

카메라 식별 가능 사진을 위한 코드삽재 및 검출

Code Embedding and Detection for Camera Identifiable Photos

배준현¹ · 홍원기² · 고재필^{1*}

¹금오공과대학교 컴퓨터공학과

²대구대학교 멀티미디어공학전공

Joon-Hyeon Bae¹ · Won-Kee Hong² · Jaepil Ko^{1*}

¹Dept. of Computer Engineering, Kumoh National Institute of Technology, Gumi, Gyeongsangbuk-do, 39177, Korea

²Dept. of Multimedia Engineering, Daegu University, Gyeongsan, Gyeongsangbuk-do, 712714, Korea

[요 약]

최근 스마트폰 카메라로 QR (quick response) 코드 사진을 촬영하여 추가 정보를 얻는 응용이 널리 사용되고 있다. QR 코드는 대용량 정보를 탑재할 수 있으나, 이를 인식하기 위해서는 반드시 코드의 네 모서리가 모두 포함되도록 촬영해야 한다. 본 논문에서는 이러한 제약을 극복할 수 있도록 디지털 영상에 코드 비트열을 삽재하고 검출하는 기법을 제안한다. 제안하는 기법은 프린트-캡처 과정에서 발생하는 심한 영상 훼손에도 불구하고 삽재된 코드가 손실되지 않도록 주파수 도메인에 코드를 삽재한다. 기존의 주파수 도메인 삽재 방법은 삽재 비트 수가 1~2비트 수준이나 본 논문에서는 16비트를 삽재할 수 있도록 설계하였다. 특히, 코드 비트열 외에 정렬 비트를 추가로 삽재하여 복호화 과정을 용이하도록 하였다. 연구실에서 수집한 영상에 대한 실험에서 임의 회전 하에서 일부만 촬영된 영상에 대해서도 평균 인식률 95.2%를 달성하였다. 이것은 평균 오류 비트의 수로 따지면 0.76개에 불과하다.

[Abstract]

Recently, applications that acquire additional information by taking a quick response (QR) code photo with a smart-phone camera is widely used. The QR codes can contain a lot of information, but in order to recognize a QR code, we need to take a photo so that all four corners of the code are covered. To overcome this limitation, we propose a method to embed a sequence of code bits to a digital image and detect from the captured image of its photos. The proposed scheme embed the code in the frequency domain so that the code is not lost despite severe image degradation issued during the print-camera process. The conventional frequency-domain approaches usually embed only 1~2 bits, however we designed to be able to contain 16 bits. In particular, we additionally embed alignment bits as well as code bits to facilitate the decoding process. In our experiments, we achieved an average accuracy of 95.2% for images partially captured under random rotation. This corresponds to only 0.76 bits in terms of the average number of error bits.

Key word : Image code, Code embedding, Watermarking, Steganography, Invisible code.

<https://doi.org/10.12673/jant.2021.25.2.156>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 16 March 2021; Revised 12 April 2021

Accepted (Publication) 28 April 2021(30 April 2021)

*Corresponding Author; Jaepil Ko

Tel: +82-54-478-7519

E-mail: nonezero@kumoh.ac.kr

1. 서론

최근 실생활에서 QR 코드 활용 사례를 쉽게 접할 수 있다. 전자출입명부, 송금 등이 대표적 사례이다. QR 코드는 불편한 스마트폰 입력을 카메라 촬영으로 대신할 수 있어, 그 활용 사례는 늘어날 것으로 예상된다. QR 코드는 대용량 정보를 탑재할 수 있으나, QR 코드를 인식하기 위해서는 반드시 코드의 네모서리가 모두 포함되도록 촬영해야 하고 심미성도 떨어지는 단점이 있다. 이에 대한 대안이 디지털워터 마킹 기술이다. 이것은 영상에 정보를 숨기고 특정키를 적용하여 정보탐재 여부를 판정하는 저작권 관리 목적으로 활발하게 개발되었으나, 2000년대 후반 개인용 모바일 장비에 탑재된 카메라로 촬영한 인쇄물을 식별하여 상품 인증 등의 용도로도 활용되기 시작했다[1]. 디지털워터마킹 기술의 장점은 인쇄하는 영상에 코드를 탑재하므로 별도로 인위적인 코드 패턴이 필요하지 않다는 점이다. 영상 자체만 사용하는 또 다른 방법은 영상을 인식하는 것이다. 최근 딥러닝 기술의 발달로 영상 인식 기술이 획기적으로 발전되었으나, 이 방법으로는 인쇄된 사진을 유일하게 구분하기 어렵다. 최근 코드 임베딩 목적의 딥러닝 모델이 발표되었으나, QR 코드와 마찬가지로 영상의 윤곽이 모두 촬영되어야 하는 제약이 있다[2].

워터 마킹 기술을 이용한 프린트-캡처 응용은 기술적으로 해결해야 할 요소가 많다. 디지털 영상을 프린터로 인쇄하는 과정에서 디지털에서 아날로그로의 변환이 발생하고, 인쇄물을 카메라로 촬영하는 과정에서 다시 아날로그가 디지털로 변환되는 과정을 거친다. 이 두 과정에서 영상 훼손이 심하므로 영상에 탑재된 코드 정보가 소실되기 쉽다. 따라서 프린트-캡처 응용에서는 영상 왜곡에 강인한 주파수 도메인 코드를 탑재 방식이 일반적이다[2].

주파수 도메인에 코드를 탑재하면 인쇄물의 일부만 촬영해도 코드를 검출할 수 있는 장점이 있으나, 탑재할 수 있는 정보의 양은 QR 코드에 비해 적을 수밖에 없다. 물론, QR 코드와 같이 영상의 모서리에 위치 구분을 위한 패턴을 추가하거나 인쇄 사진의 윤곽이 모두 촬영되어야 하는 조건에서, 원본 영상을 미리 정해진 방식으로 여러 구역으로 나누어 각각의 영역에 별도의 코드를 탑재함으로써 코드의 용량을 증가시킬 수는 있다. 그러나 최근 모바일 장치는 인터넷에 상시 연결되어 있으므로 콘텐츠 구분을 위한 키(key) 정도만 탑재하면 되므로, 영상 자체에 직접 일반 메시지와 같은 대용량 코드를 탑재할 필요성은 줄어들었다.

본 논문에서는 워터 마킹 기술 기반 프린트-캡처 응용을 위한 기법에 주목하였다. 관련 연구로 Dajun He와 Qibin Sun은 영상 훼손에 매우 강인한 코드 탑재 전략을 제시하였다[3]. 그러나 탑재할 수 있는 코드는 1비트에 불과하다. 더 많은 코드를 탑재하기 위해, [4]는 영상을 가로 및 세로 N 등분하고 각 영역에 사인 함수를 이용한 사선 패턴을 생성하여 원본 영상과 합성하였다. 사선 패턴은 크게 두 방향을 가지도록 생성하였으므로, 1비트를 표현할 수 있다. Pramila, A 등은 유사 무작위 패턴 영상을 회전시켜 얻은 영상을 원본 영상과 합성하였다[5]. 유사 무작위 패턴은 16등분 하여 사용하므로 하나의 영역은 4비트를 표현할 수 있다.

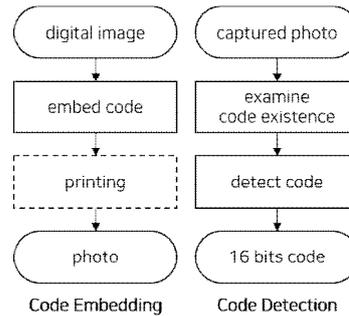


그림 1. 프로세스 구조

Fig. 1. Process architecture.

두 방법 모두 공간 도메인에 코드를 탑재하므로 영상의 왜곡에 취약하다. Ante Poljicak 등은 주파수 도메인에 코드를 탑재하는 방법으로 중간 주파수대에 환형으로 구역을 미리 정하고 이 구역을 등 간격으로 나누어 매그니튜드(magnitude) 계수 값을 크게 하거나 작게 하여 정보를 표현한다[6]. 이 방법 역시 코드 탑재 여부만 판정하고 영상의 원본 크기를 알고 있어야 한다는 제약이 있다. 우리는 이 논문에서 동기를 얻어 16비트 용량의 코드를 탑재하고 검출하는 방법을 제안한다. 본 논문의 구성은 다음과 같다. 2장에서 제안 방법론을 설명하고, 3장에서 실험 방법과 결과를 제시하며, 마지막 4장에서 결론을 내린다.

II. 환형 코드탐재 및 검출

2-1 개요

본 논문은 주어진 디지털 영상에 사용자가 정의한 16비트 코드를 탑재하고, 이를 프린터로 인쇄한 인쇄물을 다시 스마트폰 카메라로 촬영한 후, 탑재된 코드를 검출하는 방법을 설명한다. 전체 처리 과정은 그림 1과 같다.

2-2 코드탐재

코드 탑재는 인쇄물을 위한 컬러 모델인 CMYK (cyan, magenta, yellow, key plate) 컬러 모델 영상을 대상으로 하여, 코드가 탑재된 영상의 변화를 육안으로 식별하기 어렵도록 Y 채널을 이용한다 [3]. 그리고 코드가 탑재된 인쇄물의 일부만 촬영된 경우도 사진 식별이 가능하도록 주파수 도메인에 코드를 탑재한다.

코드 탑재 과정은 다음과 같다. 첫째, 입력 영상의 컬러 공간을 RGB (red, green, blue)에서 CMYK로 변환한다. 둘째, Y채널에 DFT (discrete Fourier transform) [7] 적용하여 주파수 매그니튜드 영상을 획득한다. 셋째, 이 영상의 중심으로부터 동일한 거리에 동일한 간격으로 1비트 코드를 탑재한다. 즉, 환형으로 코드를 탑재한다. 본 논문에서는 영상의 1, 2사분면에 16등분 간격으로 탑재한다. 3, 4사분면은 매그니튜드 영상의 특성상 1, 2사분면과 동일하게 탑재한다. 중심에서 떨어진 거리는 영상 높이의 0.175배로 하여, 중간 주파수 영역 대를 활용한다.

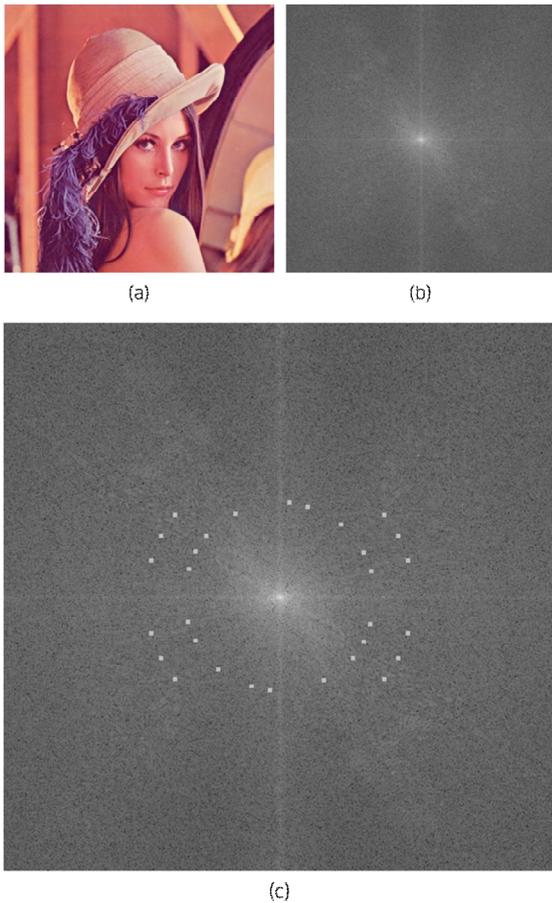


그림 2. 주파수 영역에서의 코드 삽입 영상 (a) 입력 영상 (b) 매그니튜드 영상 (c) 데이터 및 정렬 비트 삽입 영상
Fig. 2. Code embedded image in frequency domain. (a) input image (b) magnitude image (c) data and alignment bits embedded image.

중간 주파수 대역이 프린트-캡처 과정에서 코드 훼손이 적은 것으로 알려져 있기 때문이다 [6]. 1에 해당하는 비트는 앞서 설명한 구획 위치에 5x5 크기의 정사각형 영역을 설정하고 대응되는 계수를 500K만큼 증가시킨다. 한편, 코드 검출 단계에서 매그니튜드 영상은 회전에 영향을 받기 때문에, 코드 시작 위치 탐지가 요구된다. 이를 위해, 본 논문에서는 정렬을 위한 기준 비트를 삽입한다. 이는 탑재하고자 하는 16비트 데이터 비트와는 무관하도록, 데이터 비트를 위한 영역의 바깥쪽에 탑재하는 것을 제안한다. 본 논문에서는 데이터 비트보다 1.4배 바깥쪽을 사용한다. 또한, 검출의 안정성을 위해 회전 탐지를 위한 정렬 비트는 3개로 하였다. 그림 2는 Lena 영상에 대한 코드 탑재 전후 영상을 보여준다.

2-3 코드 검출

촬영된 영상에서 코드를 찾기 위한 과정은 그림 3과 같이 매그니튜드 영상 강화, 데이터 비트 반지름 추정, 기준 회전 각도 추정, 정렬 비트 검출, 데이터 비트 복호화로 구성된다.

1) 매그니튜드 강화

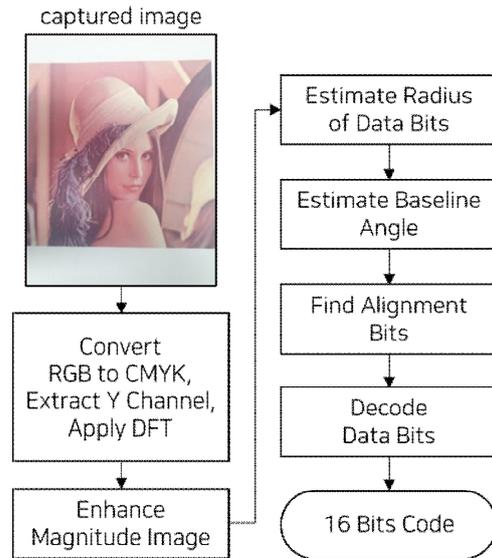


그림 3. 디코딩 과정
Fig. 3. Decoding Process.

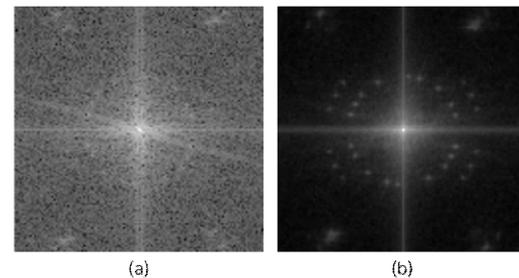


그림 4. 매그니튜드의 강화 (a) 40개 패치 중 한 패치에 대한 매그니튜드 영상 예시 (b) 제안방법 적용 강화 매그니튜드 영상
Fig. 4. Enhancement of magnitude image (a) an example magnitude image of a patch among 40 patches (b) enhanced magnitude image by our approach.

매그니튜드 영상 강화는 프린트-캡처 촬영 과정에서 훼손된 코드 비트 및 정렬 비트를 부각하기 위한 단계로, 처리순서는 다음과 같다. 먼저, 촬영된 영상의 가로 및 세로 중에서 작은 것을 기준으로 1/3 크기로 영상 패치를 40개 잘라낸다. 다음은 이렇게 잘라낸 패치에 대해 DFT를 적용하여 매그니튜드 영상을 구하고, 다시 이 매그니튜드 영상 간에 유사도를 계산하여 유사성이 높은 영상들만 누적하여 최종 매그니튜드 영상을 구한다. 이 과정을 통해 촬영 단계에서 포함될 수 있는 배경의 영향을 줄일 수 있다. 그림 4는 강화 전후 결과를 보여준다.

2) 데이터 비트열 영역의 반지름 추정

촬영 거리에 따라 매그니튜드 영상에서 비트열의 위치가 달라진다. 촬영 거리가 멀어질수록 비트열 영역은 안쪽으로 이동한다.

사용자는 임의의 거리에서 촬영할 수 있으므로 비트열이 존재하는 영역을 찾아내야 한다. 이를 위해 본 논문에서는 앞서 구한 매그니튜드 영상에 대해 다음과 같은 처리단계를 거

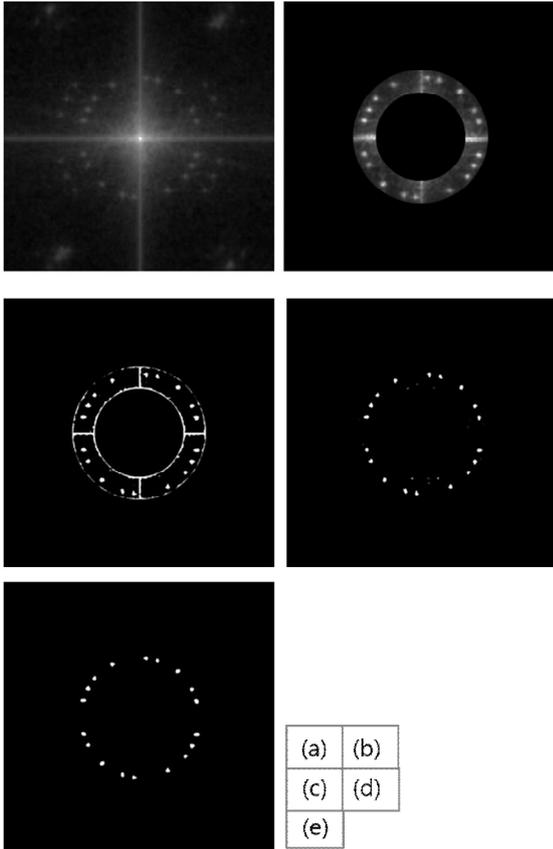


그림 5. 비트열 이진 영상 (a) 입력 매그니튜드 영상 (b) 도넛 모양의 마스크를 적용한 영상 (c) 적응적임계치를 적용한 이진 영상 (d) 허프변환 이용 선 제거 영상 (e) 모폴로지컬오픈 연산을 적용한 잡음제거 영상

Fig. 5. Bit sequence binary image (a) input magnitude image (b) donut-shape masked image (c) binary image using adopted threshold (d) line removed image using Hough transform (e) noise removed image using Morphological opening.

처 비트열이 존재하는 반지름을 추정한다. 먼저, 도넛 모양을 원점에서 바깥쪽으로 이동시켜 가며 이 영역에 한정하여 이진화를 수행한다. 그리고, 허프 (Hough) 변환을 적용하여 선 성분을 제거한다. 이렇게 획득한 이진 영상 중 흰색 픽셀의 수가 가장 많은 이진 영상을 선택한다. 각 과정에서 얻은 영상을 그림 5에 제시한다. 위 과정을 거쳐 얻은 최종 이진 영상의 흰색 픽셀에 대해 영상의 중심을 원의 중심으로 하는 원을 추정한다.

3) 기준 각도 추정

앞서 구한 이진 영상에 대해 도넛 영역을 등분하여 등분된 영역의 흰색 픽셀 비율을 고려하여 데이터 비트를 복호화 한다. 그러나, 등분하기 위한 기준 각도가 틀어지면 영역 분할이 잘못되어 올바르게 복호화를 할 수 없다. 그림 6은 올바르게 영역을 분할하는 경우와 그렇지 않은 경우를 보여준다.

기준 각도 추정을 위한 방법은 다음과 같다.

- 모든 흰색 픽셀에 대해 수평선과 이루는 각도를 계산한다.

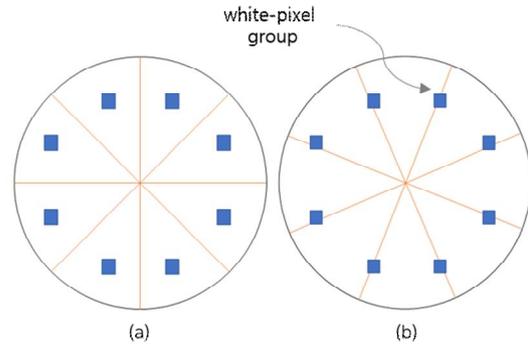


그림 6. 비트 영역분할 (a) 올바른 분할 (b) 잘못된 분할

Fig. 6. Bit-region division (a) properly divided region (b) improperly divided region.

- 각도를 11.25(180/16)도로 나눈 나머지를 계산한다.
- 나머지를 정수화하여 이 값의 빈도를 계산한다.
- 빈도가 최소인 빈(bin)의 값을 기준 각도로 설정한다.

본 논문에서는 16비트를 인코딩하므로 1, 2사분면을 대상으로 16등분 한다. 따라서 각 분할 영역은 11.25도 부채꼴 모양을 갖는다. 분할되는 경계에 흰색 픽셀이 적은 것이 좋으므로 빈의 값이 최소인 경우를 기준 각도로 설정한다.

4) 데이터 비트열 추출

데이터 비트가 위치하는 영역만 올바르게 분할되면 데이터 비트 값은 다음과 같이 간단한 방법으로 안정적으로 추출할 수 있다. 추정된 각도를 기준으로 도넛 영역을 16등분 하고 각 영역에 대해 흰색 픽셀의 비율을 계산하여 임계치보다 크면 1, 그렇지 않으면 0으로 한다. 본 논문에서는 임계치를 0.3으로 설정하였다. 이는 배경에 잡영이 일부 포함되는 경우를 고려하여 여유 있게 설정한 값이다.

5) 정렬 비트 검출 및 데이터 비트 복호화

정렬 비트는 데이터 비트열이 위치하는 바깥쪽 영역에, 구체적으로는 데이터 비트열 영역 반지름의 1.4배 위치에 인코딩하였다. 이 위치와 앞서 구한 기준 각도를 이용하면 쉽게 정렬 비트의 위치를 찾을 수 있다.

정렬 비트의 강건한 검출을 위해 1, 2사분면에 각각 3개의 점으로 구성하였다. 1사분면 3개의 점 중, 수평선에 가까운 점에 대응되는 데이터 비트를 데이터 비트의 MSB (most significant bit)로 한다.

III. 실험결과 및 검토

제안 방법은 코드가 탑재된 영상의 윤곽이 모두 포함되지 않아도, 수직으로 찍히지 않고 회전된 상태로 촬영되는 경우에도 동작하도록 설계하였다.

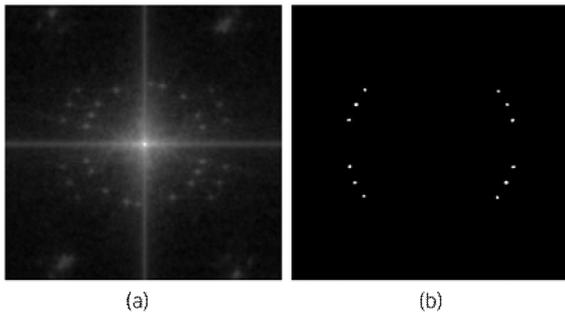


그림 7. 정렬 비트 영상 (a) 입력 매그니튜드 영상 (b) 정렬 비트의 이진화 영상
Fig. 7. Alignment bit image (a) input magnitude image (b) alignment bits binary image.

실험에서는 이러한 특성을 평가할 수 있도록 구성하였다. 실험에 사용한 원본 영상은 총 5장이다. 원본 5장에 대해 정면에서 각각 20장씩 촬영하여 총 100장을 촬영하였다. 이때, 수직 방향으로 10장, 임의의 각도로 회전하여 10장씩을 촬영하였다. 주목할 점은 수직 방향 영상은 대체로 원본 영상 전체가 촬영되었고, 회전된 영상은 원본 영상의 70% 수준에 해당하는 일부 영역만 촬영되었다는 점이다. 그림 8은 원본 5장 영상과 인쇄 후 스마트폰으로 촬영한 영상들의 일부를 보여준다.

성능평가는 삽입된 데이터 비트열과 제안 방법을 통해 복호화 데이터 비트열의 해밍거리를 활용한다. 해밍거리가 포함된 최종 유사도 계산은 수식 (1)로 정의한다.

$$Accuracy = \frac{1}{N} \left(1 - \frac{HD(x, \hat{x})}{L_x} \right) \quad (1)$$

여기서 N 은 테스트 이미지의 수, $HD(x, \hat{x})$ 는 삽입된 코드 x 와 검출된 코드 \hat{x} 간의 해밍거리, L_x 는 코드의 길이를 뜻한다.

수직 방향으로 촬영된 경우와 임의 방향으로 촬영된 경우에 대한 복호화 정확도는 표 1에 정리하였다. 인식률은 최소 94.2에서 100%로 나타났다. 이는 오류 비트의 수가 0.93개에서 0개에 해당한다. 즉, 모든 경우에 대해 평균적으로 오류 비트가 1개 이하로 나타났다.

임의 각도 방향으로 촬영한 영상에 대한 인식률은 95.2%로 수직 방향으로 촬영한 영상에 비해 3.6% 낮은 수치를 보였지만, 오류 비트 수로 따지면 각각 0.2개, 0.76개로, 0.56개 차이에 불과하다. 여기서 주목해야 할 것은 영상의 일부만, 즉 전체 대비 70% 수준만 촬영되었다는 점이다. 이것은 제안 방법론이 임의 촬영 각도와 일부 영역만 촬영되는 경우에도 성공적으로 동작한다는 것을 의미한다.

IV. 결론

본 논문에서는 프린트-캡처 응용을 위해 주파수 도메인 영상에 코드를 탑재하고 검출하는 방법을 제안하였다. 프린터-카메라 촬영 과정에서 탑재 정보가 훼손되지 않도록 주파수 도메인의 중간 주파수 영역 대에 환형으로 데이터 비트열을

배치하였고, 추가로 데이터 비트열 바깥쪽에 정렬 비트를 삽입하도록 설계하였다. 또한, 탑재된 비트열을 효과적으로 검출하기 위한 알고리즘을 제시하였다.

연구실에서 수집한 영상에 대한 실험에서 임의 회전 하에서 일부만 촬영된 영상에 대해서도 평균 오류 비트 수는 0.7개 수준으로 매우 우수한 성능을 보였다. 향후, 관련 딥러닝 기술을 도입하여 실용적 수준의 용량으로 확장할 필요가 있다 [8], [9].

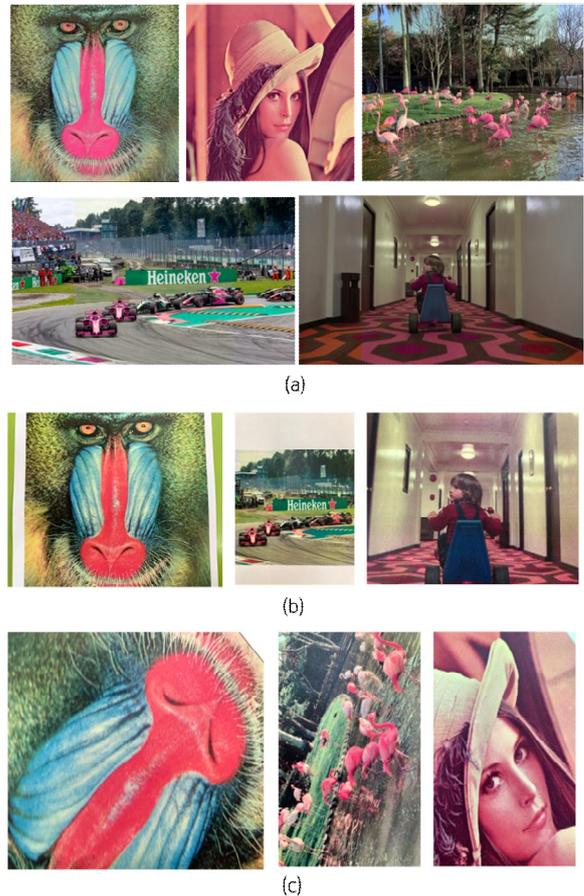


그림 8. 실험 데이터 집합 (a) 원본 영상 (b) 수직 방향으로 전체가 모두 포함되도록 촬영한 영상 (c) 무작위 방향으로 원본 영상의 일부분만 포함되도록 촬영한 영상
Fig. 8. Experimental data set (a) original images (b) fully captured images in vertical direction (c) partially captured images in random direction.

표 1. 디코드 정확도

Table 1. Decoding Accuracy.

Image Name	Shooting Direction	
	Vertical	Random
baboon	98.8	94.4
lena	100	97.7
flamingo	96.3	93.8
racing	99.4	96.2
shining	99.4	94.2
All	98.8	95.2

Acknowledgement

이 연구는 금오공과대학교 학술연구비로 지원되었음 (2018104083).

References

- [1] M. Hirokawa, and J. Iijima, "A study on digital watermarking usage in the mobile marketing field: cases in Japan," in *Proceedings of IEEE Symposium on Logistics and Industrial Informatics*, Linz: Austria, pp. 1-6, Sept. 2009.
- [2] M. Tancik, B. Mildenhall., and R. Ng., "Stegastamp: invisible hyperlinks in physical photographs", in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Seattle: WA, pp. 2117-2126, June 2020.
- [3] D. He, and Q. Sun, "A practical print-scan resilient watermarking scheme," in *Proceedings of IEEE Conference on Image Processing*, Genova: Italy, pp. 257, Sept. 2005.
- [4] T. Nakamura, A. Katayama, M. Yamamuro, and N. Sonchra,
- "Fast watermark detection scheme for camera-equipped cellular phone," in *Proceedings of ACM Conference on Mobile and Ubiquitous Multimedia*, New York: NY, pp. 101-108, Oct. 2004.
- [5] A. Pramil, A. Keskinarkaus, and T. Seppänen, "Toward an interactive poster using digital watermarking and a mobile phone camera," *Signal, Image and Video Processing*, Vol. 6, No. 2, pp. 211-222, 2012.
- [6] A. Poljicak, L. Mandic, and D. Agic, "Discrete fourier transform-based watermarking method with an optimal implementation radius," *Journal of Electronic Imaging*, Vol. 20, No. 3, pp. 033008-1-033008-8, 2011.
- [7] R. C. Gonzalez, and R. E. Woods, *Digital Image Processing*, Addison-Wesley Pub., ch. 3, pp. 120-128, 1992.
- [8] K. Zhang., A. Cuesta-Infane, L. Xu., and K. Veermachaneni., "SteganoGAN: high capacity image steganography with GANs," ArXiv, 2019. Available: <https://arxiv.org/abs/1901.03892>
- [9] J. Liu, Y. Ke, and X. Yang, "Recent advances of image steganography with generative adversarial networks," *IEEE Access*, Vol. 8, pp. 60575-60579, 2020.



배 준 현 (Joon-Hyeon Bae)

2016년 3월 ~ 현재 : 금오공과대학교 컴퓨터공학과 학사과정
 ※ 관심분야 : 컴퓨터 비전, 영상처리, 인공지능



홍 원 기 (Won-Kee Hong)

2001년 8월 : 연세대학교 컴퓨터과학과 (공학박사)
 2004년 3월 ~ 현재 : 대구대학교 정보통신대학 멀티미디어공학과 교수
 ※ 관심분야 : 모바일 컴퓨팅, IoT, 인공지능



고 재 필 (Jaepil Ko)

2004년 2월 : 연세대학교 컴퓨터과학과 (공학박사)
 2004년 3월 ~ 현재 : 금오공과대학교 컴퓨터공학과 교수
 ※ 관심분야 : 영상처리, 컴퓨터비전, 머신러닝