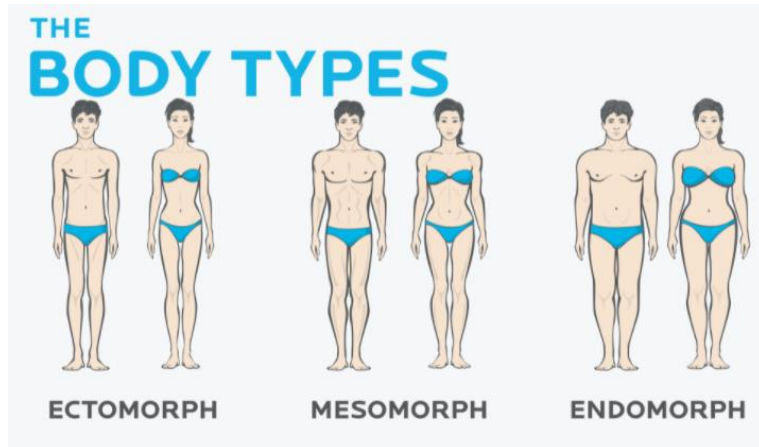


Novel Anthropometry Based on 3D-Bodyscans Applied to a Large Population Based Cohort

기존의 연구들에서는 3가지로의 체형분류 정도 시행된 바 있음.



Endomorphic(고 지방)

Ectomorphic(마름)

Endo-mesomorphic(균형)

BMI, WHtR(Waist to Height Ratio), WHR(Waist to Hip Ratio)를 사용하는 분류법들로는 "Obesity Mortality Paradox" 에서 벗어나지 못하기 때문에 다른 기준을 찾아보고자 실험.

**"Obesity Mortality Paradox" : 비만 역설. 비만이 노인이나 특정 만성 질환이 있는 사람들과 같은 특정 집단의 사람들에게 오히려 도움이 될 수 있다는 의학적 가설.

주 목적

얼마나 많은, 어떤 종류의 신체 치수가 "Body Shape"와 연관이 있을까?

특징적인 'Body type'이 존재하는가?

어떻게 그러한 'Body type'들을 강건(robust)하게 결정 및 정의 할 것인가?

사용된 데이터

Leipzig Research Center for Civilization Diseases 에서 제공된 데이터.

"Vitus Smart XXL" 이라는 3D Laser Scanner 를 이용하여 약 10000명의 독일 Leipzig 도시의 성인 (40-79세) 남녀들을 측정하였고 이들은 주로 중-유럽 인종.

각 참가자 별 140개의 measure 존재.

-97개의 길이데이터, 36개의 둘레, 2개의 각도, 무게, BMI, WHR, WHtr, "Body Shape Index(ABSI)" 존재

-140 X 9892 Matrix 생성

**ABSI : A Body Shape Index 라고 불리며 주어진 인체 신장, 질량 및 허리둘레(WC)의 건강 영향을 평가하기 위한 지표. ABSI는 신장, 체중, BMI와 아주 적은 상관관계를 갖고 있어 사망률을 예측하는 데 있어 다른 인체측정학 변수들에 비해 독립적인 경향이 크다.

$$ABSI = \frac{WC}{BMI^{\frac{2}{3}} \times Height^{\frac{1}{2}}}.$$

Data Preprocessing

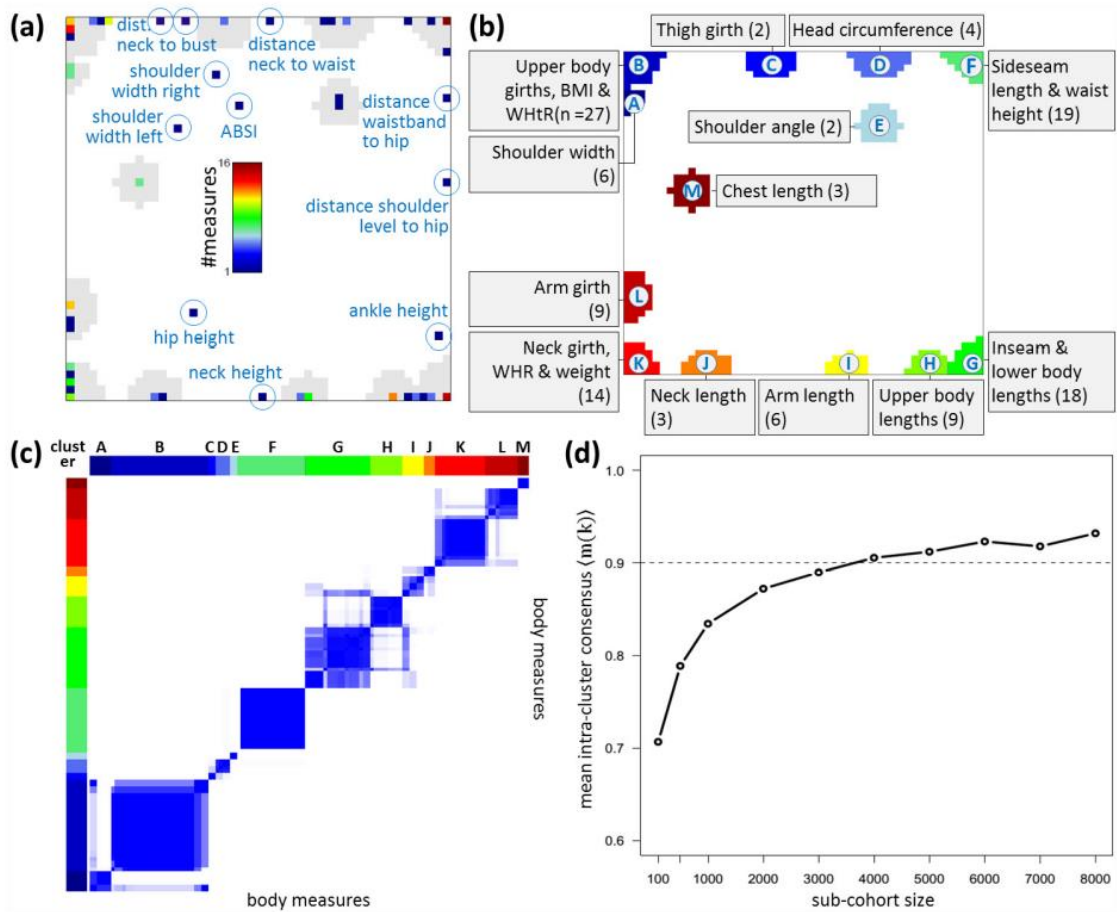
1. 50%의 넘는 결측값을 갖는 46명의 데이터 삭제 -> 5%가 넘는 결측값을 갖는 measure 6개 삭제 -> 그 외 결측값을 보유하고 있는 1347명의 데이터 삭제 -> 134 X 8499 Matrix
2. 모든 measure들을 Height(키)로 나눠서 정규화 (체형이 신장에 따라 선형으로 비례한다고 가정)
3. 각 measure에 대해 Z-Normalization 시행 -> 서로 다른 measure들을 비교할 수 있는 척도로 사용하기 위해

SOM clustering을 이용하여 "Feature Map" 생성

전처리가 끝난 134 X 8499의 데이터에 SOM 알고리즘 적용 -> SOM size : 50 X 50

분류된 cluster들 간 inter-cluster 거리가 최대가 되는 값으로 분류 -> 13개의 군집 생성 -> 13개의 meta-measure로 사용

각 meta-measure은 2~27개의 measure들이 합쳐진 값 (각 군집의 측정값에 대한 평균값으로 대체) -> 각 meta-measure들에 대해 Z-Normalization 적용

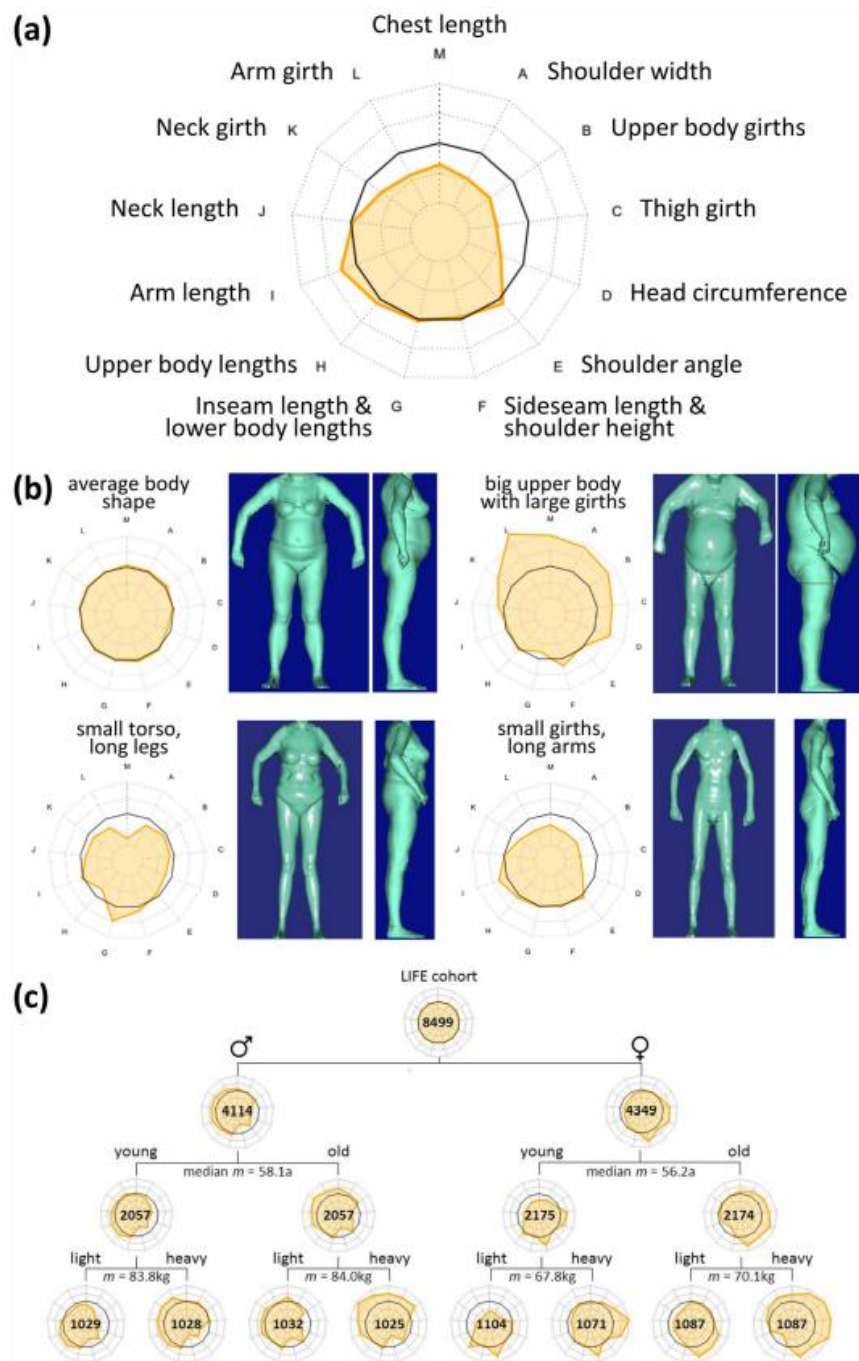


Feature map 생성 결과

W45	fs																		
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
	ID	Age	Gender	BMI	Body type	Meta-measure													
						A	B	C	D	E	F	G	H	I	J	K	L	M	
1	56AAE9332C	18 male			18 M1	-1.68612048 -1.97677614 -2.38261564 -1.44030279 -0.38132264 -0.54057937 0.83959636 -0.24752639 0.54963997 -0.08044135 -1.05953608 -1.28186417 -2.45533691													
2	18A2E17B59	19 male			20 B1	-0.6688075 -1.50143025 -1.17546191 -0.82665066 -0.90036867 0.04990012 0.11317141 0.57421513 -0.61408547 -0.96677431 -0.75353421 -0.64852458 -0.90108509													
3	D02877DF94	20 male			23 M1	-1.13685368 -0.98696203 -0.23642072 -0.71056719 0.05363031 -1.47708266 -0.71947453 0.69034756 0.95766541 0.6503583 0.20199967 -0.4225991 -0.95081014													
4	D616C4F174	20 male			20 M1	-0.79327254 -1.1985017 -1.21561017 -0.41702257 0.83807586 -1.1860855 -0.66950022 0.18218064 0.63361068 0.38073409 0.11418671 -1.1280159 -1.09112278													
5	CFF5669D7D	20 male			23 M1	-1.15625244 -1.02537592 -0.65060285 -0.75507298 0.6748655 0.0489607 0.17481949 0.09283788 1.15923118 0.62695129 -0.70152485 -0.56262415 -0.7158186													
6	C5967461B1	20 male			21 M1	-1.94139485 -1.70733544 -1.2637956 -1.08891219 -0.45282166 -0.46765293 0.36414397 1.07337735 0.31548734 -0.6034215 -1.09126793 -1.22755631 -1.33921842													
7	D0583FA68D	21 male			30 M4	0.13102267 0.14819184 -1.14781119 -0.91210022 -0.68721248 -2.89116507 -1.23992226 0.43595911 -0.71035792 -0.68015533 1.38070146 0.51955762 0.36610868													
8	C1A89851CD	21 male			30 NA	0.20578364 -0.07678685 1.17656439 -0.14310248 0.42100063 -1.10472298 -0.2585312 -0.14366389 0.70065799 0.50625879 -0.04934263 0.47082929 -0.95600511													
9	7CC88AA9C1	21 male			26 M1	-0.52875345 -0.64742554 0.52498402 -0.31719856 0.20688764 0.03219987 0.62689119 0.09860085 0.55174734 0.23209256 -0.31969163 -0.47038438 -0.04266438													
10	9D534D1D23	21 male			24 M1	-0.47406495 -1.15721025 -0.77566474 -1.65748531 -0.93387325 -1.06153584 0.89099205 0.48315491 0.2867458 -0.06057554 -0.72811045 -0.30667063 -0.49731362													

생성된 meta-measure들의 예시

Bodygram 을 통한 시각화



검은선은 $Z=0$ 인 지점을 나타냄(평균)

(C)에서는 성별, 나이, 무게에 따라 분류해 본 결과 -> 성별, 나이, 무게가 체형에 다양한 영향일 끼친다. 이전의 1 차원적인 측정에서는 체형을 설명하기에는 부족한 점이 많았음.

SOM clustering 을 이용한 'Body Map' 생성

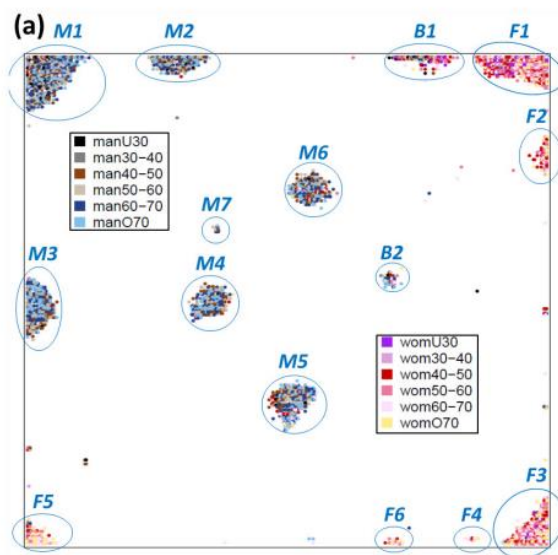
비슷한 체형을 가진 사람들을 분류하기 위해 시행 -> SOM size : 130 X 130

13개의 정규화된 meta-measure들을 이용

Robust 를 위해 한 군집에 최소한 85개의 참가자 데이터(총 집단의 1%)가 들어있도록 설정

총 15개의 군집 생성

100-fold bootstrapped clustering을 이용하여 검증하였다.

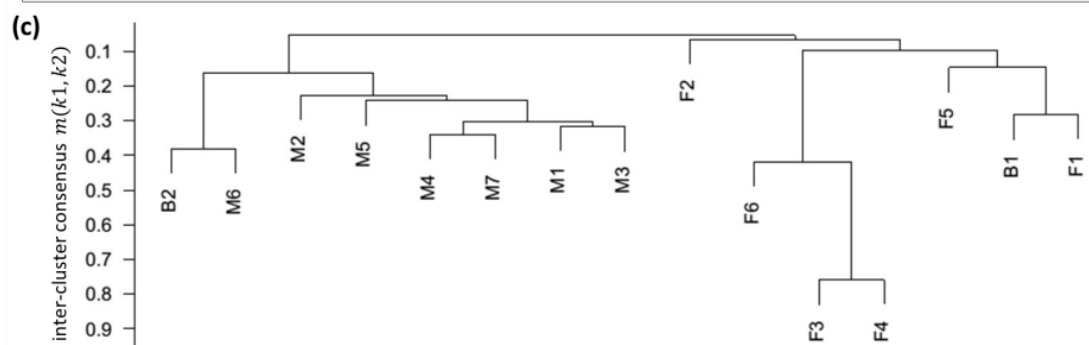


총 15개의 군집 생성 결과

B1, B2: gender unspecific body types

F1 ... F6: female specific body types

M1 ... M7: male specific body types



Hierarchical Cluster Dendrogram

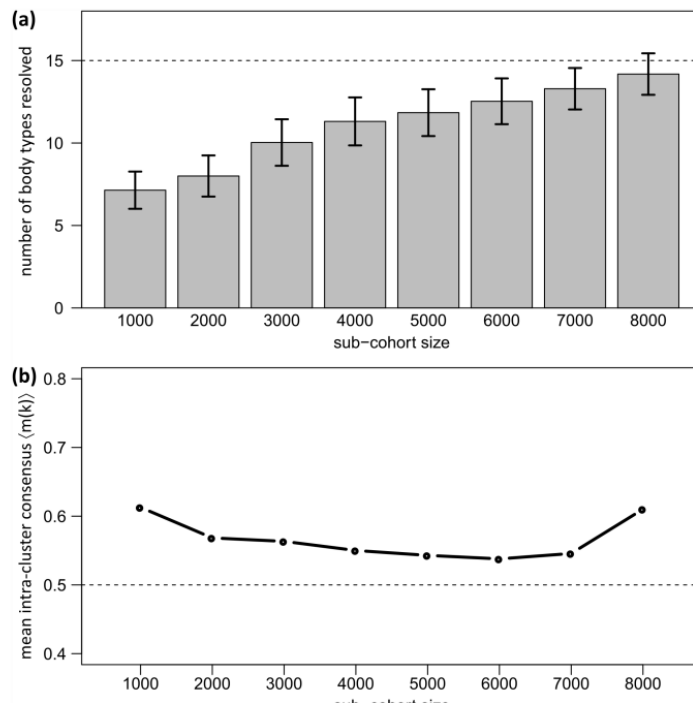
-일반적인 상식대로 크게 남성(M) 과 여성(F)의 기준으로 크게 나뉘고, 그 안에 있는 분류(1,2...)의 비율은 거의 동일하게 나타난다.

Ex) 남성과 여성으로 나뉘었을 때, 동일한 키 비율(%)에서 나타나는 체형의 수가 비슷하다.

Resolution of body types depends on cohort size

-“Stability Analyses”의 관점에서 희귀한(적은 수의) 특정한 “Body type”를 발견함에 있어서 참가자(표본)의 수가 많을수록 다양한 “Body type”이 발견됨을 알 수 있었다.

-Ex) 만약 이 연구가 8499명보다 많은 데이터가 있었다면 15개 이상의 “Body type”를 발견할 수 있었을 것. 추후 데이터가 많아지고 다시 시행할 연구에서는 더 많은 특징을 찾을 수 있을 것이다.



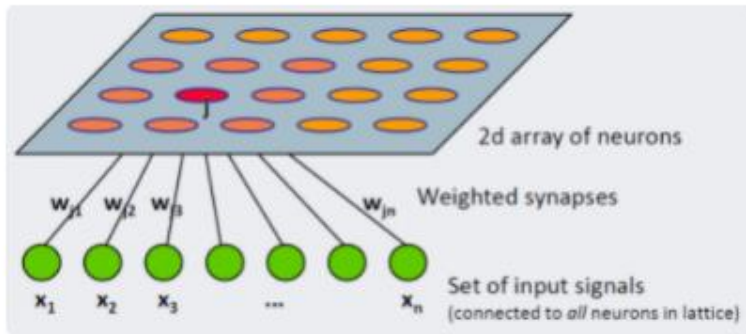
결론

8499명의 데이터를 이용하여 13개의 meta-measure를 얻어내고, 이를 통해 15개의 구체적이고 robust한 body type를 발견할 수 있었다. 이전의 3가지 분류에 비해 구체적이고, 다양한 체형을 분석할 수 있었으며 MRI-Scan 등 더 자세한 정보들을 통해 세부적 분류가 더 가능할 것으로 보인다. 이전의 연구들처럼 일차원적인(몸무게, BMI만, 키만) 고려한 것 보다, 3D 데이터를 이용한 더욱 자세한 체형분석의 가능성을 발견하였다.

연구의 한계

주어진 데이터로는 한정된 나이대(40-79)와 특정한 인종(중-유럽)에 한해서만 분석이 가능하였다= 다양한 나이대와 인종을 고려하지 못하였다.

**SOM(Self-Organizing Maps) clustering: 차원축소와 군집화를 동시에 실행해주는 알고리즘



- Unsupervised learning (인공신경망)
- 차원축소가 되므로 고차원 데이터에 유용 (차원축소에 있어 PCA와 비슷한 원리로 적용?)
- 주로 유클리디안 거리를 사용하여 분류
- 저차원 격자에서의 유사도는 고차원 입력공간에서의 유사도를 최대한 보존하도록 학습
- 처음 각 격자벡터를 랜덤으로 초기화 한 후, 학습데이터(x)를 하나씩 추가해가며 격자벡터의 위치를 x의 위치와 비슷해 지도록 업데이트 한다(instance-based learning)
- 주어진 x와 그 위치가 가장 가까운 "Winning Node"가 가장 많이 업데이트되고, 먼 곳에 위치한 격자벡터는 거의 업데이트 되지 않음.

