

빅데이터와 딥페이크 기반의 헤어스타일 추천 시스템 구현

김태국*

부경대학교 컴퓨터·인공지능공학부 교수

Implementation of Hair Style Recommendation System Based on Big data and Deepfakes

Tae-Kook Kim*

Professor, School of Computer and Artificial Intelligence Engineering, Pukyong National University

요약 본 논문에서는 빅데이터와 딥페이크 기반의 헤어스타일 추천 시스템 구현에 관해 연구하였다. 제안한 헤어스타일 추천 시스템은 사용자의 사진(이미지)을 바탕으로 얼굴형을 인식한다. 얼굴형은 타원형, 둥근형, 장방형으로 구분하며, 얼굴형에 잘 어울리는 헤어스타일을 딥페이크를 통해 합성하여 동영상으로 제공한다. 헤어스타일은 빅데이터를 바탕으로 최신 트렌드(trend)와 얼굴형에 어울리는 스타일을 적용하여 추천한다. 이미지의 분할 맵과 Motion supervised Co-Part Segmentation 알고리즘으로 같은 카테고리(머리, 얼굴 등)를 가지는 이미지들 간 요소를 합성할 수 있다. 다음으로 헤어스타일이 합성된 이미지와 미리 지정해둔 동영상을 Motion Representations for Articulated Animation 알고리즘에 적용하여 동영상 애니메이션을 생성한다. 제안한 시스템은 가상 피팅 등 전반적인 미용산업에 활용될 수 있을 것으로 기대한다. 향후 연구에서는 거울에 사물인터넷 기능 등을 적용하여 헤어스타일등을 추천해주는 스마트 거울을 연구할 예정이다.

주제어 : 빅데이터, 사물인터넷, 딥페이크, 헤어스타일, 인공지능

Abstract In this paper, we investigated the implementation of a hairstyle recommendation system based on big data and deepfake technology. The proposed hairstyle recommendation system recognizes the facial shapes based on the user's photo (image). Facial shapes are classified into oval, round, and square shapes, and hairstyles that suit each facial shape are synthesized using deepfake technology and provided as videos. Hairstyles are recommended based on big data by applying the latest trends and styles that suit the facial shape. With the image segmentation map and the Motion Supervised Co-Part Segmentation algorithm, it is possible to synthesize elements between images belonging to the same category (such as hair, face, etc.). Next, the synthesized image with the hairstyle and a pre-defined video are applied to the Motion Representations for Articulated Animation algorithm to generate a video animation. The proposed system is expected to be used in various aspects of the beauty industry, including virtual fitting and other related areas. In future research, we plan to study the development of a smart mirror that recommends hairstyles and incorporates features such as Internet of Things (IoT) functionality.

Key Words : Big data, Internet of Things (IoT), Deepfake, Hairstyle, Artificial Intelligence (AI)

*이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(RS-2023-00242528).

*교신저자 : 김태국(king@pknu.ac.kr)

접수일 2023년 5월 18일 수정일 2023년 6월 10일 심사완료일 2023년 6월 14일

1. 서론

사람 겉모습의 이미지를 결정하는 요소는 외모, 헤어스타일(hairstyle), 의류 등 다양한 요소가 있다. 특히 헤어스타일이나 의류는 적은 비용으로 사람의 이미지를 만들 수 있는 요소이다. 남성 헤어스타일의 관심도와 이미지에 관한 연구 결과에 따르면 '헤어스타일은 자신의 이미지를 결정짓는 데 중요한 역할을 한다'라는 항목이 가장 높게 나왔다[1-4]. 조사 결과는 헤어스타일이 개인의 이미지 결정에 중요한 역할을 하는 것을 시사한다. 하지만, 얼굴형과 헤어스타일에 따른 이미지 지각 연구에 따르면 헤어스타일이 정해지는 형태는 보통 헤어디자이너의 경험과 감각으로 정해진다고 한다[5]. 또 다른 헤어스타일을 바꾸는 방안으로, 사진이나 패션 잡지를 참고하여 자신과 어울리는 헤어스타일을 머릿속으로 떠올려 헤어스타일을 선택하기도 한다. 하지만 사진이나 패션 잡지는 2차원 이미지가므로 하나의 각도만 나타낼 수 있고 입체적이지 않다. 따라서 머릿속으로 떠올렸던 스타일과 실제로 바꾼 스타일이 다르게 느껴져 어울리지 않는 문제가 발생하고 매년 변해가는 스타일링 트렌드에 맞춰 헤어스타일을 결정하기에는 시간과 경제적인 어려움이 있다.

주관적인 판단으로 정해지는 헤어스타일에 대한 문제점과 헤어스타일링 결정에 대한 어려움을 해결하기 위해 헤어스타일을 미리 체험하는 제품 및 연구들이 많이 이뤄졌다[6][7]. 그러나, 체험을 위해선 헤어스타일 체험용 스마트 거울이 있는 미용실을 방문하거나, 헤어스타일링 스마트 거울을 구매해야 하는 문제가 있다. 이러한 문제를 해결하기 위해 본 연구에서는 빅데이터와 딥페이크(deepfakes)를 기반으로 사용자의 사진을 업로드하면 사용자의 얼굴형에 잘 어울리는 헤어스타일을 합성하여 동영상으로 제공한다. 따라서 집에서도 간단히 사진만 업로드하면 추천 헤어스타일이 적용된 좌측 얼굴, 우측 얼굴을 확인할 수 있다.

2. 관련 연구

본 연구에서 헤어스타일 추천 영작 제작을 위해 OpenCV Face Recognition, Dlib, Motion supervised Co-Part Segmentation, Motion Representations for Articulated Animation 기술을 활용하였다.

2.1 OpenCV

OpenCV(Open Source Computer Vision)는 실시간 이미지 처리에 중점을 둔 프로그래밍 라이브러리이다. OpenCV는 머신 러닝에서의 영상, 이미지 인식, 검색 등에 이용되고 있으며 TensorFlow, Torch, PyTorch 및 Caffe의 딥러닝 프레임워크를 지원한다[8].

2.2 OpenCV Face Recognition

파이썬에서 무료로 제공되는 Face Recognition 라이브러리는 딥러닝 기반으로 제작된 Dlib의 최첨단 얼굴 인식 기능을 사용하여 구축된 모델이다[9]. Face recognition 라이브러리의 작동 순서는 다음과 같다. 첫째, 사진에서 사람 얼굴에 해당하는 모든 얼굴을 찾는다. 둘째, 사진에 있는 얼굴의 특징(턱의 위치, 윤곽, 눈, 코, 입)을 찾고 도출한다. 셋째, 사진 속 얼굴의 신원을 확인해 누구인지 인식한다.

2.3 Dlib

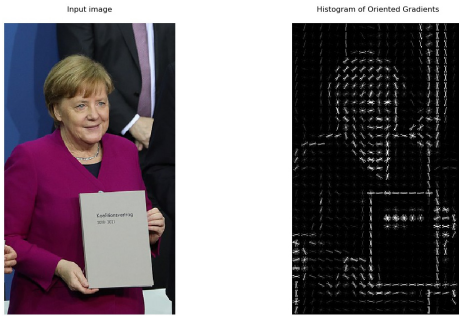
Dlib 라이브러리에는 딥러닝을 포함한 다양한 머신러닝 알고리즘 기능을 포함하고 있으며, 얼굴 인식에 쓰이는 대표적인 오픈소스이다[10]. [Fig. 1]은 Dlib을 사용하여 얼굴에서 68개의 특징점 검출을 나타낸다. 이를 통해 얼굴형을 인식할 수 있다.



[Fig. 1] Dlib 68 feature point detection

2.4 Dlib Face Detector

Dlib 라이브러리의 Face Detector는 HOG(Histogram of Oriented Gradients)와 SVM(Support Vector Machine)을 사용해서 얼굴을 찾는다[11]. HOG는 이미지에서 색상의 변화량을 나타낸 것이다. 색상 변화량을 계산해 이미지로부터 물체의 특징을 찾아낼 수 있다.



[Fig. 2] HOG(Histogram of Oriented Gradients)

SVM은 기계 학습의 분야 중 하나로 패턴 인식, 자료 분석을 위한 지도학습 모델이며, 주로 분류와 회귀 분석을 위해 사용한다. 두 카테고리 중 어느 하나에 속한 데이터의 집합이 주어졌을 때, SVM 알고리즘은 주어진 데이터 집합을 바탕으로 하여 새로운 데이터가 어느 카테고리에 속할지 판단하는 비확률적 선형 분류 모델을 만든다. 만들어진 분류 모델은 데이터가 사상된 공간에서 경계로 표현되는데 SVM 알고리즘은 그 중 가장 큰 폭을 가진 경계를 찾는 알고리즘이다. SVM은 선형 분류기로 한 이미지를 다차원 공간의 한 벡터라고 보면 여러 이미지는 여러 벡터가 된다. 이 벡터들을 구분하는 방법이다. SVM을 이용하여 얼굴이 맞았는지 틀렸는지를 구분한다.

[Fig. 2]는 HOG를 나타낸다. 이미지의 색상만으로는 SVM이 완벽하지 못하다. 따라서 SVM이 잘 작동하기 위해서는 이미지를 HOG를 통해 벡터로 만든다.

2.5 Motion supervised Co-Part Segmentation

공동 부분 분할이란 하나의 이미지 혹은 비디오 프레임에서 영역별로 객체를 분할하는 것을 뜻한다. 하나의 이미지에서 눈, 코, 귀 등의 객체를 영역에 따라 분할하는 기술이다. 대부분의 공동 분할 방법은 많은 양의 레이블이 달린 데이터가 필요한 지도학습으로 학습된다. 이와 달리 Motion supervised Co-Part Segmentation은 분할을 위해 자기 지도학습 모델을 기반으로 레이블이 없는 대량의 비디오에서 모션 정보를 활용해 객체에 대한 분할 정보를 추출한다. 자기 지도학습을 기반으로 동작하는 Motion supervised Co-Part Segmentation에선 하나의 이미지에 대해 하나 혹은 다수의 요소를 바꿔서 비디오를 만드는 기술을 설명한다. Motion supervised Co-Part Segmentation을 적용하면 단일 이미지에 대한 부분 분할 맵을 얻을 수 있고, 분할 맵을 기반으로 원하는 요소가 덮어써진 이미지를 얻을 수 있다[12].

2.6 Motion supervised Co-Part Segmentation

애니메이션화를 위해 Motion Representations for Articulated Animation 모델을 사용한다. Motion Representations for Articulated Animation에선 비지도 학습법으로 하나의 이미지와 목표 비디오의 관절형 객체를 통해 이미지를 애니메이션화하는 모션 표현 모델을 제안한다.

객체를 형상과 포즈로 분리한 후, 로컬 아핀 변환(affine transformation)으로 입력 이미지에 대한 가려진 영역의 합성을 통해 최종적으로 하나의 이미지를 렌더링하며, 각 단계마다 객체의 부분들을 식별해, 목표 비디오로 추적함으로 목표 비디오를 기반으로 한 이미지를 만들어가며 하나의 이미지로 목표 비디오 기반 애니메이션을 얻을 수 있다[13].

3. 얼굴형 인식 구현

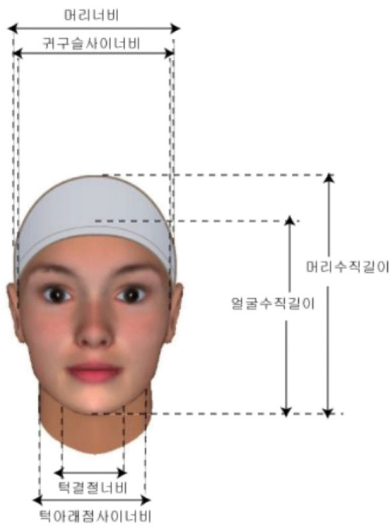
3.1 얼굴형 인식 및 분류

Facial landmark detection을 이용하여 얼굴형을 인식한다. 얼굴형은 얼굴 수직 길이(Face vertical length), 콧볼 사이 너비(귀 구슬 사이 너비, Width between the earlobes), 턱 아래점 사이 너비(Width between lower chin)를 기준으로 분류하였다. <Table 1>은 20대 여성 얼굴 및 이목구비의 유형 분류의 얼굴 유형별 계측 데이터이다[14][15]. Dlib 라이브러리를 통해 추출한 랜드마크 68개의 특징점을 활용하여 얼굴 유형을 인식한다.

<Table 1> Figures by face type (unit: mm)

Item	Oval	Round	Rectangle
Face vertical length	212.5	193.8	219.4
Width between the earlobes	158.2	149.8	175.6
Width between lower chin	125.2	114.6	139.5

얼굴 수직 길이, 콧볼 사이 너비, 턱 아래점 사이의 너비는 [Fig. 3]의 측정기준점에 따른 너비 항목과 길이 항목을 따른다.



[Fig. 3] Width item by measurement base point

타원형(oval face)은 얼굴 너비 대 얼굴 길이의 비율이 약 68.2%로 다른 유형에 비해 얼굴 너비가 얼굴 길이에 비해 가장 좁은 형태로 아래턱 사이 너비 대 얼굴 길이의 비율도 낮아 계란형에 가까운 형태로 이상적인 얼굴 비례를 가진다[14].

둥근형(round face)은 얼굴 너비 대 얼굴 길이의 비율이 약 70.0%로 다른 유형에 비해 턱 결절 너비가 좁은 형태로 얼굴 너비 대 얼굴 길이의 비율이 다른 유형에 비해 넓은 형태의 얼굴 비례를 가진다[14].

장방형(직사각형, rectangular face)은 얼굴 너비 대 얼굴 길이의 비율이 약 75.4%로 유형 중 가장 크게 나타나며 다른 유형에 비해 턱 결절 너비도 큰 유형으로 전체적인 길이가 길어 성숙하고 지적인 얼굴 비례를 가진다 [14].



[Fig. 4] 68 dotted facial landmarks

[Fig 4]는 68개 랜드마크에 대한 인덱스 번호이다. 얼굴 너비는 [Fig. 4]의 랜드마크 배열 0번 x 좌표와 배열 16번 x 좌표 사이의 거리로 측정하였다. 배열 0번과 배열 16번 좌표의 y 좌표의 값은 동일하다.

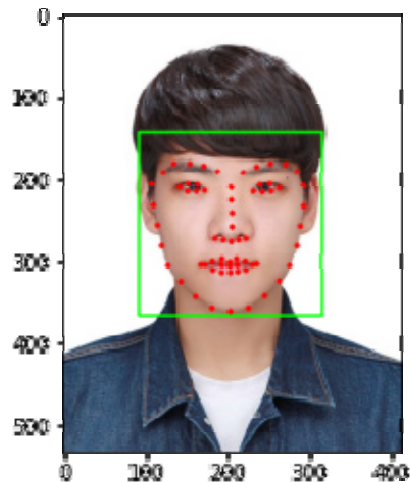
얼굴 길이는 턱 밑 랜드마크 배열 8번 y 좌표와 랜드마크 0~68번 좌표가 아닌, 새로 구한 좌표로 이마 중앙 위 머리 시작점을 나타내는 y 좌표 사이의 거리로 측정하였다. 배열 8번과 머리 시작점의 x 좌표의 값은 같다.

3.2 얼굴형 인식 및 분류 결과

얼굴 너비 대 얼굴 길이 비율이 약 68.2% 이하라면, 타원형으로 분류된다. 얼굴 너비 대 얼굴 길이 비율이 약 68.2%보다 크거나 약 70.0%보다 작다면, 둥근형으로 분류된다. 얼굴 너비 대 얼굴 길이 비율이 약 75.4% 이상이라면, 장방형으로 분류된다.

<Table 2> Face type recognition and classification result

1	rectangles[[92, 142) (315, 365)]]
2	
3	The point above the forehead : (205, 108)
4	The point below the chin : (205, 360)
5	point to the left of your face : (108, 206)
6	point to the right of your face : (293, 206)
7	face ratio : (0.734127)
8	
9	Face Type : a round face



[Fig. 5] Face type recognition and classification result

<Table 2>와 [Fig. 5]는 얼굴형 인식 및 분류 결과를 나타낸다. [Fig. 5]와 같이 사용자의 사진을 통해 68개의

배열로 나타낸 얼굴 랜드마크를 추출한다. 그리고 <Table 2>와 같이 68개의 포인트를 통해 얼굴 너비 대 얼굴 길이의 비율을 계산하여 약 73.4%의 결과값을 얻을 수 있다. 따라서 얼굴형은 둥근형으로 올바르게 분류된 것을 확인할 수 있다.

4. 헤어스타일 적용 및 애니메이션화 구현

4.1 헤어스타일 적용

헤어스타일 추천 및 애니메이션화를 구현하기 위해 목표 동영상에서 물체를 구성하는 부분들을 분할한다. 목표 동영상의 모션은 유지한 상태로 원본 이미지에서 특정 부분의 정체성을 가져와 적용할 수 있다.

모델에 이미지와 영상을 넣기 전, 전처리 과정이 필요하다. 모델에 들어갈 이미지의 사이즈는 (256, 256)이 되어야 하고, 모델에 담기 위해 목표 이미지를 영상으로 재구성 해야한다. 이를 위해 frames 배열에 목표 이미지 하나를 넣음으로 차원을 하나 더 높이는 방법을 사용했다. 앞선 과정에서 매핑된 맞춤형 헤어스타일의 이미지를 source_image로 받아서 모델에 넣기 위해 이미지 사이즈를 조정한다.



[Fig. 6] Face segmentation module results

[Fig. 6]은 얼굴 요소 분할에 사용되는 모듈이 분류해주는 요소에 대한 결과이다. 0번 인덱스부터 18번 인덱스까지 총 19가지의 얼굴 요소를 구분한다. 이 모듈을 사용하여 원본 이미지에 헤어스타일을 덮어씌울 수 있다.



[Fig. 7] Result of applying hairstyle

[Fig. 7]은 맞춤형 헤어스타일 적용 모듈에 대한 결과이다. 왼쪽부터, 목표 헤어스타일, 사용자 이미지, 맞춤형 헤어스타일 적용된 이미지를 나타낸다. 헤어스타일은 빅데이터를 바탕으로 최신 트렌드와 얼굴형에 어울리는 스타일을 적용하여 추천한다.

4.2 애니메이션화

Motion Representations for Articulated Animation 알고리즘을 통해 애니메이션 동영상을 만들 수 있다. 참고 동영상으로부터 모션 정보를 추출하고 원본 이미지의 정체성은 유지한 채로 참고 동영상 모방을 통해 딥페이크 영상을 만들 수 있다.

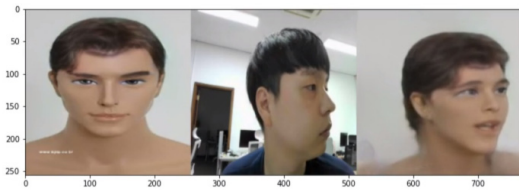
<Table 3>은 모델이 예측한 애니메이션화에 대한 결과를 출력하는 함수의 소스코드이다. 호출 시 넘어오는 매개 변수들을 통해 모델이 동작하며, Motion Representations for Articulated Animation 알고리즘은 이미지 생성 단계에서 원본 이미지, 영역에 대한 정보, 움직임에 대한 정보를 통해 이미지를 생성하는 것을 코드상으로 확인할 수 있다.

<Table 3> Code for Animating

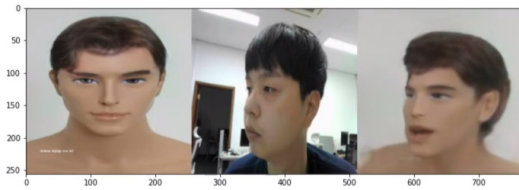
```

1 def make_animation(source_image, driving_video, generator,
2                   region_predictor, avd_network,
3                   animation_mode='standard', cpu=False):
4     with torch.no_grad():
5         predictions = []
6         source = torch.tensor(source_image[np.newaxis]
7                               .astype(np.float32)).permute(0, 3, 1, 2)
8         if not cpu:
9             source = source.cuda()
10        driving = torch.tensor(np.array(driving_video)
11                               [np.newaxis].astype(np.float32))
12                               .permute(0, 4, 1, 2, 3)
13        source_region_params = region_predictor(source)
14        driving_region_params_initial = region_predictor(
15            driving[:, :, 0])
16
17        for frame_idx in tqdm(range(driving.shape[2])):
18            driving_frame = driving[:, :, frame_idx]
19            if not cpu:
20                driving_frame = driving_frame.cuda()
21            driving_region_params = region_predictor(
22                driving_frame)
23            new_region_params = get_animation_region_params
24                (source_region_params,
25                 driving_region_params,
26                 driving_region_params_initial,
27                 avd_network=avd_network,
28                 mode=animation_mode)
29            out = generator(source,
30                            driving_region_params=new_region_params)
31            source_region_params=source_region_params,
32            driving_region_params=new_region_params)
33            predictions.append(np.transpose(out['prediction']
34                                          .data.cpu().numpy(),
35                                          [0, 2, 3, 1])[0])
36        return predictions
    
```

[Fig. 8]과 [Fig. 9]은 헤어스타일 추천 시스템의 최종 결과물인 동영상이다. [Fig. 8]은 애니메이션화가 적용된 좌측 얼굴의 헤어스타일 적용 이미지, [Fig. 9]는 애니메이션화가 적용된 우측 얼굴의 헤어스타일 적용 동영상의 이미지를 나타낸다. 즉, 애니메이션 동영상상 자신의 얼굴형에 맞는 가상의 헤어스타일을 확인할 수 있다.



[Fig. 8] Animated image (left face)



[Fig. 9] Animated image (right face)

5. 결론

본 논문에서는 얼굴형을 인식 및 분류하고, 빅데이터와 딥페이크를 이용한 헤어스타일 추천 기법을 구현하였다. 구현한 기법은 사용자가 입력하는 이미지 한 장을 통해 추출한 랜드마크 68개의 특징점을 활용하여 타원형, 둥근형, 장방형 중 하나로 얼굴형을 인식 및 분류한다. 인식 및 분류 결과에 따라 각자 얼굴형에 맞는 헤어스타일이 합성된 딥페이크 동영상을 생성한다. 헤어스타일은 빅데이터를 바탕으로 최신 트렌드와 얼굴형에 어울리는 스타일을 적용하여 추천한다. 이를 통해 본인의 얼굴형에 어울리는 헤어스타일을 정적인 이미지가 아닌 동영상상을 통해 여러 방향에서도 헤어스타일을 확인할 수 있다.

본 논문에서 제안한 알고리즘을 활용해 전신 이미지에 적용하면 사용자의 전신 이미지와 옷을 합성하여 입체적인 가상 피팅 서비스를 만들 수 있다. 따라서 제안한 기법은 가상 헤어스타일링, 가상 피팅 등의 미용산업에 활용될 수 있을 것으로 기대한다.

REFERENCES

- [1] B.H.Kang, "A study on the interest and image of men's hairstyles," Nambu university, Doctoral thesis, 2022.
- [2] H.M.Kim, J.S.Lee, "A Study on the Male Hair Cut Section Division," Journal of the Korean Society of Cosmetology, Vol.23, No.4, pp.811-823, 2017.
- [3] Y.S.Ku, Y.J.Lee, and T.G.Choo, "A Study on Appearance Management Behavior of Male Consumers(I)," The Society of Fashion and Textile Industry, Vol.12, No.4, pp.459-466, 2010.
- [4] E.J.Im, M.W.Kim, "The Correlation between Recognition of Hair style and Body Image of Men in his Twenties," Journal of the Korean Society of Design Culture, Vol.20, No.2, pp.523-533, 2014.
- [5] J.Y.Son, "A Study on the Perception of Image According to Face and Hair Style," Konkuk university, Doctoral thesis, 2011.
- [6] H.M.Noh, H.W.Joo, Y.S.Moon, K.S.Kong, "Smart Mirror to support Hair Styling," The Journal of the Institute of Internet, Broadcasting and Communication, Vol.20, No.1, pp.127-133, 2020.
- [7] S.B.Kang, S.H.Kwon, Y.H.Kim, S.I.Lee, Y.O.Han, "Smart Mirror for Styling," The Journal of the Institute of Internet, Broadcasting and Communication, Vol.16, No.6, pp.1271-1278, 2021.
- [8] OpenCV, <https://opencv.org/>
- [9] OpenCV Face Recognition, <https://opencv.org/opencv-face-recognition/>
- [10] Dlib C++ Library, <http://dlib.net/>
- [11] Dlib Face Detector, http://dlib.net/face_detector.py.html/
- [12] A.Siarohin, S.Roy, S.Lathuilière, S.Tulyakov, E.Ricci, N.Sebe, "Motion-supervised Co-Part Segmentation", ICPR 2021.
- [13] A.Siarohin, O.J.Woodford, J.Ren, M.Chai, S.Tulyakov, "Motion Representations for Articulated Animation", CVPR 2021.
- [14] J.H.Kim, "An Classification of Face, Ears, Eyes, Mouth and Nose Type of Women in Twenty", The Catholic university of Korea, Doctoral thesis, 2008.
- [15] M.Y.Song, O.L.Park, "A Study on Women's Face Types Classification and Shape Differences," Journal of Fashion Business, Vol.8, No.1, pp.76-90, 2004.

김 태 국(Tae-Kook Kim)

[종신회원]



- 2004년 8월 : 고려대학교
전기전자전파공학부(공학사)
- 2006년 8월 : 고려대학교
메카트로닉스학과(공학석사)
- 2014년 8월 : 고려대학교
모바일솔루션학과(공학박사)
- 2016년 3월 ~ 2022년 2월 : 동
명대학교 AI학부 교수
- 2022년 3월 ~ 현재 : 국립부경대학교 컴퓨터·인공지능공
학부 교수

<관심분야>

사물인터넷(IoT), 콘텐츠 전송 네트워크(CDN), 이동성,
인공지능(AI), 빅데이터(Big data)