에 곱하면 이미지의 밝기 성

분이 개선된다[16]. 식 (3)에서부터 식(7)까지의 단계를 거치면, 원본 이미지에서 고주파 성분이 추출되고, 이를 다시 원본 이미지에 더해주는 과정인 High-pass filter를 실행하게 된다. 마지막으로 High-pass filter를 통과한 이미지를 다시 공간 도메인으로 변환한다.

$$img_{exp} = \exp(img) - 1 \quad (8)$$

식 (8)은 주파수 도메인 이미지를 지수함수인  $\exp(x)$  함수로 역변환하여 공간 도메인으로 가져오는 수식이다. 이렇게 얻은 HF의 최종 이미지는 식 (7)의 감마값 설정에 따라 다른 결과를 갖는다. 감마값 설정 과정은 정형화된 가이드라인이 없고 수동으로 이뤄지기 때문에 일부 이미지는 HF의 감마값 설정 과정을 생략하기도 한다. 그러나 이렇게 얻은 이미지는 반사율과 주변광의 조절이 각각 가능한 HF의 특수성을 이용하지 않으며 저주파와 고주파의 세밀한 조절이 이뤄지지 않은 결과물로 나타난다.

본 논문에서는 학습데이터 이미지의 밝기를 조정하기 위하여 감마값을 수동 조절해야 하는 문제점을 해결하고자 학습데이터 이미지의 특성을 이용한 감마값의 범위를 설정 방법을 제안한다. 학습데이터 이미지의 인식률에 영향을 주는 특성으로는 휘도(Luminance), 대비(Contrast), 엔트로피(Entropy)가 있다. 휘도는 이미지의 밝기 수준을 의미하며, 휘도를 통해 이미지의 밝기를 분석할 수 있다[ 17]. 대비는 이미지의 밝고 어두움의 차이를 나타내며 대비가 높을수록 대상 정보와 배경을 명확하게 구분하는 데 도움을 준다[18]. 엔트로피는 정보의 무질서도를 나타내며, 엔트로피가 높을수록 이미지 내에 더 많은 다양한 정보를 포함하고 있다는 것을 의미한다[19]. 이러한 학습데이터 이미지의 특성들은 피어슨 상관분석법[20]을 사용해 상관관계를 분석하고 분석 결과는 감마값의 범위를 설정하기 위해 사용한다. 이와 관련된 내용은 III 장에서 자세히 다루도록 한다.

### III. The Proposed Scheme

본 장에서는 제안한 Homomorphic Filtering (HF)의 최적 감마값 범위를 설정하기 위해 실험 과정을 추출과 분석의 단계로 나누어 진행한다. 추출 단계에서는 HF 감마값들을 0.1부터 10까지 0.1 간격으로 텍스트 인식이 되지 않는 학습데이터 이미지에 적용하여 변환된 이미지의 휘도, 대비, 엔트로피를 추출한다. 추출된 데이터 간의 상관관계 및 휘도, 대비, 엔트로피의 변화량 분석을 통해 감마

값 설정에 대한 보다 정밀한 기준을 제시한다.

Table 1. Experimental Setup

Component	Specification
CPU	AMD Ryzen 7 5800U with Radeon Graphics 1.90 GHz
RAM	16.0GB
Operating system	Windows 11
Programming language	Python ver. 3.11.5 (with Visual Studio Code)

Table. 1은 본 논문의 실험 환경을 나타낸다. 학습 데이터 이미지는 텍스트 인식 여부를 tesseract OCR[21]을 사용하여 인식이 되지 않는 텍스트 이미지를 선별하였다.

Table 2. Luminance, contrast, entropy and size of each original image

No.	Size	luminance	contrast	entropy
(a)	505×335	61.9113	1.0000	6.8011
(b)	1024×602	44.8291	0.9318	4.3871
(c)	1280×720	118.6573	0.9318	7.2385
(d)	1200×680	51.8146	0.8731	3.2244
(e)	1040×693	36.7535	1.0000	4.9820
(f)	1280×720	45.2326	1.0000	6.1163
(g)	1280×720	40.0266	1.0000	3.9773
(h)	1280×720	104.7465	0.6416	6.8542
(i)	1280×720	50.7848	1.0000	6.1320
(j)	1280×720	39.5337	1.0000	5.7664
(k)	750×500	53.2680	0.9844	6.8993
(l)	750×500	56.5099	0.9844	6.9797



Fig. 2. Test images before Homomorphic filtering

Table. 2는 12장의 학습데이터 이미지들의 크기, 휘도, 대비와 엔트로피의 값을 나타내며 Fig. 2는 실험에 사용된 이미지를 보여준다. 실험 이미지는 HF의 특성을 고려하여 단색의 배경에 텍스트 주변이 빛의 산란으로 인해 인식을 방해받는 이미지를 선별했다.

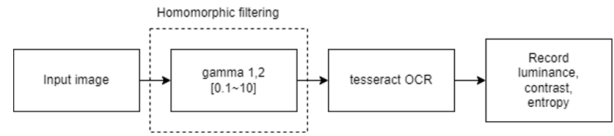


Fig. 3. Feature extraction process after Homomorphic filtering.

Fig. 3은 추출 단계의 Flowchart를 보여준다. 첫 번째로 실험은 텍스트 인식이 되지 않는 학습데이터 이미지를 입력하고, 두 번째로 HF의 감마값  $\gamma_1$ 와  $\gamma_2$ 을 0.1에서 10까지 0.1씩 증가시키며 HF를 실행한다. 이때 감마값 범위는 10을 넘어선 큰 값은 이미지에 과도한 잡음을 줄 수 있고, 0 이하의 값은 주요한 정보를 반전시키거나 왜곡해 인식을 방해할 수 있다. 이에 최소 0에서 최대 10까지 실험에 사용할 감마값 범위를 고정한다. 또한 큰 단위는 최적의 감마값에 도달하기 전 중간값 범위를 뛰어넘는 경우가 발생하므로 단위를 0.1로 설정해 점진적인 변화를 준다.

이 과정을 통해  $\gamma_1$ 와  $\gamma_2$ 의 변화에 따라 각 학습데이터 이미지 당 10,000장의 이미지가 생성되며 총 120,000장의 결과 이미지를 얻게 된다. 세 번째로 120,000장의 결과 이미지를 tesseract OCR을 통해 텍스트 인식 여부를 확인한다. Fig. 4는 두 번째 과정을 통해 얻은 120,000장의 결과 이미지 중 텍스트 인식이 가능한 이미지로 변환이 성공한 결과 이미지의 일부를 보여준다. 마지막으로 분석을 위해 120,000장의 결과 이미지의 휘도, 대비, 엔트로피의 값 및 HF에 사용된 감마값  $\gamma_1$ 와  $\gamma_2$ 을 기록한다.

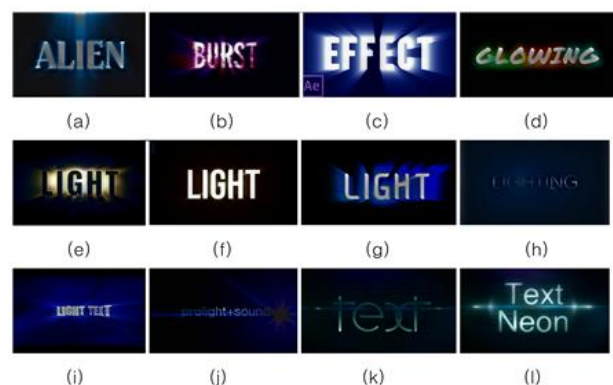


Fig. 4. Test images after Homomorphic filtering

분석 단계에서는 추출된 120,000장의 결과 이미지의 데이터를 바탕으로 두 가지 주요 분석을 수행한다. 첫 번째는 피어슨 상관분석법을 적용하여 감마값과 이미지의 휘도, 대비, 엔트로피 간의 상관관계를 분석한다. 두 번째는 각각의 감마값 변화에 따른 원본 및 필터링된 이미지의 휘도, 대비, 엔트로피의 변화량과 바뀐 값을 분석한다.

분석의 첫 번째 단계는 결과 이미지들의 휘도, 대비, 엔트로피와 감마값  $\gamma_1$ ,  $\gamma_2$ 을 피어슨 상관 분석법으로 분석하여 결과 이미지의 특성과 감마값의 선형성을 계산한다. 피어슨 상관분석은 변수 간의 선형적인 상관관계를 측정하는 유용한 방법으로, 상관계수는 -1부터 1까지의 범위에서 강도와 방향성을 나타낸다. 상관계수가 양수일 경우 두 변수가 양의 선형 관계에 있음을 나타내며, 상관계수가 음수일 경우 음의 선형 관계에 있음을 나타낸다. 또한 상관계수의 절댓값이 클수록 두 변수 간의 선형 관계가 강함을 의미한다. 만약 휘도와 감마 값 간에 강한 음의 상관관계가 있다면, 휘도를 높이기 위해서는 감마 값을 낮추면 된다. 이와 마찬가지로 대비와 엔트로피에 대한 감마값과의 상관관계를 측정하면 감마값을 어떻게 조절할지에 대한 정보를 얻을 수 있다.

피어슨 상관계수는 두 변수의 공분산을 각각의 표준편차로 나눈 값으로 수식으로 나타내면 아래와 같다.

$$r_{XY} = \frac{\sum_i^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_i^n (X_i - \bar{X})^2} \sqrt{\sum_i^n (Y_i - \bar{Y})^2}} \quad (9)$$

식 (9)에서  $r_{XY}$ 는 피어슨 상관계수를 나타내며  $X$ 와  $Y$ 에는 상관계수를 계산하게 될 값들이 대입된다. 예를 들어  $X$ 에 휘도 값이 대입되고  $Y$ 에  $\gamma_1$ 이 대입되면 휘도와  $\gamma_1$ 의 피어슨 상관계수를 구하게 된다. 본 단계에서는 결과 이미지의 휘도, 대비, 엔트로피와 감마값  $\gamma_1$ ,  $\gamma_2$ 의 상관계수를 구하기 위해 변수  $X$ 에는 결과 이미지의 특성인 휘도, 대비, 엔트로피가 대입되고  $Y$ 에는 감마값  $\gamma_1$ ,  $\gamma_2$ 가 대입된다.  $\bar{X}$ 와  $\bar{Y}$ 는  $X$ 와  $Y$ 의 평균값을 나타내며, 데이터 포인트  $i$ 는 감마값을 조절한 후의 측정된 휘도, 대비, 엔트로피의 값을 의미한다.

Table 3은 120,000장의 결과 이미지의 특성과 감마값  $\gamma_1$ ,  $\gamma_2$  간의 피어슨 상관계수를 나타낸다.

Table 3. Pearson correlation coefficient between gamma value and image characteristics

Characteristic/ Gamma	luminance	contrast	entropy
Gamma1( $\gamma_1$ )	-0.563	0.054	-0.445
Gamma2( $\gamma_2$ )	-0.827	0.096	-0.677

Table 3을 보면  $\gamma_1$ 의 경우 휘도와 -0.563의 음의 상관계수를, 대비와는 0.054의 매우 약한 양의 상관계수를, 엔트로피와는 -0.445의 음의 상관계수를 가진다. 이는  $\gamma_1$ 과 휘도, 엔트로피가 중간 정도의 음의 선형성을 갖으며 대비와는 비선형적인 관계임을 알 수 있다.  $\gamma_2$ 의 경우 휘도와 -0.827의 매우 강한 음의 상관계수를, 대비와는 0.096의 약한 양의 상관계수를, 엔트로피와는 -0.677의 음의 상관계수를 가진다. 이는 검은 배경에 고주파 성분인 빛의 산란이 포함된 학습데이터의 특성상  $\gamma_1$ 에 비해 고주파 정보를 제어하는  $\gamma_2$ 가 휘도, 엔트로피와 유의미한 음의 선형성을 가지고 있는 것으로 파악된다. 이를 통해 감마값들은 휘도와 엔트로피는 0.5 정도로 유의미한 선형성을 갖는 반면, 대비와는 두 감마값 모두 0.1 이하로 상대적으로 낮은 선형성을 갖는 것으로 정리할 수 있다.

분석의 두 번째 단계는 감마값 변화에 따른 원본 및 필터링된 이미지의 휘도, 대비, 엔트로피의 변화량과 바뀐 값을 분석한다. Fig. 5.는 원본 이미지와 120,000장의 결과 이미지의 감마값 변화에 따른 휘도, 대비, 엔트로피의 값과 변화량을 나타낸 그래프이다.

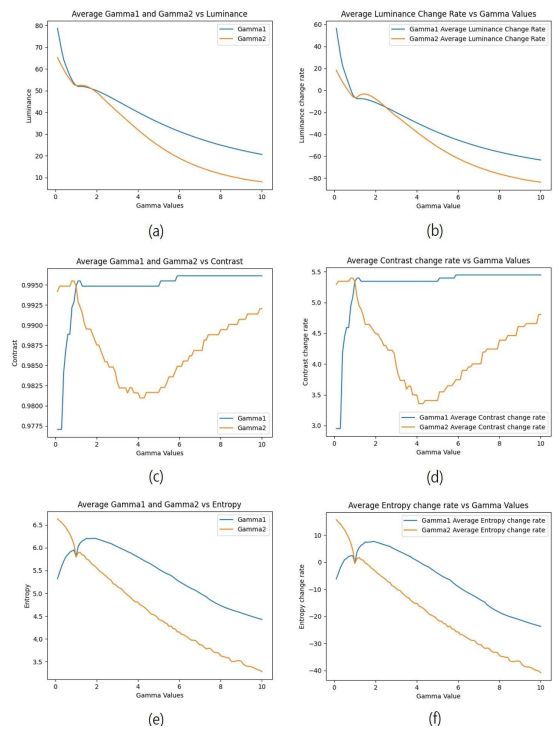


Fig. 5. Graph of values and changes in luminance, contrast, and entropy according to gamma 1,2

Fig. 5의 (a),(b)는 감마값 변화에 따른 결과 이미지의 평균 휘도의 값과 휘도의 변화량을 나타낸다. 감마값은 휘도

와 음의 선형성을 가지므로 0~1에서는 휘도가 증가되는 경향을 보이지만, 1부터 값이 증가함에 따라 휘도가 감소하게 된다. 또한  $\gamma_2$ 가  $\gamma_1$ 에 비해 큰 상관계수를 가지므로 감마값이 증가할 때 변화량이 더 커짐을 알 수 있다. Fig. 5의 (c),(d)는 감마값 변화에 따른 결과 이미지의 평균 대비와 대비의 변화량을 나타낸다. 감마값은 대비와 낮은 선형성을 가지므로 그래프의 모양에서는 크게 변한 것으로 보이지만 변화량은 크지 않음을 알 수 있다.  $\gamma_1$ 의 경우 0~1 범위에서 휘도가 증가함에 따라 대비의 비율이 급격히 증가했으며, 1부터 값이 증가함에 따라 대비의 변화율이 일정하게 유지된다.  $\gamma_2$ 의 경우 0~1 범위에서 대비의 비율이 일정하게 유지되다가 1~5까지 변화율이 소폭 감소하고 5~10까지 변화율이 점차 증가함을 알 수 있다. 이는 위 상관계수에서 감마값이 대비와 상관성이 낮음을 나타낸 결과와 일치함을 알 수 있다. 첫 번째 분석에서 사용한 피어슨 상관계수와 마찬가지로 그래프에서도 휘도, 엔트로피에 비해 변화율의 폭이 작고 따라서 선형성이 낮다는 첫 번째 분석에 부합한다. 이를 통해 이미지의 대비 변화에는  $\gamma_1$ ,  $\gamma_2$ 의 범위를 적절히 조절한다면 높은 대비의 유지가 가능함을 알 수 있다. Fig. 5의 (e),(f)는 감마값 변화에 따른 결과 이미지의 평균 엔트로피와 엔트로피의 변화량을 나타낸다. 엔트로피는 저주파 성분이 많을수록 낮아지고 고주파 성분이 많을수록 높아지게 된다. HF에서  $\gamma_1$ 은 저주파에 가중치를  $\gamma_2$ 은 고주파에 가중치를 주게 되는데 감마값이 2 이상일 때 휘도의 감소는 고주파 성분의 손실을 초래한다. 이는 세부적인 텍스처와 가장자리 정보의 감소로 이어져 감마값이 2 이상일 때의 엔트로피 값을 하락시킨다. 따라서, 엔트로피가 감마값이 2 이상일 때 감소하는 그래프를 그린다.

Table 4. Gamma's Threshold Range Setting Guidelines

luminance	contrast	Gamma1	Gamma2
100~	-	4~6	
50~100	-	3~5	
~50	-	1~10	
-	keep 1.0		1~3, 8~10

Table. 4.는 Fig. 5.의 그래프를 기반으로 설정한 Gamma 값 범위 표를 보여준다.

휘도의 경우, HF를 통해 실험 이미지의 빛 정보가 조절되며, 휘도의 경우 학습데이터가 검은 배경을 가진 만큼 전체적으로 휘도 값이 낮다. 따라서 Gamma 값 범위 표 또한 학습데이터의 특성에 기반한 범위를 설정하였다. 텍

스트 추출에 성공한 대부분 결과 이미지에서 휘도가 원본에 비해 절반 이하로 감소하는 경향을 보인다. 특히, 휘도가 높은 이미지의 경우  $\gamma_1$ 을 4~6 범위로 설정하여 휘도를 절반 수준으로 줄이는 것이 효과적이었다. 반면, 휘도가 50 이상 100 이하의 경우  $\gamma_1$ 을 3~5 범위로 설정하여 휘도를 20에서 50 정도 줄이는 것이 적절했으며, 휘도가 50 이하의 경우에는 휘도를 낮출 필요가 없으므로  $\gamma_1$  값을 1~10 범위로 넓게 잡아 대비를 높이는 데 중점을 두었다. 대비는 원본 이미지가 HF 후에도 텍스트와 배경을 뚜렷하게 구분해주는 높은 대비 값을 유지하는 것이 텍스트 인식에 긍정적인 영향을 미쳤다. 높은 대비를 위해 감마값의 조정에서 1.0에 근접하거나, 이미 1.0이었던 실험 이미지에서는 상대적으로 변함없는 값을 유지하는 것을 목표로 하였다. 높은 대비를 위해  $\gamma_2$ 을 1~3 및 8~10 범위로 설정하여 대비를 최적화할 수 있었다.

엔트로피의 경우, 고주파에 미치는 휘도의 영향으로 1 이상의 감마값에서 감소하는 형태의 그래프를 그린다. 이에 따라서 가이드라인에서 제시하는 휘도 범위의 영향으로 엔트로피 조절이 가능하다.

#### IV. Experiments

본 장에서는 3장에서 제안한 Homomorphic Filtering (HF)의 최적 감마값 범위를 사용하여 텍스트 인식이 되지 않는 실험 이미지에 적용하여 실험을 진행한다.

Table 5. Luminance, contrast, entropy and size of each experiments image

No.	Size	luminance	contrast	entropy
(a)	480X360	27.8658	1.0	5.2417
(b)	736X414	14.8178	1.0	4.9521
(c)	474X271	31.5766	1.0	4.7339
(d)	570X358	13.2290	1.0	3.8432
(e)	493X275	10.3101	1.0	1.6766

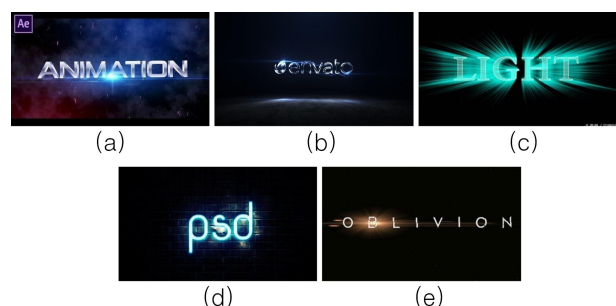


Fig. 6. Experiments images before HF filtering

Table. 5.는 5장의 실험 데이터 이미지들의 크기, 휘도, 대비와 엔트로피의 값을 나타내며 Fig. 6은 실험에 사용된 이미지를 보여준다. 본 연구는 Table. 4.의 가이드라인에 따라 다섯 개의 이미지 (a), (b), (c), (d), (e)의 감마값 범위를 설정한다. 다섯 개 이미지 모두 휘도가 50 이하인 범위에 속하므로  $\gamma_1$ 의 값을 1에서 10 사이로 넓게 설정하여 대비를 강화하는 데 중점을 두었다. 또한 모든 이미지에서 대비가 1.0으로 일정하므로,  $\gamma_2$  값을 1에서 3 그리고 8에서 10 사이로 조정하여 대비를 최적화하였다.

Table 6. Image result after setting the gamma value range according to the guidelines

No.	All gamma values			Application of the suggested gamma range		
	All image	Recognized image	Recognition rate (%)	All image	Recognized image	Recognition rate (%)
(a)	10000	5707	57.07	3822	2455	64.23
(b)	10000	4608	46.08	3822	1187	31.06
(c)	10000	1708	17.08	3822	672	17.58
(d)	10000	1645	16.45	3822	201	5.26
(e)	10000	4731	47.31	3822	1309	34.25

Table. 6.는 본 연구에서 수행된 감마값 가이드라인 적용의 효과를 분석한 결과이다. 모든 원본 이미지에서 tesseractOCR이 실패했으나 HF 필터링을 적용한 이후 텍스트 인식에 성공하는 모습을 보였다. 가이드라인을 통해 감마값 범위를 적용한 후에도 (b)를 제외하고 인식률이 15% 이상으로 감소하지 않는 일정한 수준을 유지함을 알 수 있었다.

위 결과를 통해 HF의 감마값 가이드라인을 이용한 전처리가 텍스트 인식률을 개선함과 동시에 HF만을 사용하는 전처리 방법보다 효율적임을 알 수 있었다.

## V. Conclusion and Future Work

본 논문은 AI-OCR의 전처리를 위한 학습데이터 이미지의 밝기를 조정하기 위하여 감마값을 수동 조절해야 하는 문제점을 해결하고자 학습데이터 이미지의 특성을 이용한 감마값의 범위 설정 방법을 제안하였다. 이를 위해 Homomorphic filtering을 통해 텍스트 인식이 가능하게 된 이미지의 휘도, 대비, 엔트로피와 감마값의 상관관계

및 변화량을 분석하여 감마값 범위를 설정하였다. 인식되지 않는 학습데이터 이미지의 특성을 이용하여 제시한 감마값 범위를 적용한 이후 텍스트 인식에 성공함을 확인할 수 있다. 또한 모든 감마값의 범위를 적용하여 인식된 이미지의 최대 43%를 최소 12%를 추출됨을 알 수 있었다.

이 결과를 통해 이미지의 특성을 이용한 감마값의 범위를 설정은 텍스트 인식률을 개선하는 데 있어 유효한 전략이 될 수 있으며, 이는 더 넓은 범위의 이미지 처리 및 텍스트 인식 문제에 적용될 수 있다. 본 연구는 특정한 특성을 가진 이미지를 사용한다. 감마값 조정이 모든 유형의 이미지에 대해 동일한 결과를 제공하지 않는다는 점을 인지하고, 모든 이미지에 대한 인식율을 높이기 위한 추가적인 실험과 분석을 통한 일반화가 추후 연구로 필요하다.

## ACKNOWLEDGEMENT

This research was supported by the MISP(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation) (2019-0-01219).

## REFERENCES

- [1] Kim, Won Jun, Lee, Sang Kon and Pyo, Sung Kuk, "A Study on the Prediction for the OCR Technology Development Trajectory based on the Patent and Article Information.", JITS, 6, 80, 39-51, 2022, DOI: 10.9716/KITS.2022.21.6.039
- [2] Kim, Pil Ho, "A Study on the Improvement of OCR Hangul Recognition Rate for Efficient Search Utilization of Records.", Joogbu University: Department of Records Management, 2022.
- [3] Sejin Ahn, Hyunho Hwang and Yim, Jin Hee "A Case Study on the Application of AI-OCR for Data Transformation of Paper Records", JKOSIM, 3, 125, pp. 165-193, 2022, DOI: 10.3743/KOSIM.2022.39.3.165
- [4] Young Jae Park, Dae Won Lee, "OCR system through automatic selection of deep learning-based image processing techniques.", Proceedings of KIIT Conference,(),336-339, 2021.
- [5] Kim, Da Young, "Artificial Intelligence-based document image preprocessing method for improving OCR accuracy on scanned documents." Graduate School of Information and Telecommunications, Konkuk University, 2022. Seoul.
- [6] J. He, J. Liu, H. Chen, Y. Zhang and W. Wu, "FMR Low-light

- Tomato Image Enhancement Algorithm Based on Optimized Homomorphic Filter.", International Conference on Computer and Communications (ICCC), 7th, pp. 725-730, 2021, DOI: 10.1109/ICCC54389.2021.9674365.
- [7] P. Venkatappareddy and B. Lall, "A Novel Thresholding Methodology Using WSI EMD and Adaptive Homomorphic Filter.", IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 67, no. 3, pp. 590-594, March 2020, DOI: 10.1109/TCSII.2019.2918444.
- [8] Liao D., Cui Z., Li J., Li W., Wang W., "Surface defect detection of Si3N4 ceramic bearing ball based on improved homomorphic filter-Gaussian filter coupling algorithm.", AIP Advances, 12, no. 025325, 2022., DOI: 10.1063/5.0082702
- [9] T. B. Adji, F. Rahadian, H. A. Nugroho and A. G. Persada, "Negative content filtering based on skin texture, homomorphic filter and localizations." 2014 International Conference on Electrical Engineering and Computer Science (ICEECS), Kuta, Bali, Indonesia, pp. 1-6, 2014, DOI: 10.1109/ICEECS.2014.7045242.
- [10] Ebrahim Bidokh, Hamid Hassanpour, "Enhancing Wireless Capsule Endoscopy images from intense illumination specular reflections using the homomorphic filter", Biomedical Signal Processing and Control, Volume 84, 2023, 104723, ISSN 1746-8094.
- [11] Lina Xu, Yujuan Si, Saibiao Jiang, Ying Sun, Homayoun Ebrahimiyan, "Medical image fusion using a modified shark smell optimization algorithm and hybrid wavelet-homomorphic filter", Biomedical Signal Processing and Control, Volume 59, 2020, 101885, ISSN 1746-8094,
- [12] S. Alotaibi, N. Alharbi and H. Kurdi, "Face recognition under varying illumination based on homomorphic filter and local binary patterns," 2017 Intelligent Systems Conference (IntelliSys), London, UK, pp. 585-591, 2017, DOI: 10.1109/IntelliSys.2017.8324354.
- [13] M. Petrou and C. Petrou, "Image Processing: The Fundamentals: Second Edition." Wiley, 2011.
- [14] P. Hill, H. Bhaskar, M. Al-Mualla and D. Bull, "Improved illumination invariant homomorphic filtering using the dual tree complex wavelet transform", 2016 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), 2016.
- [15] Dogra, A., & Bhalla, P., "Image sharpening by gaussian and butterworth high pass filter." Biomedical & Pharmacology Journal, 7(2), 707-713, 2014, DOI: <https://doi.org/10.13005/bpj/545>.
- [16] Jin Chen Lin, Lin Jin Chen, Yang Junjie, Yu Ting and Geng Chuanping, "Image Dehazing Algorithm Based On Improved Guided Filtering[J]", IOP Conference Series: Earth and Environmental Science, vol. 571, no. 1, 2020.
- [17] Mahani, Z., Zahid, J., Saoud, S., El Rhabi, M., Hakim, A. "Text Enhancement by PDE's Based Methods." In: Elmoataz, A., Mammas, D., Lezoray, O., Nouboud, F., Aboutajdine, D. (eds) Image and Signal Processing. ICISP 2012. Lecture Notes in Computer Science, vol 7340. Springer, Berlin, Heidelberg. DOI: 10.1007/978-3-642-31254-0\_8
- [18] Christian Wolf, Jean-Michel Jolion, "Extraction and Recognition of Artificial Text in Multimedia Documents.", Pattern Analysis and Applications, 4, 6, pp.309-326, 2004.
- [19] Yanglem Loijing Khomba Khuman, H. Mamata Devi, N. Ajith Singh, "Entropy-based skew detection and correction for printed Meitei/Meetei script OCR system", Materials Today: Proceedings, Volume 37, Part 2, 2021, p.p. 2666-2669, ISSN 2214-7853, DOI: 8443/10.1016/j.matpr.2020.08.522.
- [20] Sang-Il Lee, Daeheon Cho, Minpa Lee, "Spatializing the Pearson's Correlation Coefficient : An Experimental Comparison of Three Relevant Techniques.", Journal of the Korean Geographical Society, 53(5), 761-776.
- [21] Tesseract-ocr, tesseract, <https://github.com/tesseract-ocr/tesseract>

## Authors



Heeyeon Jo entered the Division of Digital Healthcare at Yonsei University, Wonju, Korea, in 2021. She is currently a student in the Division of Digital Healthcare, Yonsei University.

She is interested in data analysis, data science and image processing.



Jeongwoo Lee entered the Division of Digital Healthcare at Yonsei University, Wonju, Korea, in 2021. She is currently a student in the Division of Digital Healthcare, Yonsei University.

She is interested in artificial intelligence, biosignal processing, and image processing.



Hongrae Lee received the B.S., M.S. and Ph.D. degrees in Computer Science from Yonsei University, Korea, in 2010, 2012 and 2019, respectively. Dr. Lee joined the faculty of the Division of Computer and

Telecommunications Engineering at Yonsei University, Wonju, Korea, in 2020. He is currently an Assistant Professor in the Software Division, Yonsei University. He is interested in image processing, image communication technology and computer vision system.