

3차원 그래픽스 기반 디지털 휴먼 애니메이션의 동적 리얼리즘 기술 분류 연구

표정·실시간 근육/피부 변형·본(Bone) 구조를 중심으로

Technical Classification of Dynamic Realism in 3D Digital Human Animation

Focusing on Facial Expression, Real-time Muscle/Skin Deformation, and Bone Structure

주 저 자 : 박가현 (Park, Gha Hyun) 광운대학교 미디어커뮤니케이션학부 외부참여연구원

공 동 저 자 : 이학범 (Lee, Hak Bum) 광운대학교 전자재료공학과 박사 과정

교 신 저 자 : 오문석 (Oh, Moon Seok) 광운대학교 미디어커뮤니케이션학부 교수
motion@kw.ac.kr

<https://doi.org/10.46248/kidrs.2026.2.177>

접수일 2026. 05. 18. / 심사완료일 2026. 06. 01. / 게재확정일 2026. 06. 08. / 게재일 2026. 06. 30.

본 논문은 2025학년도 광운대학교 교내학술연구비를 지원받아 수행되었음

Abstract

This study aims to systematically classify the dynamic realism of 3D digital human animation from a technical perspective. The dynamic expression of digital humans is divided into three core axes: facial expression animation, real-time muscle and skin deformation, and bone-based animation structure. Each axis is classified into three levels of realism: hyper-realistic, realistic, and semi-realistic. Key technologies including FACS, blendshapes, anatomical muscle simulation, dual quaternion skinning, ML deformer, and full-body IK are analyzed, and the validity of the classification framework is verified through recent production cases such as Wētā FX's APFS, MetaHuman, and Unreal Engine 5. As a result, a technical classification matrix for dynamic realism is proposed. The study also establishes that visual stylization and technical realism level are independent variables, clarifying that the realism stage should be determined by the actual driving technology rather than the external appearance of the character. This framework is expected to serve as a practical guideline for defining target quality levels and allocating resources efficiently in digital human production across film, games, XR, and metaverse industries.

Keyword

Digital Human(디지털 휴먼), Dynamic Realism(동적 리얼리즘), Facial Expression Animation(표정 애니메이션), Muscle Simulation(근육 시뮬레이션), Bone Structure(본 구조), Rigging(리깅)

요약

본 논문은 3차원 그래픽스 기반 디지털 휴먼 애니메이션의 동적 리얼리즘을 기술적 관점에서 체계적으로 분류한다. 디지털 휴먼의 동적 표현을 표정 애니메이션, 실시간 근육 및 피부 변형, 본(Bone) 기반 애니메이션 구조의 세 가지 핵심 축으로 구분하고, 각 축을 극사실(Hyper-Realistic), 사실적(Realistic), 준사실적(Semi-Realistic)의 세 단계로 분류한다. FACS, 블렌드셰이프, 해부학적 근육 시뮬레이션, 듀얼 쿼터니언 스킨링, ML 디포머, 풀바디 IK 등 주요 기술을 분석하고, Weta FX의 APFS, MetaHuman, Unreal Engine 5 등 최신 사례를 통해 분류 체계의 타당성을 검증한다. 연구 결과로 세 축과 세 단계를 교차한 동적 리얼리즘 기술 분류 매트릭스를 제안하였으며, 외형적 스타일화와 기술적 리얼리즘 단계가 독립 변수임을 밝혀 동적 리얼리즘의 분류 기준이 캐릭터의 외형이 아닌 실제 구동 기술의 수준에 기반해야 함을 확인하였다. 본 연구는 영화, 게임, XR, 메타버스 등 다양한 산업 분야에서 디지털 휴먼 제작 시 목표 품질 수준의 설정과 리소스 배분의 실무적 기준으로 활용될 수 있을 것으로 기대된다.

목차

1. 서론

- 1-1. 연구 배경 및 필요성
- 1-2. 연구의 목적 및 범위
- 1-3. 논문의 구성

2. 이론적 배경 및 선행 연구

- 2-1. 동적 표현 기술의 발전 흐름

- 2-2. 최근 동적 표현 기술의 패러다임 전환

- 2-3. 선행 연구의 한계와 본 연구의 위치

3. 디지털 휴먼 애니메이션의 동적 표현 요소

- 3-1. 동적 표현의 세 축과 상호 의존 구조
- 3-2. 표정 애니메이션의 기술적 구조
- 3-3. 실시간 근육 및 피부 변형의 기술적 구조
- 3-4. 본(Bone) 기반 애니메이션 구조의 기술적 구조

3-5. 세 축의 기술적 상호 의존 관계와 분류의 필요성

4. 동적 리얼리즘의 기술적 분류

4-1. 분류 기준과 리얼리즘 단계의 정의

4-2. 동적 리얼리즘 기술 분류 매트릭스

4-3. 사례 분석: 축 간 조응 관점에서의 검증

4-4. 분류 매트릭스의 실무적 활용과 의의

5. 결론

참고문헌

1. 서론

1-1. 연구의 배경 및 필요성

그동안 디지털 휴먼의 사실성 논의는 주로 모델링, 텍스처, 셰이딩 등 정적 시각 요소에 집중되어 왔다. 그러나 정적 외형이 아무리 정교하더라도 표정이 어색하거나 움직임이 부자연스럽다면 관객은 즉각적인 위화감을 경험한다. Tinwell은 얼굴 표정에서 눈과 입의 비동기화, 미세 표정의 부재가 심각한 불쾌감을 유발함을 실험적으로 입증하였으며,¹⁾ MacDorman과 Ishiguro는 위화감이 정적 유사성보다 동작·다중 모달리티의 일관성 결여에서 비롯됨을 지적하였다.²⁾ 이러한 연구들은 디지털 휴먼의 품질이 동적 표현의 사실성에 결정적으로 의존함을 시사한다.

특히 디지털 휴먼의 동적 표현은 활용되는 기술에 따라 극사실, 사실적, 준사실적 단계로 명확하게 구분될 수 있으며, 각 단계는 고유한 기술 요소와 구현 방법론을 요구한다. 그러나 이러한 동적 리얼리즘의 기술적 구분을 체계적으로 정립한 연구는 부족한 상태이며, 제작 실무에서는 개별 프로젝트의 경험과 직관에 의존하여 동적 요소의 품질 수준을 결정하고 있다.

1-2. 연구의 목적 및 범위

본 연구는 디지털 휴먼 애니메이션의 동적 표현을 표정 애니메이션(Facial Expression Animation), 실시간 근육 및 피부 변형(Real-time Muscle and Skin Deformation), 본(Bone) 기반 애니메이션 구조의 세

축으로 구분하고, 각 축을 극사실·사실적·준사실적의 세 단계로 분류하는 기술적 프레임워크를 수립하는 것을 목적으로 한다. 최신 기술 사례를 통해 분류 체계의 타당성을 검증하며, 궁극적으로 디지털 휴먼의 동적 리얼리즘을 평가할 수 있는 기술 분류 매트릭스를 제안한다. 연구 범위는 3차원 그래픽스 기반의 디지털 휴먼 애니메이션에 한정하며, 음성 합성, 자율 행동, 인터랙션 디자인 등은 제외한다.

1-3. 논문의 구성

본 논문은 총 5장으로 구성된다. 2장에서는 동적 표현 기술의 발전 흐름과 선행 연구를 고찰한다. 3장에서는 세 가지 동적 표현 축의 기술적 구조를 분석한다. 4장에서는 기술 분류 매트릭스를 제시하고 사례를 통해 검증한다. 5장에서는 연구 결과를 종합하고 향후 과제를 제시한다.

2. 이론적 배경 및 선행 연구

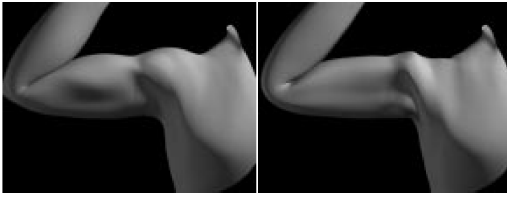
2-1. 동적 표현 기술의 발전 흐름

디지털 휴먼의 동적 표현 기술은 키프레임 기반 수작업 애니메이션에서 출발하여, 해부학적 시뮬레이션과 신경망 기반 변형 근사에 이르기까지 단계적으로 발전해 왔다.

초기 디지털 캐릭터 애니메이션은 선형 블렌드 스킨닝(Linear Blend Skinning, LBS)에 의존하였다. 그러나 LBS는 관절이 크게 회전할 때 부피 손실과 캔디 래퍼 현상을 일으키는 근본적 한계를 가졌고, Lewis 등은 [그림 1]과 같이 포즈 스페이스 디포메이션(Pose Space Deformation, PSD)을 제안하여 특정 포즈에서의 보정 변형을 도입하였다.³⁾

1) Tinwell, A., The Uncanny Valley in Games and Animation, A K Peters/CRC Press, 2014, pp.72-73.

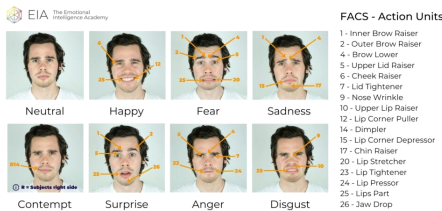
2) MacDorman, K. F., Ishiguro, H., 'The Uncanny Advantage of Using Androids in Cognitive and Social Science Research', Interaction Studies, Vol.7, No.3, 2006, p.308.



[그림 1] Pose Space Deformation 적용 그림⁴⁾

Kavan 등은 듀얼 쿼터니언 스킨닝(Dual Quaternion Skinning, DQS)을 제안하여 회전 변환의 보간 품질을 수학적으로 개선하였으며,⁵⁾ 이는 현재까지 실시간 환경의 표준 기법으로 활용되고 있다.

표정 애니메이션 영역에서는 Ekman과 Friesen이 개발한 FACS(Facial Action Coding System)가 이론적 토대를 제공하였다.⁶⁾ FACS는 [그림 2]와 같이 얼굴 움직임을 44개의 AU(Action Unit)로 분해하여 모든 표정을 해부학적으로 기술할 수 있는 체계로, 디지털 휴먼 표정 시스템의 사실상의 표준이다.



[그림 2] FACS(Facial Action Coding System)의 Action Unit 체계⁷⁾

Lewis 등은 블렌드쉐이프 모델의 이론과 실무를 종합적으로 정리하면서, 이 기법이 영화 산업의 표준이지

3) Lewis, J. P., Cordner, M., Fong, N., 'Pose Space Deformation: A Unified Approach to Shape Interpolation and Skeleton-driven Deformation', SIGGRAPH, 2000, pp.165-172.

4) Lewis, J. P., Cordner, M., Fong, N., Ibid.

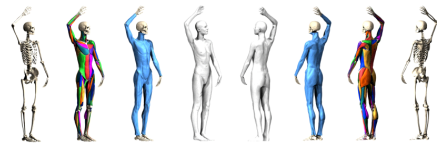
5) Kavan, L., Collins, S., Žára, J., O'Sullivan, C., 'Geometric Skinning with Approximate Dual Quaternion Blending', ACM Transactions on Graphics, Vol.27, No.4, 2008, pp.105:1-105:4.

6) Ekman, P., Friesen, W. V., Facial Action Coding System, Consulting Psychologists Press, 1978.

7) Ekman, P., Friesen, W. V., Ibid.

만 타겟 형상 수와 선형 보간의 한계를 가짐을 지적하였다.⁸⁾

근육 및 피부 변형 영역에서는 Sifakis와 Barbič이 유한요소법(FEM)을 연조직 시뮬레이션의 표준 방법론으로 정립하였으며,⁹⁾ Jacobs 등이 발표한 Ziva Dynamics의 'How to Build a Human'은 [그림 3]과 같이 골격·근육·지방·피부의 다층 구조를 FEM으로 통합 시뮬레이션하는 프로덕션 파이프라인을 제시하였다.¹⁰⁾



[그림 3] 다층 구조 시뮬레이션 프로덕션 파이프라인¹¹⁾

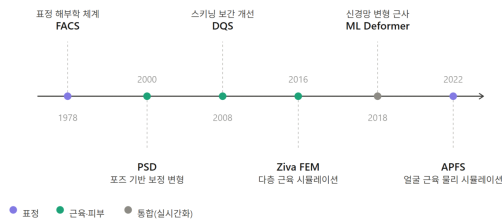
[그림 4]는 이상의 발전 흐름을 연대순으로 정리한 것이다. 1978년 FACS의 등장 이후 약 40여 년간 동적 표현 기술은 표정과 근육·피부 영역에서 각각 독립적으로 발전해 왔으며, 초기의 기하학적 근사(PSD, DQS)에서 물리 기반 시뮬레이션(FEM, Ziva)으로 패러다임이 전환되어 왔음을 확인할 수 있다. 특히 2016년 이후 해부학적 시뮬레이션이 실제 프로덕션에 정착하기 시작하였고, 2018년 이후에는 신경망 기반 변형 근사가 등장하여 오프라인 품질의 실시간 이전이라는 새로운 흐름이 형성되었다. 다음 절에서는 이러한 최근의 패러다임 전환을 구체적으로 살펴본다.

8) Lewis, J. P., Anjyo, K., Rhee, T., Zhang, M., Pighin, F., Deng, Z., 'Practice and Theory of Blendshape Facial Models', Eurographics State of the Art Reports, 2014, pp.199-218.

9) Sifakis, E., Barbič, J., 'FEM Simulation of 3D Deformable Solids: A Practitioner's Guide to Theory, Discretization and Model Reduction', ACM SIGGRAPH 2012 Courses, 2012, pp.20:1-20:50.

10) Jacobs, J., Barbič, J., Edwards, E., Doran, C., van Straten, A., 'How to Build a Human', Proceedings of DigiPro '16, 2016, pp.1-4.

11) Jacobs, J., Barbič, J., Edwards, E., Doran, C., van Straten, A., Ibid.

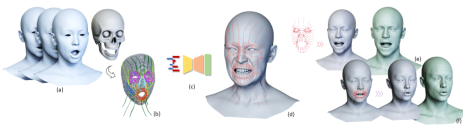


[그림 4] 디지털 휴먼 동적 표현 기술의 발전 흐름

2-2. 최근 기술의 패러다임 전환

2010년대 후반 이후 동적 표현 기술은 두 방향에서 패러다임 전환을 겪고 있다.

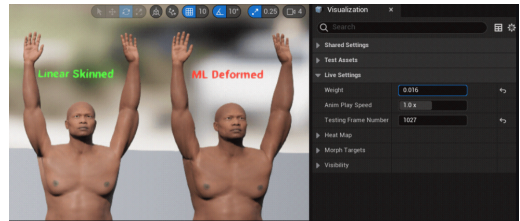
첫째, 해부학 기반 시뮬레이션의 프로덕션 정착이다. Choi 등이 발표한 'Animatomy'는 Wētā FX가 <Avatar: The Way of Water>에서 활용한 APFS(Anatomically Plausible Facial System)의 기반 기술로, 기존 FACS 기반 블렌드셰이프에서 벗어나 [그림 5]와 같이 얼굴 근육 섬유 곡선의 물리적 시뮬레이션을 도입하였다.¹²⁾ 이는 블렌드셰이프의 선형성 한계를 극복한 최초의 대규모 프로덕션 적용 사례로 평가된다.¹³⁾



[그림 5] APFS 근육 섬유 기반 얼굴 변형 시스템¹⁴⁾

둘째, 신경망 기반 변형 근사를 통한 실시간화이다. Bailey 등은 영화급 리그의 비선형 변형을 신경망으로 학습하여 실시간 환경에서 5~17배 속도 향상을 달

성하였다.¹⁵⁾ 이 접근은 Unreal Engine 5의 ML Deformer로 상용화되어,¹⁶⁾ 오프라인 품질의 실시간 이전이 [그림 6]과 같이 가능해졌다.



[그림 6] Unreal Engine 5 ML Deformer 적용 예시¹⁸⁾

2-3. 선행 연구의 한계와 본 연구의 위치

이상의 선행 연구들은 개별 기술 요소에 대한 깊이 있는 통찰을 제공한다. 김릉희는 MetaHuman 리그와 ARKit 리그의 표정 정합성을 비교 분석하여 도구 간 구현 능력의 차이를 밝혔으며,¹⁹⁾ Shakir와 Al-Azza는 얼굴 모델링 및 애니메이션 기법을 종합적으로 분류하였다.²⁰⁾ 그러나 이들은 대체로 개별 기술이나 개별 도구에 초점을 맞추고 있으며, 표정·근육 및 피부 분 구 조라는 동적 표현의 세 핵심 축을 통합적 관점에서 리얼리즘 단계에 따라 체계적으로 분류한 연구는 찾아보기 어렵다.

15) Bailey, S. W., Otte, D., D Lorenzo, P., O'Brien, J. F., 'Fast and Deep Deformation Approximations', ACM Transactions on Graphics, Vol.37, No.4, 2018, pp.119:1-119:12.

16) Bailey, S. W., Omens, D., D Lorenzo, P., O'Brien, J. F., 'Fast and Deep Facial Deformations', ACM Transactions on Graphics, Vol.39, No.4, 2020.

17) Machine Learning Deformer in Unreal Engine, (2026.01.15.)
<https://dev.epicgames.com/documentation/unreal-engine/ml-deformer-framework-in-unreal-engine>

18) Ibid.

19) 김릉희, '사실적인 디지털 휴먼 디자인을 위한 표정 정합성 비교 연구 - 메타휴먼 리그와 ARKit 리그를 중심으로', 커뮤니케이션디자인학연구, Vol.85, 2023, pp.40-43.

20) Shakir, S. D., Al-Azza, A. A., 'Facial Modeling and Animation Approaches: An Overview of the State-of-the-Art', Iraqi Journal for Electrical & Electronic Engineering, Vol.18, No.1, 2022, pp.28-33.

12) Choi, B., Eom, H., Mouscadet, B., Cullingford, S., Ma, K., Gassel, S., Kim, S., Moffat, A., Maier, M., Lewis, J. P., Cho, K., Singh, K., 'Animatomy: an Animator-centric, Anatomically Inspired System for 3D Facial Modeling, Animation and Transfer', SIGGRAPH Asia 2022 Conference Papers, 2022.

13) Letteri, J., 'Technology and Innovation in Avatar: The Way of Water', Proceedings of DigiPro '23, 2023.

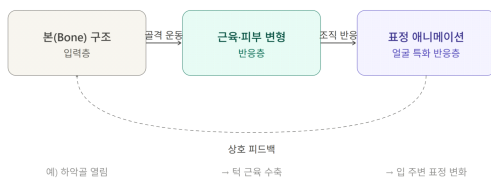
14) Choi, B., Eom, H., Mouscadet, B. 외, Op. cit., 2022.

기존 연구들이 블렌드셰이프, FEM, IK 등 개별 기술의 원리와 성능을 분석하는 데 집중한 반면, 본 연구는 이들 기술을 표정·근육 및 피부·본 구조의 세 축으로 재배치하고, 각 축의 기술 수준을 교차 비교할 수 있는 통합적 분류 체계를 수립한다는 점에서 차별화된다. 이러한 교차 분류 관점은 개별 기술의 우수성만으로는 포착되지 않는 축 간 조응 문제를 드러내며, 이는 디지털 휴먼 제작 실무에서 반복적으로 관찰되는 품질 불균형 현상을 설명하는 틀이 된다.

3. 디지털 휴먼 애니메이션의 동적 표현 요소

3-1. 동적 표현의 세 축과 상호 의존 구조

디지털 휴먼의 동적 표현은 상호 의존적인 기술 레이어가 계층적으로 결합되어 최종 움직임을 생성한다. 본 연구는 이를 기능적 역할에 따라 표정 애니메이션, 실시간 근육 및 피부 변형, 본(Bone) 기반 애니메이션 구조의 세 축으로 [그림 7]과 같이 분류한다.



[그림 7] 동적 표현의 세 축과 상호 의존 구조

이 세 축은 독립적으로 존재하지 않는다. 본 구조가 움직임을 입력을 제공하면 근육·피부 변형이 연조직의 반응을 처리하고, 표정 애니메이션은 얼굴 영역에서 이 과정이 고도로 집약된 형태로 작동한다. 즉 본 구조는 동적 표현의 입력층, 근육·피부 변형은 반응층, 표정 애니메이션은 얼굴 특화 반응층에 해당한다. 예를 들어 턱뼈(하악골)의 열림은 본 구조에 속하지만, 턱 근육의 수축과 볼 살의 변형은 근육·피부 변형에, 그에 따른 입 주변 표정 변화는 표정 애니메이션의 영역에 해당한다. 세 축은 해부학적으로 분리 불가능한 하나의 시스템이지만, 기술적 구현 관점에서는 독립 모듈로 설계·최적화되어야 한다.

3-2. 표정 애니메이션의 기술적 구조

얼굴은 인간의 뇌에 특화된 인식 영역(방추상회 얼굴 영역, 상측두구)이 존재하기 때문에, 미세한 동적 이상도 즉각적으로 감지된다. 따라서 표정 애니메이션은 세 축 중 가장 높은 기술적 정밀도를 요구한다.

표정 애니메이션의 기술적 구조는 [그림 8]과 같이 데이터 입력 → 변형 구동 → 출력 보정의 세 단계로 구성된다.



온사실적 → 사실적 → 극사실적 단계에 따라 각 영역의 기술 복잡도 증가

[그림 8] 표정 애니메이션의 기술적 구조

데이터 입력 단계에서는 수동 키프레임, 성능 기반 캡처(HMC 시스템, 멀티카메라), ARKit 등 모바일 캡처, AI 기반 표정 생성 등이 활용된다. 김릉희의 연구는 MetaHuman 리그와 ARKit 리그의 단독 표정에서는 유의미한 차이가 없으나, 복합 표정에서는 ARKit의 한계가 드러남을 규명하였다.²¹⁾

변형 구동 단계에서는 블렌드셰이프, FACS 기반 리그, 해부학적 근육 시뮬레이션(APFS류)이 활용된다. 블렌드셰이프는 타겟 형상의 가중치 혼합으로 작동하며 제어가 직관적이거나, 타겟 간 선행 보간의 한계로 복잡한 근육 상호작용을 정확히 재현하기 어렵다. APFS는 근육 섬유 곡선의 물리적 시뮬레이션으로 이 한계를 극복한다.²²⁾

출력 보정 단계에서는 코렉티브 블렌드셰이프, 동적 주름 맵(wrinkle map) 전환이 활용되며, 극사실 수준에서는 혈류에 따른 홍조, 모세혈관 변화까지 재현된다.

3-3. 실시간 근육 및 피부 변형의 기술적 구조

이 축은 골격 움직임에 따른 근육의 수축·이완, 피

21) 김릉희, Op. cit., 2023, p.42

22) Choi 외, Op. cit., 2022

부의 슬라이딩, 지방의 흔들림(jiggle) 등을 재현한다. 기술적 구조는 변형 계산의 물리적 충실도를 기준으로 네 수준으로 계층화된다.

첫째, 기하학적 스키닝이다. LBS는 각 정점을 본의 변환 행렬로 가중 선형 합산하여 변형시키며, 빠르지만 관절 대변형 시 부피 손실이 발생한다. DQS는 회전 보간을 수학적으로 개선하나, 근육 웅기나 피부 독립 거동은 여전히 재현하지 못한다.²³⁾

둘째, 보정 기반 변형이다. PSD와 코렉티브 블렌드 셰이프는 기하학적 스키닝의 한계를 수동 반자동으로 보정한다.²⁴⁾ 시각적으로 그럴듯한 근육 반응이 표현되지만 물리 시뮬레이션에 기반하지 않으므로 예상치 못한 포즈에서 시각적 오류가 발생할 수 있다.

셋째, 물리 기반 시뮬레이션이다. Ziva Dynamics의 FEM 기반 파이프라인이 대표적으로,²⁵⁾ 골격 위에 근육·힘줄·지방·피부를 독립 유한요소 메시로 구축하여 물리적으로 정확한 변형을 시뮬레이션한다. 계산 비용이 매우 높아 주로 오프라인 렌더링 환경에서 활용된다.

넷째, 신경망 기반 변형 근사이다. 물리 기반 시뮬레이션의 결과를 신경망으로 학습하여 실시간 환경에서 유사한 품질을 재현한다. UE5의 ML Deformer가 대표적이다.²⁶⁾ 학습 범위를 벗어난 극단적 포즈에서는 아티팩트가 발생할 수 있다.

이 네 수준은 상호 배타적이지 않으며, 실제 제작에서는 기본 변형을 DQS로 처리하되 가시성 높은 부위에 PSD를 추가하고, 클로즈업에서 물리 시뮬레이션을 덧씌우는 하이브리드 방식이 일반적이다.

3-4. 본(Bone) 기반 애니메이션 구조의 기술적 구조

본 구조는 동적 표현이 시작되는 기반층이다. 실제 인간의 골격은 206개의 뼈로 구성되며, 각 관절은 고유한 자유도(DOF)와 운동 범위를 가진다. 본 구조의 기술적 구조는 스켈레톤 복잡도, 운동학적 제어 시스템, 보조 메커니즘으로 구성된다.

스켈레톤 복잡도는 본의 수와 계층 구조의 세밀함을

로 결정된다. 30~50개의 단순 구조에서 300개 이상의 해부학적 완전 재현까지 다양한 수준이 존재한다.

운동학적 제어 시스템은 FK(순운동학)와 IK(역운동학)가 기본이며, 사실적 이상에서는 IK/FK 스위칭, 극 사실에서는 풀바디 IK가 도입된다. Aristidou 등이 정리한 바와 같이 IK 기법은 CCD에서 야코비안 기반, 물리 기반, 학습 기반으로 발전해 왔다.²⁷⁾

보조 메커니즘으로는 트위스트 본(축 회전 분산), 벤디 본(곡선 변형), 제약(constraint) 기반 리깅(본 간 자동 연동)이 활용된다. 극사실 수준에서는 생체역학 기반 제약이 도입되어 실제 인간의 관절 운동 범위가 리깅에 내장된다.

3-5. 세 축의 기술적 상호 의존 관계와 분류 필요성

세 축은 파이프라인 상에서 긴밀하게 연동된다. 본 구조의 복잡도는 근육·피부 변형의 정밀도를 결정하는 전제 조건이다. 예를 들어 건갑골이 독립 본으로 구현되지 않은 상태에서 어깨 근육의 해부학적 시뮬레이션을 수행하면, 입력 자체가 불충분하여 시뮬레이션 결과의 정확성이 저하된다. 한 축만 극사실 수준으로 끌어올리고 나머지 축이 낮은 수준에 머무를 경우, 기술적 투자 대비 효과가 감소할 뿐 아니라 위화감을 오히려 증가시킬 수 있다.

이상의 분석을 종합하면, 동적 리얼리즘의 기술적 분류는 다음의 세 가지 원칙 위에 설계되어야 한다. 첫째, 세 축을 독립적으로 평가하되 축 간 조응 수준을 함께 고려해야 한다. 둘째, 각 축의 기술 수준은 구현에 활용된 실제 기술 스택(스키닝 기법, 리그 규모, 시뮬레이션 방식 등)에 의해 결정되어야 하며, 캐릭터의 외형적 스타일과는 독립적으로 판단되어야 한다. 셋째, 동적 리얼리즘의 실질적 수준은 가장 낮은 축의 수준에 의해 지배되는 경향이 있으므로, 축 간 균형 유지가 제작 전략의 핵심이 된다. 4장에서는 이 세 원칙을 바탕으로 기술 분류 매트릭스를 도출한다.

23) Kavan 외, Op. cit., 2008

24) Lewis 외, Op. cit., 2000

25) Jacobs 외, Op. cit., 2016

26) Epic Games, Op. cit.

27) Aristidou, A., Lasenby, J., Chrysanthou, Y., Shamir, A., 'Inverse Kinematics Techniques in Computer Graphics: A Survey', Computer Graphics Forum, Vol.37, No.6, 2018, pp.35-58.

4. 동적 리얼리즘의 기술적 분류

4-1. 분류 기준과 리얼리즘 단계의 정의

본 연구의 분류 기준은 현재 산업에서 활용되는 대표적 도구 및 프로덕션 파이프라인의 공개된 기술 사양에 기반하여 설정하였다. 사실적 단계의 기준점은 Epic Games의 MetaHuman Creator가 제공하는 표준 리그 사양(FACS 기반 블렌드셰이프 약 80~150개, DQS+PSD 기반 변형, 80~150개 수준의 본 구조)을 참조하였으며, 극사실 단계의 기준점은 Weta FX가 APFS 및 Tissue 시스템을 통해 공개한 프로덕션 기술 보고를 참조하였다. 준사실적 단계는 모바일 플랫폼의 하드웨어 제약(단일 캐릭터당 연산 예산, 실시간 60fps 이상 요구)에 의해 결정되는 기술적 상한을 기준으로 설정하였다. 사례 선정은 각 단계를 대표하는 프로덕션으로서 기술 사양이 공개 문헌이나 공식 문서를 통해 확인 가능한 경우로 한정하였다. 다만 각 단계의 경계 값은 연속적 스펙트럼 위의 참조점으로 가능하며, 절대적 경계가 아님을 밝혀 둔다.

3장에서 정의한 세 기술 축과 리얼리즘 단계를 교차시켜 동적 표현을 체계적으로 분류한다. 본 분류의 핵심 원칙은 축 간 조율으로, 동적 리얼리즘의 실질적 수준은 가장 낮은 축의 수준에 의해 지배되는 경향이 있다.

극사실(Hyper-Realistic): 세 축 모두가 해부학적·물리적 사실성에 기반하여 구축된 수준이다. FEM 기반 시뮬레이션, 해부학적 얼굴 근육 시뮬레이션(APFS류), 풀바디 IK와 생체역학 제약이 통합 적용된다. 영화 VFX의 주연급 디지털 캐릭터에 주로 적용된다.

사실적(Realistic): 실시간 환경의 제약 안에서 인간 움직임을 근사하는 수준이다. DQS, PSD, ML 디포머의 조합, FACS 기반 블렌드셰이프 리그, IK/FK 전환 시스템이 활용된다. MetaHuman과 UE5의 표준 워크플로우가 기준점이며, AAA 게임과 고품질 XR 콘텐츠에 적용된다.

준사실적(Semi-Realistic): 하드웨어·대역폭 제약이 엄격한 환경에서 작동하는 수준이다. LBS 기반 스키닝, 제한된 블렌드셰이프, 단순 스켈레톤으로 구성된다. 모바일 게임, 소셜 VR 아바타, VTuber 플랫폼에 해당한다.

여기서 중요하게 구분되어야 할 점은 외형적 스타일화와 기술적 리얼리즘 단계가 독립 변수라는 것이다. 외형적으로 스타일화된 캐릭터라도 구동 기술(얼굴 리그 규모, 근육 시뮬레이션 수준, 본 복잡도)이 극사실

단계에 해당할 수 있으며, 반대로 외형이 사실적이라도 구동 기술이 준사실적 단계에 머무르는 경우가 존재한다. 본 연구의 분류는 외형이 아닌 실제 구동 기술의 수준에 기반한다.

4-2. 동적 리얼리즘 기술 분류 매트릭스

위의 단계 정의를 세 축의 구체적 기술 명세로 교차시키면 [표 1]~[표 4]의 매트릭스가 도출된다. 이 매트릭스는 디지털 휴먼 제작 실무에서 목표 품질을 정의하고, 각 축의 기술 선택이 다른 축에 미치는 영향을 사전에 검토하는 도구로 활용될 수 있다.

[표 1] 동적 리얼리즘 기술 분류 매트릭스: 표정

구분	준사실적	사실적	극사실적
데이터 입력	수동 키프레임	마커리스 캡처 또는 ARKit 기반	고프레임 HMC + 멀티카메라 퍼포먼스 캡처
변형 구동	단순 블렌드셰이프	FACS 기반 블렌드셰이프 리그	얼굴 근육 시뮬레이션(APFS류) 또는 FACS+물리 하이브리드
블렌드셰이프 규모	15~30개	80~150개	300개 이상 또는 근육 섬유 기반 대체
AU 제어 범위	기본 감정 중심	주요 AU 대부분 지원	전체 AU + 비대칭 독립 제어
출력 보정	정적 텍스처	주름 맵 정적 블렌드	주름 맵 동적 전환 + 혈류 발한 반응
립싱크 정밀도	단순 입 모양	음소 단위 싱크	공명·혀 위치·미세 입술 긴장 포함

[표 2] 동적 리얼리즘 기술 분류 매트릭스: 근육 및 피부

구분	준사실적	사실적	극사실적
변형 구동	LBS	DQS + PSD + ML 디포머	FEM 기반 해부학적 시뮬레이션
레이어 구조	단일 메시 + 가중치	주요 관절 보정 + 부분 레이어	골격·근육·힘줄 지방·피부 다층 독립
피부 슬라이딩	미구현	제한적 구현(주요 부위)	완전한 다층 슬라이딩
실시간성	실시간 60fps+(모바일 포함)	실시간 30~60fps	주로 오프라인 (또는 ML 근사로 실시간화)
대표 기술	표준 LBS, 가중치 기반	UE5 ML Deformer, PSD	Ziva VFX, Weta Tissue,

	기본 스키닝	파이프라인	APFS 하위 시뮬레이션
--	--------	-------	---------------

[표 3] 동적 리얼리즘 기술 분류 매트릭스: 본

구분	준사실적	사실적	극사실적
본 수	30~50개	80~150개 + 보조 조인트	300개 이상 + 생체역학 제약 내장
척추·쇄골	단일 본 또는 2~3분할	분절 그룹, 보조 조인트 근사	추골 개별 제어, 견갑골 활주 구현
운동학 시스템	기본 IK 또는 FK	IK/FK 전환 + 기본 제약	풀바디 IK + 생체역학 제약
보조 메커니즘	최소 또는 없음	트위스트 본 + 기본 제약	트위스트-벤다 제약 통합 운용

[표 4] 동적 리얼리즘 기술 분류 매트릭스: 통합

구분	준사실적	사실적	극사실적
제작 리소스	낮음 (수일~수주)	중간 (수주~수개월)	매우 높음 (수개월~수년)
주요 활용 분야	모바일 게임, 소셜 VR 아바타, VTuber 플랫폼	AAA 게임, 가상 인플루언서, XR	영화 VFX, 하이엔드 시네마틱

4-3. 사례분석

본 절에서는 분류 매트릭스를 실제 디지털 휴먼 사례에 적용하여, 세 축의 기술 수준이 어떻게 조응하고 있는지, 그리고 그 조응이 최종 품질에 어떻게 반영되는지를 분석한다. 본 연구가 주목하는 것은 각 사례의 시각적 완성도 자체가 아니라, 세 축의 수준이 일치하는 사례와 불일치하는 사례가 어떻게 구분되는가이다.

4-3-1. 극사실적 단계: <Avatar: The Way of Water>의 Na'vi 캐릭터

Weta FX가 구현한 Na'vi 캐릭터는 [그림 9]와 같이 세 축 모두 극사실 단계에 도달한 대표적 사례이다.

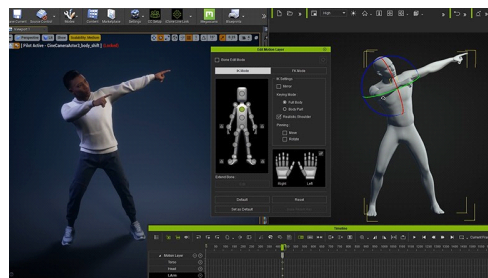


[그림 9] 영화 <Avatar: The Way of Water>(2022)의 Na'vi 캐릭터²⁸⁾

APFS는 얼굴 근육 섬유 곡선을 직접 시뮬레이션하여 FACS 블렌드셰이프의 선형성 한계를 극복하였으며,²⁹⁾ 전신에 FEM 기반 해부학 시뮬레이션이 적용되고,³⁰⁾ 수중 환경의 부력·저항과 연동된 근육 반응까지 물리적으로 재현되었다. 본 구조 역시 Na'vi의 해부학적 설정에 맞춘 커스텀 스켈레톤으로 구축되었다. 세 축이 동일 수준에 있을 때 기술적 투자가 상호 증폭된다는 점이 이 사례의 핵심 시사점이다. 얼굴 근육 시뮬레이션이 정밀하게 작동할 수 있는 이유는 하악골 경추 연결부의 본 구조가 해부학적으로 정확하기 때문이다.

4-3-2. 사실적 단계 사례: MetaHuman 기반 게임·XR 캐릭터

MetaHuman Creator와 UE5 기반 디지털 휴먼은 [그림 10]과 같이 사실적 단계에 해당한다.



[그림 10] MetaHuman Creator 기반 디지털 휴먼³¹⁾

FACS 기반 블렌드셰이프 리그(80~150개), DQS + PSD 기반 변형에 ML Deformer를 결합하며, 표준 MetaHuman 리그로 주요 관절 운동을 재현한다. 세 축이 사실적 단계에서 일관되게 조응하며, 실시간 제약 안에서 가능한 최대한의 자연스러움을 추구한다. 일부

28) Seymour, M., 'Exclusive: Joe Letteri Discusses Weta FX's New Facial Pipeline on Avatar 2', fxguide, (2026.01.15.)
<https://www.fxguide.com/featured/exclusive-joe-letteri-discusses-weta-fxs-new-facial-pipeline-on-avatar-2/>

29) Choi 외, Op. cit., 2022

30) Letteri, Op. cit., 2023

31) Epic Games, 'MetaHuman Documentation', (2026.01.15.)
<https://www.dev.epicgames.com/documentation/en-us/metahuman>

프로덕션이 MetaHuman의 표정 리그(사실적)에 극사실급 근육 시뮬레이션을 결합하려 시도한 경우, 표정 구동의 해상도 부족으로 시뮬레이션의 효과가 충분히 드러나지 않는 현상이 관찰되며, 이는 축 간 조응의 실무적 중요성을 뒷받침한다.

4-3-3. 준사실적 단계 사례: 모바일 게임과 실시간 아바타 플랫폼

모바일 RPG(〈원신〉, 〈붕괴: 스타레일〉 등), VTuber 플랫폼, [그림 11]과 같은 소셜 VR 아바타(VRChat, ZEPETO 등)는 준사실적 단계에 해당한다.



[그림 11] ZEPETO 캐릭터 애니메이션 예시³²⁾

본 구조는 30~60개의 단순화된 스켈레톤, 근육 피부 변형은 LBS 기반 스키닝, 표정은 15~30개 블렌드 셰이프 또는 텍스처 스위치에 의존한다. 관절 대변형의 부피 손실은 물리 보정이 아닌 아트 스타일(혈령한 의상, 셀 셰이딩)로 위장된다.

이러한 기술 수준은 하드웨어 제약에 대응한 의도적 설계의 결과이다. 모바일 디바이스의 전력·발열 제약, 소셜 VR의 다수 아바타 동시 렌더링 요구 등이 단일 캐릭터당 연산 예산을 극도로 제한한다. 동일 스튜디오가 극사실 단계의 PC·콘솔 타이틀과 준사실적 단계의 모바일 타이틀을 병행 제작하는 경우가 흔하며, 이는 동적 리얼리즘 단계가 제작 역량이 아닌 배포 환경의 요구에 따라 결정됨을 보여준다.

4-3-4. 축간 불일치 사례

BioWare의 〈Mass Effect: Andromeda〉(2017)는 정적 요소와 동적 요소, 그리고 동적 축들 내부에서 기술 수준이 불일치할 때 어떤 결과가 발생하는지를 보여주는 대표적 사례이다. 이 타이틀은 캐릭터 모델링과 피부 텍스처에서 상당한 수준의 사실적 정적 품질을

갖추었으나, 표정 애니메이션 시스템은 이에 미치지 못하는 수준으로 구현되었다. 출시 직후 '죽은 눈(dead eyes)'이라는 표현으로 집약되는 광범위한 비판이 제기되었으며,³³⁾ [그림 12]와 같이 대화 장면에서 눈동자의 초점 부재, 감정 전환 시 표정의 급격한 단절, 입과 눈의 비동기화 등이 주요 원인으로 지적되었다.



[그림 12] 〈Mass Effect: Andromeda〉(2017)의 표정 애니메이션³⁴⁾

본 연구의 분류 체계 관점에서 분석하면, 이 사례의 문제는 단순한 기술적 미완성이 아니라 축 간 수준의 불일치에서 비롯된다. 근육·피부 변형과 본 구조는 당시 AAA 타이틀 기준의 사실적 단계에 근접하였으나, 표정 애니메이션은 블렌드셰이프의 수, AU 제어 범위, 출력 보정 측면에서 준사실적 단계에 가까운 수준에 머물렀다. 사실적 수준의 정적 외형과 결합된 준사실적 수준의 표정 구동이 관객의 기대를 심각하게 충족시키지 못한 것이다. 특히 얼굴은 인간의 시선이 가장 집중되는 영역으로, 세 축 중 표정 축의 수준 저하가 전체 품질에 미치는 영향이 가장 크다는 점이 이 사례를 통해 확인된다.

이 사례는 분류 매트릭스가 단순히 목표 수준을 제시하는 데 그치지 않고, 제작 과정에서 축 간 균형을 진단하는 도구로 기능할 수 있음을 시사한다.

4-4. 분류 매트릭스의 실무적 활용과 의의

이상의 분석을 통해, 본 연구가 제안하는 동적 리얼리즘 기술 분류 매트릭스는 다음의 네 가지 실무적 활용 가치를 가진다.

33) Tinwell, Op. cit., 2014, pp.85-92

34) Mass Effect Andromeda Facial Animation, (2025. 01. 13)
https://www.reddit.com/r/Warframe/comments/1i0nbp2/sorry_de_but_these_faces_are_on_a_mass_effect/

32) ZEPETO, (2018. 09. 20)

https://www.youtube.com/watch?v=n0llsGIXT_U

첫째, 프로젝트 초기의 목표 품질 정의. 제작 시작 시점에서 세 축 각각에 대해 목표 단계를 명시함으로써, 이후의 기술 선택·인력 구성·일정 계획이 일관된 기준 아래 이루어질 수 있다. 목표가 '극사실'인지 '사실적'인지 모호한 채로 제작이 진행된다면, 축별로 다른 수준의 결과물이 혼재하여 최종 품질이 저하되는 사례가 빈번하다.

둘째, 축 간 조응의 사전 검토. 특정 축에만 기술적 투자가 집중되어 다른 축이 병목으로 작용하는 상황을 사전에 식별할 수 있다. 예를 들어, 고급 얼굴 근육 시뮬레이션 도입을 검토할 때 본 구조가 이를 지탱할 수 있는지 먼저 확인하는 것이 리소스 낭비를 방지한다.

셋째, 제한된 리소스의 전략적 배분. 모든 축에서 극사실을 추구하는 것이 이상적이지만 현실적으로 불가능한 경우, 본 매트릭스는 어느 축을 낮출 때 전체 품질에 미치는 영향이 가장 적은지를 판단하는 기준을 제공한다. 일반적으로 관객의 시선이 집중되는 표정 축의 수준을 다른 축보다 반 단계 높게 설정하고, 전신 근육·피부 변형은 사실적 단계로 두는 비대칭적 배분이 효율적인 절충안이 된다.

넷째, 기술 발전에 따른 지속적 업데이트 기반. 매트릭스의 각 셀에 기재된 기술 명세는 현재 시점의 산업 표준을 반영한 것이며, 새로운 기술이 등장하면 해당 셀이 업데이트될 수 있는 유연한 구조이다. 예를 들어, ML Deformer 계열 기술이 더욱 성숙하여 FEM 품질의 변형을 완전히 실시간화할 수 있게 되면, 극사실 단계의 '주로 오프라인' 조건이 '실시간 가능'으로 이전되고, 사실적 단계는 현재의 준사실적 단계로 하향 이동할 수 있다. 이러한 동적 업데이트 가능성은 본 매트릭스가 고정된 규범이 아닌 살아있는 프레임워크로 기능할 수 있게 한다.

5. 결론

본 연구는 디지털 휴먼 애니메이션의 동적 사실성을 표정, 근육 및 피부 변형, 본 구조의 세 축과 극사실·사실적·준사실적의 세 단계로 교차 분류하는 기술적 프레임워크를 제안하였다. <Avatar: The Way of Water>, MetaHuman 기반 캐릭터, 모바일 게임·VTuber 아바타 등의 사례 분석을 통해 분류 체계의 타당성을 검증하였으며, 축 간 기술 수준의 조응이 동적 사실성의 결정 요인임을 확인하였다.

본 연구의 기여는 다음과 같다. 첫째, 동적 사실성을 세 축의 조응으로 성립하는 통합적 영역으로 정립하였다. 둘째, 기술 축과 리얼리즘 단계의 교차 구조를 갖는 분류 매트릭스를 수립하여 프로젝트 간 품질 수준을 동일 기준으로 비교할 수 있게 하였다. 셋째, 외형적 스타일화와 기술적 리얼리즘 단계가 독립 변수임을 밝혀 동적 리얼리즘 논의에서 자주 발생하는 범주 혼동을 해소하였다.

본 연구는 다음의 한계를 가진다. 첫째, 매트릭스의 기술 명세는 현재 시점의 산업 표준을 반영한 것으로 지속적 갱신이 필요하다. 둘째, 각 단계의 경계값(블렌드쉐이프 수, 본 수 등)은 대표적 도구와 프로덕션의 공개 기술 사양을 참조하여 설정하였으나, 연속적 스펙트럼 위의 참조점으로 기능하며 절대적 경계가 아니다. 향후 전문가 델파이 조사나 사례별 정량 비교를 통한 경계 기준의 객관화가 필요하다. 셋째, 사용자 지각 기반의 실증적 검증은 수행하지 않았다. 넷째, 음성·자율 행동 등 비시각적 동적 요소는 다루지 않았다.

향후 연구에서는 각 단계의 디지털 휴먼에 대한 사용자 지각 실험을 통해 매트릭스의 타당성을 정량적으로 검증하고, AI 기반 디지털 휴먼 생성 기술(NeRF, 가우시안 스피클레팅, 대규모 모션 생성 모델 등)이 본 분류 체계에 어떻게 편입될 수 있는지 연구할 계획이다. 나아가 시각적 동적 사실성과 음성·행동 축을 통합하는 다중 모달리티 분류 프레임워크의 개발이 요구된다. 또한 본 연구의 기술 분류 체계를 디자인 의사결정 이론과 연계하여, 프로젝트 초기의 품질 목표 설정에서 활용 가능한 디자인 프레임워크로 확장하는 연구도 요구된다.

참고문헌

- Ekman, P., Friesen, W. V., Facial Action Coding System, Consulting Psychologists Press, 1978.
- Tinwell, A., The Uncanny Valley in Games and Animation, A K Peters/CRC Press, 2014.

3. 김릉희, '사실적인 디지털 휴먼 디자인을 위한 표정 정합성 비교 연구 — 메타휴먼 리그와 ARKit 리그를 중심으로', 커뮤니케이션디자인학회연구, Vol.85, 2023.
4. Aristidou, A., Lasenby, J., Chrysanthou, Y., Shamir, A., 'Inverse Kinematics Techniques in Computer Graphics: A Survey', Computer Graphics Forum, Vol.37, No.6, 2018.
5. Bailey, S. W., Otte, D., Dilozenzo, P., O'Brien, J. F., 'Fast and Deep Deformation Approximations', ACM Transactions on Graphics, Vol.37, No.4, 2018.
6. Bailey, S. W., Omens, D., Dilozenzo, P., O'Brien, J. F., 'Fast and Deep Facial Deformations', ACM Transactions on Graphics, Vol.39, No.4, 2020.
7. Choi, B., Eom, H., Mouscadet, B., Cullingford, S., Ma, K., Gassel, S., Kim, S., Moffat, A., Maier, M., Lewis, J. P., Cho, K., Singh, K., 'Animatomy: an Animator-centric, Anatomically Inspired System for 3D Facial Modeling, Animation and Transfer', SIGGRAPH Asia 2022 Conference Papers, 2022.
8. Jacobs, J., Barbič, J., Edwards, E., Doran, C., van Straten, A., 'How to Build a Human', Proceedings of DigiPro '16, 2016.
9. Kavan, L., Collins, S., Žára, J., O'Sullivan, C., 'Geometric Skinning with Approximate Dual Quaternion Blending', ACM Transactions on Graphics, Vol.27, No.4, 2008.
10. Letteri, J., 'Technology and Innovation in Avatar: The Way of Water', Proceedings of DigiPro '23, 2023.
11. Lewis, J. P., Anjyo, K., Rhee, T., Zhang, M., Pighin, F., Deng, Z., 'Practice and Theory of Blendshape Facial Models', Eurographics State of the Art Reports, 2014.
12. Lewis, J. P., Cordner, M., Fong, N., 'Pose Space Deformation: A Unified Approach to Shape Interpolation and Skeleton-driven Deformation', SIGGRAPH, 2000.
13. MacDorman, K. F., Ishiguro, H., 'The Uncanny Advantage of Using Androids in Cognitive and Social Science Research', Interaction Studies, Vol.7, No.3, 2006.
14. Shakir, S. D., Al-Azza, A. A., 'Facial Modeling and Animation Approaches: An Overview of the State-of-the-Art', Iraqi Journal for Electrical & Electronic Engineering, Vol.18, No.1, 2022.
15. Sifakis, E., Barbič, J., 'FEM Simulation of 3D Deformable Solids: A Practitioner's Guide to Theory, Discretization and Model Reduction', ACM SIGGRAPH 2012 Courses, 2012.
16. Seymour, M., 'Exclusive: Joe Letteri Discusses Wētā FX's New
17. fxguide.com
18. dev.epicgames.com
19. dev.epicgames.com
20. youtube.com
21. reddit.com