



# 역방향 하프토닝을 위한 다중 손실 계층 및 영상 구조 맵 예측기에 기반한 다중 스트림 네트워크

손 창 환\*

## Multi-Stream Networks with Multi-Loss Layers and Image Structure Map Predictor for Inverse Halftoning

Chang-Hwan Son\*

---

이 논문은 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임  
(No. 2017R1D1A3B03030853)

---

### 요 약

이 논문에서는 하프톤 영상에서 연속 계조 영상을 복원하는 역방향 하프토닝 기법에 대해 소개하고자 한다. 최근 영상 복원 분야에서 큰 주목을 받고 있는 심층 합성곱 신경망 기법을 역방향 하프토닝 분야에 적용할지라도 평탄 영역에서 하프톤 패턴의 불완전한 제거나 에지 및 텍스처 영역에서 디테일 표현 부족은 여전히 현안으로 남아 있다. 이러한 문제를 해결하고자 이 논문에서는 다중 손실 계층을 도입해서 영상 구조 맵과 연속 계조 영상을 동시에 추정이 가능한 다중 스트림 기반의 심층 합성곱 신경망을 새롭게 제안하고자 한다. 그리고 실험 결과를 통해, 제안한 기법이 기존의 최첨단 기법들보다 화질 성능 측면에서 더 우수한 결과를 달성할 수 있음을 보이고자 한다.

### Abstract

This paper introduces the inverse halftoning method for reconstructing the continuous-tone images from the halftoned images. Even though recently-introduced deep convolutional neural networks having drawn much attention from image restoration areas are applied for the inverse halftoning, it still remains a pending issue to remove the halftone patterns completely in flat regions and improve the details in edge and texture regions. To address this problem, this paper presents a new multi-stream-based deep convolutional neural network, which enables the image structure map and the continuous-tone image to be estimated jointly by using multi-loss layers. Through the experimental results, it is also confirmed that the proposed method achieves better results than the conventional state-of-the-art methods in terms of visual quality performance.

### Keywords

inverse halftoning, image restoration, convolutional neural networks, multi-loss layers, dictionary learning

---

\* 군산대학교 소프트웨어융합공학과  
- ORCID<sup>1</sup>: <https://orcid.org/0000-0001-7077-3074>

• Received: Dec. 21, 2018, Revised: Feb. 26, 2019, Accepted: Mar. 01, 2019  
• Corresponding Author: Chang-Hwan Son  
Department of Software Convergence Engineering Republic of Korea,  
Tel.: +82-63-469-8915, Email: [cson@kunsan.ac.kr](mailto:cson@kunsan.ac.kr)

## 1. 서론

디지털 하프토닝(Digital Halftoning)이란 디지털 영상을 흰색과 검정색의 도트 패턴으로 구성된 하프톤 영상(Halftoned Images)을 생성하는 기법을 말한다[1]. 하프톤 영상에서 흰색과 검정색으로 구성된 도트 패턴은 실제 프린터의 토너나 잉크가 찍히는 위치를 나타낸다. 따라서 디지털 하프토닝 기술은 주로 복합기, 프린터, 또는 팩스와 같은 인쇄 장치에 사용된다. 그리고 플라즈마 디스플레이에서 저계조 표현시에 발생하는 의사윤곽(False Contour)을 저감하기 위해 사용하기도 한다[2].

역방향 하프토닝(Inverse Halftoning)이란 디지털 하프토닝의 역과정으로써, 흰색과 검정색 도트 패턴으로 구성된 하프톤 영상에서 연속계조 영상(Continuous-tone Images)를 복원하는 과정이다[3]. 여기서 연속계조 영상이란  $n$ 개의 밝기 레벨로 구성된 디지털 영상을 말한다. 이러한 역방향 하프토닝은 디스크리닝(Descreening)이라 불리기도 한다. 역방향 하프토닝은 팩스 전송시 하프톤 영상의 압축 효율화, 스캐닝 시 모아레(Moire) 결함 제거, 스캔된 만화(Comic Books)의 화질 개선, 그리고 최근에는 밀집 바이너리 센서(Denser Binary Sensors)의 광대역(High Dynamic Range) 실현을 위한 영상 복원 과정에서도 꾸준히 활용되고 있다[4]-[6].

역방향 하프토닝은 이진 계조에서  $n$ 개의 연속 계조를 추정하는 문제를 푸는 과정이므로 불량 설정 문제(ill-posed Problem)에 속한다. 이 문제를 푸는 기법에는 가장 단순한 가우시안 필터링에서부터 최대사후확률(Maximum-A-Posterior)[7], 룩업 테이블(Look-up Table)[3], 비등방성 필터링(Anisotropic Filtering)[8] 접근법 등이 있다. 근래에는 사전 학습(Dictionary Learning) 기반의 역방향 하프토닝 기법[9][10]이 주류를 이루고 있는 가운데 신경망(Neural Networks) 기반의 접근 방식도 고려되고 있다[11]. 최근 영상 복원 분야에서 심층 합성곱 신경망(DCNN, Deep Convolutional Neural Networks) 기법이 성능 평가에서 큰 성과를 거두고 있다[12]. 물론 이 심층 합성곱 신경망 기법을 역방향 하프토닝 분야에 적용하면 사전 학습 기법보다 더 우수한 성능을 얻을 수 있다. 하지만 디테일 영역에서 선명도

저하나 영상구조의 표현 능력은 여전히 개선해야 될 사안으로 남아 있다. 따라서 이 논문에서는 기존의 심층 합성곱 신경망 아키텍처를 수정하여, 즉 다중 손실 계층과 영상 구조 맵 예측기를 갖는 다중 스트림 기반의 심층 합성곱 신경망을 새롭게 제안하여 디테일 영역의 선명도 향상과 영상구조 표현 능력을 강화하고자 한다. 좀 더 설명하자면, 제안한 심층 합성곱 신경망은 입력 하프톤 영상에서 원본 영상의 영상구조 맵을 추정하기 위한 서브 네트워크와, 그리고 이 서브 네트워크의 결과와 입력 하프톤 영상을 스택으로 쌓아 연속계조 영상을 복원하기 위한 서브 네트워크로 구성된다. 두 종류의 서브 네트워크가 존재하므로 다중 손실계층으로 구성되고, 영상구조 맵 추정을 위한 서브 네트워크의 결과가 연속계조 영상 복원을 위한 서브 네트워크에게 라인 또는 텍스처가 존재하는 영역을 제공할 수 있기 때문에 복원된 연속계조 영상에서 영상구조의 선명도와 디테일을 강화할 수 있다.

## II. 최신 관련 연구

### 2.1 사전 학습 기반의 역방향 하프토닝

사전 학습이란 기저벡터들의 선형 결합으로 입력 영상 패치를 표현하는 기법을 말한다[13]. 이때 기저벡터의 집합은 사전이라 불리며 푸리에 변환 또는 웨이블릿 변환의 기저벡터로 대체 가능하다. 하지만 최근에는 고정된 형태의 사전이 아닌, 훈련 집합으로부터 학습된 사전을 사용하고 있다. 사전 학습 모델은 다음과 같이 표현된다.

$$\operatorname{argmin}_{D^h, D^c, A} \left\| \begin{bmatrix} Y^h \\ Y^c \end{bmatrix} - \begin{bmatrix} D^h \\ D^c \end{bmatrix} A \right\|_2^2 \quad (1)$$

$$\text{subject to } \forall j, \|\alpha_j\|_0 \leq TH$$

여기서  $Y^h$ 와  $Y^c$ 는 각각 하프톤 패치와 그에 대응하는 연속계조 패치를 포함하고 있다. 좀 더 자세하게 말하자면, 연속계조 패치들을 열벡터로 모양을 변경한 다음에  $Y^c$  행렬의 각각의 열에 저장한다.

그리고 연속계조 패치들을 하프토닝을 적용한 후, 생성된 하프톤 패치들도 열벡터로 변경한 다음

$Y^h$  행렬의 각각의 열에 저장한다. 이때 연속계조 패치와 하프톤 패치는  $Y^c$ 와  $Y^h$ 의 동일한 열에 저장되어야 한다. 식 (1)에서  $\alpha_j$ 는  $A$ 행렬의  $j$ 번째 열벡터에 대응한다. 이때  $\alpha_j$ 는 희소 제약(Sparse Constraint) 조건, 즉  $\|\alpha_j\|_0 \leq TH$ 을 만족해야 된다. 이는  $\alpha_j$  열벡터의 원소의 대부분이 영이 된다는 것을 의미한다. 식 (1)은 K-SVD[13]와 같은 사전 학습 알고리즘으로 풀 수 있고, 이를 통해 최종  $D^h$ 와  $D^c$  행렬, 즉 하프톤 패치와 연속계조 패치를 표현할 수 있는 사전을 생성할 수 있다.

$D^h$ 와  $D^c$ 가 주어진 경우, 입력 하프톤 영상으로부터 연속계조 영상을 복원할 수 있다. 복원 과정은 패치 기반으로 수행된다.

$$\begin{aligned} \operatorname{argmin}_{\alpha_{i,j}} \|D^h \alpha_{i,j} - R_{i,j} Y\|_2^2 \\ \text{subject to } \|\alpha_{i,j}\|_0 \leq TH \end{aligned} \quad (2)$$

여기서  $Y$ 는 입력 하프톤 영상이고  $R_{i,j}$ 는  $(i,j)$  위치에서 패치를 추출하기 위한 연산자이다. 따라서 식 (2)는 하프톤 패치  $R_{i,j} Y$ 가  $D^h \alpha_{i,j}$ 로 모델링할 수 있음을 의미한다. 즉 입력 하프톤 패치가 하프톤 사전의 선형조합으로 표현 가능하다는 것이다. 이때  $\alpha_{i,j}$ 는 희소 제약을 만족하는 벡터이고, 정합 추적(Matching Pursuit)[13] 알고리즘을 사용해서 추정할 수 있다. 끝으로 패치 기반으로 복원된 연속계조 패치들은 패치 간 중복영역에서 평균화 과정을 통해 최종 입력 영상과 동일한 크기의 연속계조 영상을 만든다. 이런 사전 학습 기반의 역방향 하프토닝 기법은 영상 구조에 최적화 된 사전을 생성하기 때문에 기존의 최대사후확률, 룩 업 테이블, 비등방성

필터링 기반의 접근 방식보다 영상 구조 표현에 있어서 더 우수한 결과를 도출할 수 있다. 식 (1)과 (2)에서 스파스 한 특성을 보존한다면,  $TH$ 의 값은 신호 차원의 반, 즉 패치 크기의 반까지 설정할 수 있다. 실제로 식 (1)과 (2)는 특정 표현 오차가 만족 될 때까지 기저벡터의 개수를 증가하는 방식으로 구현되었다. 자세한 구현 사항을 알고 싶다면, 참고 문헌 [9][13]의 공개 소스코드를 보시기 바란다.

## 2.2 심층 신경망 기반의 역방향 하프토닝

심층 합성곱 신경망 기법은 최근 영상복원 및 인식 분야에서 탁월한 성능을 보이고 있다[14][15]. 따라서 심층 합성곱 신경망 기법을 역방향 하프토닝 분야에도 그대로 적용이 가능하다. 심층 합성곱 신경망의 아키텍처는 그림 1과 같이 합성곱(Convolution) 계층과 정류선형유닛(ReLU, Rectified Linear Unit) 계층으로 구성된다.

그리고 맨 마지막 계층은 학습할 때만 사용되는 손실 계층으로써, 원본 영상과 복원된 영상간의 오차를 계산한다. 주로 유클리디언 거리가 오차의 척도로 사용된다. 테스트 단계에서는 손실 계층을 제거한 후, 입력 하프톤 영상을 학습된 신경망에 입력으로 주면 연속계조영상을 획득할 수 있다. 이와 같은 심층 합성곱 신경망 기법을 역방향 하프토닝에 적용하면, 기존의 사전 학습 기법보다 더 우수한 결과를 얻을 수 있다. 왜냐하면 사전 학습은 얇은 학습(Shallow Learning)에 속하기 때문에 심층 합성곱 신경망처럼 선형과 비선형을 조합으로 저수준에서 고수준까지의 매핑 관계를 모델링하지 못하기 때문이다.

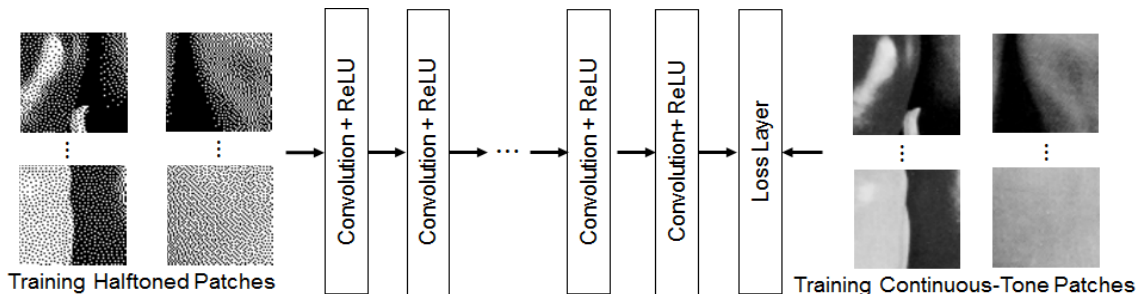


그림 1. 역방향 하프토닝을 위한 심층 합성곱 신경망의 아키텍처  
Fig. 1. Architecture of deep convolutional neural networks for inverse halftoning

하지만 에지나 텍스처 영역에서 디테일 표현 능력은 여전히 개선할 여지가 있다. 그림 1에 보듯이 기존의 심층 합성곱 신경망은 입력 하프톤 패치의 영상구조를 고려하지 않기 때문에 디테일 표현 능력을 강화할 여지가 있다. 따라서 이 논문에서는 기존의 심층 합성곱 신경망에 영상 구조 맵을 추정하는 서브 네트워크를 추가해서 다중 손실 계층 기반으로 에지나 텍스처의 표현 능력을 개선하고자 한다.

### III. 제안한 역방향 하프토닝 기법

#### 3.1 다중 스트림 네트워크 아키텍처

그림 2는 학습 단계에서 사용할 제안한 다중 스트림 네트워크의 아키텍처를 보여주고 있다. 제안한 신경망의 아키텍처는 크게 3개의 서브 네트워크로 구성된다. 먼저, 좌측 상단의 파란색 박스에 표시된 서브 네트워크는 미리 학습된 모델로써, 입력 하프톤 패치로부터 초기 연속계조 영상 패치를 추정한다. 그리고 우측 상단의 파란색 박스에 표시된 서브 네트워크는 앞단의 선행 학습의 결과로부터 영상 구조 맵을 예측한다. 이를 위해 손실 계층에서 입력 하프톤 패치에 대응하는 호그(HOG, Histogram of Oriented Gradients)[16] 패치를 입력으로 받는다.

여기서 호그 패치란 영상 구조에 관한 정보를 담고 있는 패치를 일컫는다.

마지막으로 우측 하단의 서브 네트워크는 우리가 원하는 연속계조 패치를 복원한다. 이를 위해 손실 계층에서는 원본 연속계조 패치를 입력으로 받고, 입력 계층에서는 연결 계층(Concatenation Layer)을 통과한 2개의 특징 맵(Feature Maps)을 입력으로 받는다. 연결 계층에서는 입력 하프톤 패치와 좌측 상단의 서브 네트워크의 결과인 특징 맵을 스택으로 쌓는 작업을 한다.

그림 2에 제시된 제안한 다중 스트림 네트워크의 기본적인 학습 유도 전략은 좌측 상단의 미리 학습된 서브 네트워크를 통해서 초기 연속계조 영상을 추정한 후, 이 영상으로부터 영상구조 맵을 예측하면서 동시에 최종 연속계조 영상을 복원하는 것에 있다. 이를 통해 디테일 영역의 표현 능력을 강화하고 평탄 영역의 도트 패턴 제거 능력을 제고하고자 한다.

그림 1의 기존의 심층 합성곱 신경망은 입력 하프톤 영상에서 원본 영상의 영상구조의 추정 없이, 즉 입력 하프톤 영상에서 평탄한 영역, 라인 및 텍스처가 존재하는 영역을 구별하지 않고 훈련패치를 학습한다.

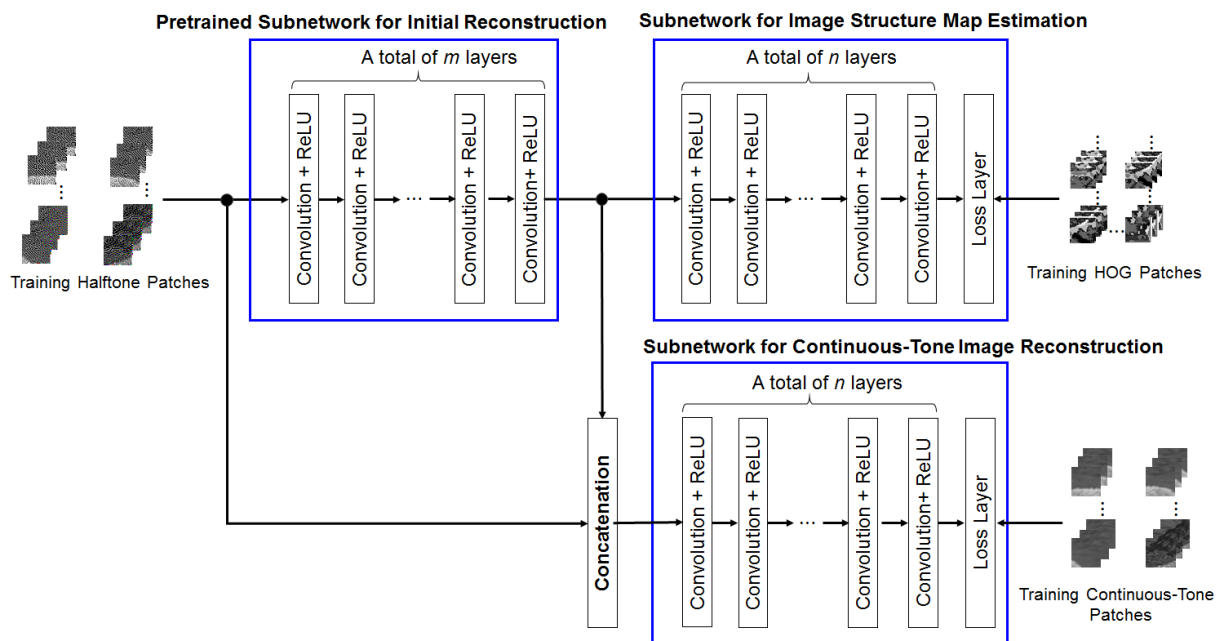


그림 2. 제안한 다중 손실 계층과 영상 구조 맵 예측기에 기반한 다중 스트림 네트워크  
Fig. 2. Proposed multi-stream network via multi-loss layers and image structure map predictor

반면에, 제안한 심층 합성곱 신경망에서는 영상 구조 맵 예측을 위한 서브 네트워크를 추가함으로써, 영상구조 맵의 결과가 연속계조 영상 복원을 위한 서브 네트워크에게 평탄 영역, 그리고 라인 및 텍스처가 존재하는 영역을 알려줄 수 있기 때문에 복원된 연속계조 영상에서 영상구조의 선명도와 디테일을 강화할 수 있다.

### 3.2 훈련 패치 생성 과정

학습 단계에서 필요한 패치는 연속계조 패치, 하프톤 패치, 그리고 영상구조 정보를 포함한 호그 패치이다. 이 3종류의 패치들을 생성하기 위해서는 먼저 연속계조 영상들을 수집해야 한다. 연속계조 영상들은 인터넷 웹사이트나 영상처리 분야에서 주로 사용되는 기준 데이터베이스에서 수집이 가능하다. 이 논문에서는 총 500장의 영상을 인터넷 사이트에서 수집했다. 하프톤 영상은 수집된 연속계조 영상에 오차 확산(Error Diffusion) 하프토닝 기법[1]을 적용해서 생성한다. 그리고 연속계조 영상에서 영상구조 정보를 갖는 호그 영상은 3단계를 거쳐 생성된다. 그림 3은 호그 맵 생성 과정에 대한 전체 블록도를 보여주고 있다.

**1단계:** 연속계조 영상에서  $8 \times 8$  크기의 윈도우 창을 래스터 스캐닝하면서 윈도우 창안의 각 픽셀 위치에서 그라디언트의 크기(Gradient Magnitude)와

방향(Gradient Orientation)을 계산한다.

$$m(i, j) = \sqrt{H(i, j)^2 + V(i, j)^2} \quad (3)$$

$$\theta(i, j) = \tan^{-1} \frac{V(i, j)}{H(i, j)} \quad (4)$$

$$H(i, j) = F(i + 1, j) - F(i - 1, j) \quad (5)$$

$$V(i, j) = F(i, j + 1) - F(i, j - 1) \quad (6)$$

여기서  $F(i, j)$ 는  $(i, j)$  픽셀 위치에서 연속계조 영상의 밝기 값이고,  $m(i, j)$ 와  $\theta(i, j)$ 는 각각 그라디언트의 크기와 방향을 나타낸다.

**2단계:** 그라디언트의 방향을 9개로 양자화한 후,  $8 \times 8$  윈도우 창안에서 그라디언트 방향에 따른 그라디언트 크기를 누적해서 히스토그램을 생성한다. 래스터 스캐닝 방향을 따라 동일한 과정을 반복하면서 모든 연속계조 영상에서  $8 \times 8$  윈도우 창에서 히스토그램을 생성한 후, K-평균 군집화(K-means Clustering) 알고리즘을 적용해서 16개의 군집의 센터 벡터를 뽑아낸다.

**3단계:** 수집된 연속계조 영상에서 호그 맵을 생성하기 위해, 1단계처럼  $8 \times 8$  윈도우 창을 연속계조 영상에서 썬 후, 래스터 스캐닝을 하면서 히스토그램을 추출한다.

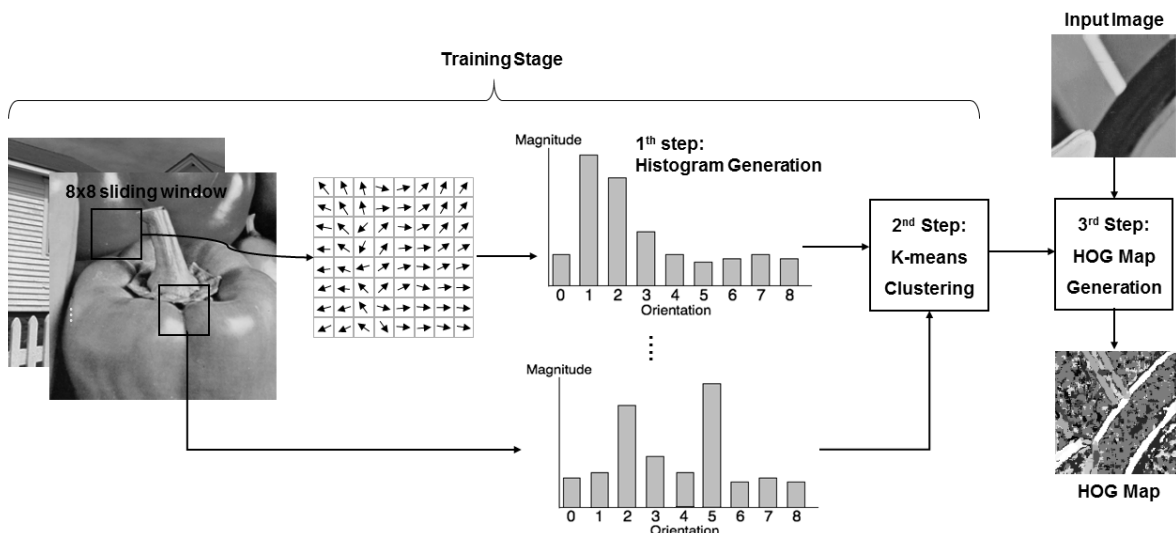


그림 3. 호그 맵 생성 과정  
Fig. 3. HOG map generation

그리고 2단계에서 생성된 16개의 군집센터 벡터와 비교해서 가장 유사한 센터 벡터의 인덱스 값으로 사상한다. 여기서 유사성 척도는 유클리디언 거리에 반비례하도록 모델링된다. 따라서 호그 맵은 결국 16개의 센터 벡터에 대한 인덱스 값으로 구성된다. 이 인덱스 값은 8×8 국부 영역에서의 그라디언트 분포 특성을 반영한 것으로 볼 수 있다. 예시로써, 그림 3의 우측에 입력 연속계조 영상에 대한 호그 맵이 주어져 있다. 그림에서 볼 수 있듯이, 라인이나 평탄한 영역들이 유사한 인덱스 값으로 군집을 이룬 것을 볼 수 있다.

수집된 연속계조 영상들에 대한 하프톤 영상과 호그 맵이 확보되면 동일한 위치에서 패치를 추출할 수 있다. 이 논문에서 사용된 패치 크기는 영상 복원 분야에서 주로 사용되는 32×32로 선택했다. 이렇게 추출된 연속계조 패치, 하프톤 패치, 호그 패치들은 최종적으로 그림 2에서 제시된 다중 스트림 네트워크를 학습하기 위해 사용된다.

### 3.3 학습 및 테스트 단계

그림 2에서 합성곱과 정류선형유닛 쌍의 개수는 초기 영상 복원을 위한 서브 네트워크의 경우에는 17개, 그리고 영상 구조 예측과 연속계조 영상 복원을 위한 서브 네트워크의 경우에는 8개이다. 그리고 합성곱 계층의 필터의 크기는 5×5, 필터 개수는 64개, 패딩 크기는 2로 설정된다. 심층 합성곱 신경망

의 파라미터는 확률적 경사 하강(Stochastic Gradient Descent) 기법을 통해 학습된다. 즉, 배치 크기가 64이고, 학습률(Learning Rate)은  $10^{-5}$ , 총 에폭(Epochs)은 250회이다. 이때 에폭 당 64,000개의 32×32 패치를 500장의 훈련영상 셋에서 랜덤하게 추출한다. 손실함수는 추정된 패치와 원본 패치 또는 추정된 호그 맵과 원본 호그 맵과의 유클리디언 거리의 총합으로 계산된다. 그리고 다중 손실함수의 가중치 비율은 동일하게 설정된다. 제안한 다중 스트림 네트워크 기반의 심층합성곱 신경망이 학습되면 입력 하프톤 영상을 연속계조 영상으로 변환할 수 있다. 테스트 단계에서 최종 복원 영상의 결과는 서브 네트워크(그림 2의 우측 하단) 손실 계층의 입력 특징 맵에 해당한다. 참고로 제안한 다중 스트림 네트워크는 그레이 영상으로 학습된다. 따라서 테스트 영상이 칼라인 경우는 3채널로 분리해서 최종 결과 영상을 복원한다.

## IV. 실험 및 결과

### 4.1 실험 환경

제안한 다중 스트림 네트워크 기반의 역방향 하프토닝 기법을 평가하기 위해, 훈련 집합에 포함되지 않은 총 10장의 영상을 테스트했다. 그림 4는 이 논문에서 사용된 테스트 영상 집합을 보여주고 있다.



그림 4. 테스트 영상 집합  
Fig. 4. Test image set

그림에서 보듯이 테스트 영상이 다양한 영상 구조, 예를 들어 라인, 곡선 등을 포함하고 있는 것을 볼 수 있다. 그리고 테스트 영상을 구분하기 위해 영상 하단에 인덱스 번호를 붙여 놓았다. 비교 대상으로 최첨단 기법의 사전 학습 기반의 역방향 하프토닝 기법[9]과 심층 합성곱 신경망 기반의 역방향 하프토닝 기법[14]을 선택했다.

그리고 정량적인 화질 평가를 위해, 영상 복원 분야에서 널리 사용되는 최대 신호 대 잡음비 (PSNR, Peak Signal-to-Noise Ratio)와 구조 유사성 (SSIM, Structure Similarity)[17] 척도를 사용했다. 최대 신호 대 잡음비는 원본 영상과 복원 영상간의 픽셀 값의 오차가 얼마나 나는지를 판별하는 척도이며 반면에 구조 유사성은 두 영상간의 라인, 곡선, 패턴의 모양이 얼마나 유사한지를 반영할 수 있는 척도이다. 두 척도 모두 수치가 높을수록 복원 영상과 원본 영상이 유사하다는 것을 의미한다. 제안한 기법은 매트랩 언어로 구현되었고 1080Ti GPU 2개로 구성된 워크스테이션에서 실행되었다.

#### 4.2 실험 결과

그림 5는 기존의 방법과 제안한 방법을 사용해서 복원된 결과 영상을 비교해주고 있다. 맨 좌측부터 입력 하프톤 영상, 사전학습 기반의 역방향 하프토닝을 적용한 결과 영상, 심층 합성곱 신경망 기반의 역방향 하프토닝을 적용한 결과 영상, 제안한 기법을 적용한 결과 영상, 그리고 원본 영상이 차례대로 나열되어 있다.

그림 5의 빨간색 박스에서 보듯이, 제안한 기법이 기존의 최첨단 기법들보다 영상구조 표현 능력이 더 우수하다는 것을 알 수 있다. 특히 실험 영상 4번과 7번의 경우, 기존의 기법들은 바지의 줄무늬나 모자의 패턴을 거의 복원하지 못한 것을 볼 수 있다.

또한 8번 영상과 같은 경우에는 선명도가 전반적으로 떨어지는 것을 볼 수 있다. 이는 제안한 다중 스트림 기반의 역방향 하프토닝 기법은 영상 구조 맵을 추정하고 이 맵 정보를 입력 하프톤 영상과 함께 최종 연속계조 영상을 복원하기 때문이다. 즉 영상 구조 맵을 학습시에 활용하기 때문에 복원 영상에서 라인, 곡선, 패턴의 선명도가 강화되거나 복구 능력이 개선될 수 있다.

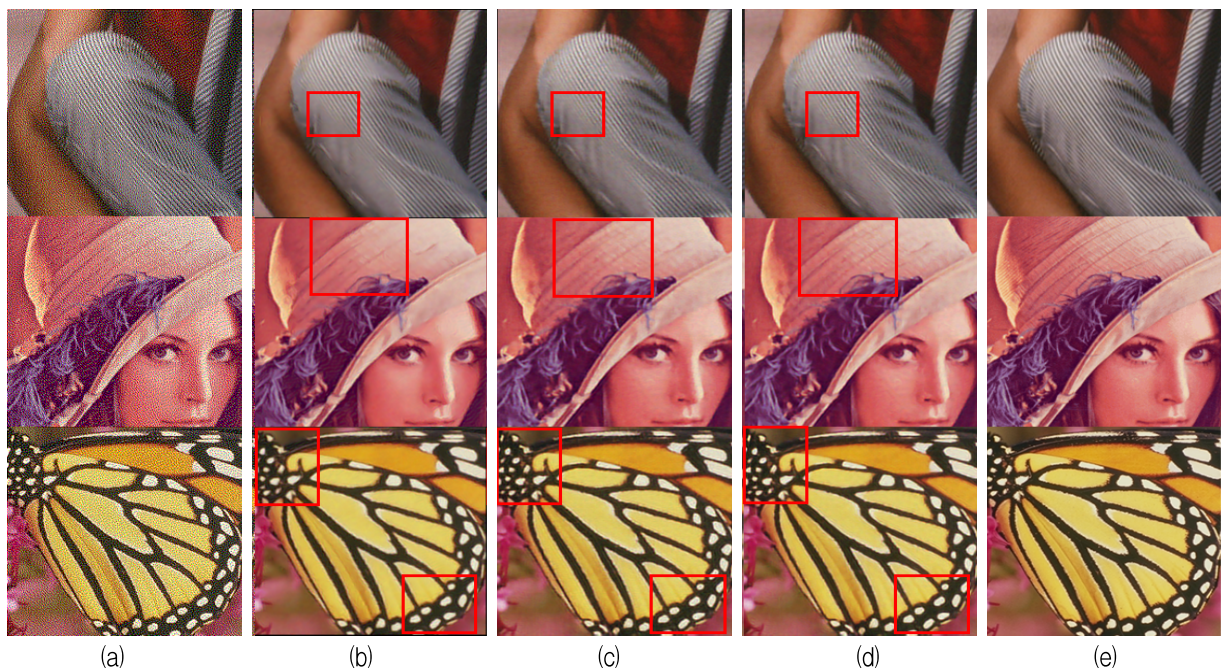


그림 5. 실험결과: (a) 입력 하프톤 영상, (b) 사전 학습 기반의 역방향 하프토닝 [9], (c) 심층 합성곱 신경망 기반의 역방향 하프토닝 [14], (d) 제안한 다중 스트림 네트워크 기반의 역방향 하프토닝, (e) 원본 연속계조 영상  
 Fig. 5. Experimental results: (a) Input halftoned images, (b) Inverse halftoning via dictionary learning[9], (c) Inverse halftoning via DCNN [14], (d) Inverse halftoning via proposed multi-stream networks, (e) Original continuous-tone images (left to right)

표 1. 정량적인 화질 평가

Table 1. Quantitative image quality evaluation

Test Images		1	2	3	4	5	6	7	8	9	10	AVG.
Proposed Method	PSNR	<b>25.467</b>	25.530	<b>25.201</b>	<b>29.624</b>	31.738	<b>25.722</b>	<b>31.221</b>	<b>27.782</b>	<b>23.596</b>	<b>29.472</b>	<b>27.535</b>
	SSIM	<b>0.815</b>	<b>0.903</b>	<b>0.856</b>	<b>0.883</b>	0.978	<b>0.898</b>	<b>0.980</b>	<b>0.965</b>	<b>0.952</b>	<b>0.930</b>	<b>0.916</b>
DCNN [14]	PSNR	25.181	25.395	24.810	28.608	<b>31.818</b>	25.370	31.084	27.275	23.413	29.214	27.217
	SSIM	0.808	0.900	0.846	0.854	<b>0.979</b>	0.890	0.979	0.959	0.949	0.928	0.909
Dictionary Learning [9]	PSNR	25.016	<b>25.541</b>	24.654	28.500	31.023	25.150	30.514	27.563	22.845	28.645	26.945
	SSIM	0.799	0.818	0.798	0.846	0.870	0.869	0.844	0.919	0.765	0.865	0.839

표 1은 정량적 화질 평가의 결과이다. 표에서 볼 수 있듯이 제안한 기법이 최대 신호 대 잡음비와 구조 유사성 평가에서 기존의 기법보다 평균치 측면에서 더 높은 것을 볼 수 있다. 즉, 제안한 기법이 기존의 방법에 비해, 원본 영상과의 오차가 더 작으면서도 라인이나 텍스처와 같은 영상구조는 더 가깝게 표현했음을 의미한다. 특히 심층 합성곱 신경망 기반의 역방향 하프토닝 기법과 비교했을 때, 최대 신호 대 잡음비는 평균치가 0.318, 구조 유사성은 평균치가 0.007로 증가한 것을 볼 수 있다. 그림 5의 결과 영상을 육안으로 비교했을 때, 선명도나 영상구조 차이를 구별할 수 있으므로 최소식별 차이(Just Noticeable Difference)를 넘어선 것으로 볼 수 있다. 표 1에서 2번과 5번에 대한 성능이 좋지 않은 이유는 심층 합성곱 신경망을 학습할 때 확률적 경사 하강 방식을 사용해서 훈련 배치에 대한 총 손실함수가 국부적으로 최소가 되도록 학습하기 때문에 모든 패치에 대해 화질 개선을 유도할 수는 없는 것으로 분석된다.

## V. 결 론

이 논문에서는 하프톤 영상에서 연속계조 영상을 복구할 때에 라인이나 텍스처와 같은 영상구조 표현 능력을 강화하기 위한 심층 합성곱 신경망 아키텍처를 제안했다. 특히 영상 구조 맵 예측기와 다중 손실 계층을 사용한 다중 스트림 네트워크를 소개했다. 제안한 기법은 영상구조 맵 정보를 활용하여 입력 하프톤 패턴을 연속계조 영상으로 변환할 시 영상 구조를 강화할 수 있도록 학습을 유도하였다.

제안한 기법의 성능 평가를 위해, 기존의 최첨단 사전학습 기반의 역방향 하프토닝 기법과 심층 합성곱 신경망 기반의 역방향 하프토닝 기법과 비교하였다. 총 10장의 테스트 영상에 대해, 제안한 기법이 기존의 기법보다 구조 유사성과 최대 신호 대 잡음비 측면에서 더 우수한 정량적 평가를 달성하였다. 특히 심층 합성곱 신경망 기반의 역방향 하프토닝 기법과 비교했을 때, 최대 신호 대 잡음비의 경우 평균치가 0.318, 구조 유사성의 경우 평균치가 0.007만큼 개선할 수 있었다.

## References

- [1] H. R. Kang, "Digital Color Halftoning", SPIE Press, 1999.
- [2] H. C. Do, B. G. Cho, S. I. Chien, and H. S. Tae, "Improvement of low gray-level linearity using perceived luminance of human visual system in PDP-TV", IEEE Transactions on Consumer Electronics, Vol. 51, No. 1, pp. 204-209, Feb. 2005.
- [3] M. Mese and P. P. Vaidyanathan, "Look-up table (LUT) method for inverse halftoning", IEEE Transactions on Image Processing, Vol. 10, No. 10, pp. 1566-1578, Oct. 2001.
- [4] B. Sun, S. Li, and J. Sun, "Scanned image descreening with image redundancy and adaptive filtering", IEEE Transactions on Image Processing, Vol. 23, No. 8, pp. 3698-3710, Aug. 2014.
- [5] T. Remez, O. Litany, and A. Bronstein, "A picture



- is worth a billion bits: Real-time image reconstruction from dense binary threshold pixels", IEEE International Conference on Computational Photography, Evanston, IL, U.S.A, pp. 1-9, May 2016.
- [6] J. Kopf and D. Lischinsk, "Digital reconstruction of halftoned color comics", ACM Transactions on Graphics, Vol. 31, No. 6, article no. 140, Nov. 2012.
- [7] R. Stevenson, "Inverse halftoning via MAP estimation", IEEE Transactions on Image Processing, Vol. 6, No. 4, pp. 574-583, Apr. 1997.
- [8] A. Foi, V. Katkovnik, K. Egiazarian, and J. Astola, "Inverse halftoning based on the anisotropic LPA-ICI deconvolution", Proc. Int. TICSP Workshop Spectral Meth. Multirate Signal Process., Vienna, Austria, pp. 49-56, Sep. 2004.
- [9] C.-H. Son and H. Choo, "Local learned dictionaries optimized to edge orientation for inverse halftoning", IEEE Transactions on Image Processing, Vol. 23, No. 6, pp. 2542-2556, Jun. 2014.
- [10] Y. Zhang, E. Zhang, W. Chen, Y. Chen, and J. Duan, "Sparsity-based inverse halftoning via semi-coupled multi-dictionary learning and structural clustering", Engineering Applications of Artificial Intelligence, Vol. 72, pp. 43-53, Jun. 2018.
- [11] F. Pelcastre-Jimenez, et al., "An inverse halftoning algorithms based on neural networks and atomic functions", IEEE Latin America Transactions, Vol. 15, No. 3, pp. 488-495, Mar. 2017.
- [12] Y. Choi and C.-H. Son, "Rain removal via deep convolutional neural networks considering orientation and strength of rain streak", Journal of Korean Institute of Information Technology, Vol. 17, No. 1, pp. 85-98, Jan. 2019.
- [13] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation", IEEE Transactions on Signal Processing, Vol. 54, no. 11, pp. 4311-4322, Nov. 2006.
- [14] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition", International Conference on Learning Representations, San Diego, CA, May 2015.
- [15] G. T. Park and C. H. Son, "Image denoising via structure-aware deep convolutional neural networks", Journal of Korean Institute of Information Technology, Vol. 16, No. 11, pp. 85-95, Nov. 2018.
- [16] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection", IEEE Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. 886-893, June 2005.
- [17] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity", IEEE Transactions on Image Processing, Jan. 2004.

### 저자소개

손 창 환 (Chang-Hwan Son)



2002년 2월 : 경북대학교

전자전기공학부(공학사)

2004년 2월 : 경북대학교

전자공학과(공학석사)

2008년 8월 : 경북대학교

전자공학과(공학박사)

2008년 8월 ~ 2009년 12월 :

삼성전자 프린팅사업부 책임연구원

2015년 3월 ~ 2017년 2월 : 캐나다 Ryerson 대학교

포닥연구원

2017년 4월 ~ 현재 : 군산대학교 소프트웨어융합공학과

조교수

관심분야 : 컴퓨터 비전, 영상처리, 기계학습, 딥 러닝