

하이퍼그래프 모델을 이용한 fMRI Brain Network의 허브 모듈 분석

김준식^{1,0}, 임병권¹, 김은솔¹, 양진산¹, 장병탁^{1,2,3}

¹바이오지능 연구실, 서울대학교 전기컴퓨터공학부

²서울대학교 인지과학협동과정

³서울대학교 뇌과학협동과정

{jskim, bklim, eskim, jsyang, btzhang}@bi.snu.ac.kr

Searching for the Hub Module of fMRI

Data with the Hypergraph Model

Joon Shik Kim^{1,0}, Byoung-Kwon Lim¹, Eun-Sol Kim¹, Jinsan Yang¹,

Byoung-Tak Zhang^{1,2,3}

¹Biointelligence Lab, School of Computer Science and Engineering, Seoul National University

²Interdisciplinary Program in Cognitive Science, Seoul National University

³Interdisciplinary Program in Brain Science, Seoul National University

요 약

본 논문에서는 하이퍼그래프의 고유벡터를 척도로 하여 fMRI기반 Brain Network를 분석하여 중요한 허브노드를 찾는 방법론을 제시한다. 이 방법을 비디오게임을 수행하면서 촬영한 기능적 자기뇌영상(fMRI) 데이터인 PBAIC 2007 데이터셋에 대하여 그 유용성을 검증하였다. 이 데이터는 각 20분씩 세 세션을 촬영한 것이며 처음 두 세션에는 13가지의 감정 항목의 평가치가 각 스캔마다 주어진다. 한 피험자의 첫 번째 세션 데이터로부터 13가지 감정 항목에 대하여 상관관계가 높은 각각의 복셀(voxel)들을 추출하였다. 이 13가지의 복셀들의 집합들을 각각 하이퍼에지로 보고 하이퍼그래프를 구성하였다. 하이퍼그래프로부터 인접 행렬(adjacency matrix)를 구성한 후 고유치(eigenvalue)와 고유벡터(eigenvector)를 구하였다. 여기서 고유치가 가장 큰 고유벡터의 원소들은 각 복셀들의 중앙성(centrality), 즉 중요성을 나타내며 이로부터 감정과 관련된 중요한 허브 복셀들과 그들의 국소적 집합인 모듈을 찾았다. 모듈들은 감정 및 작업기억(working memory)과 관련된 뇌 영역들의 클러스터(cluster)로 추정된다.

1. 서 론

기능적 뇌영상(fMRI)의 연구의 중요한 분야의 하나는 측정된 볼드(blood oxygen level dependent, BOLD) 신호를 바탕으로 피험자의 심리상태를 추측하는 연구이다[1]. 이때 볼드 시그널은 헤모글로빈과 결합된 산소의 농도를 나타내며 물분자의 수소, 즉 양성자의 스핀의 양자역학적 세기를 제어서 측정한다[2]. 공급되는 산소, 즉 헤모글로빈과 결합된 산소가 혈중에 많을수록 전자기적 균일성이 높아져서 교란되지 않은 높은 스핀의 세기가 관측된다[3].

이 볼드 신호를 대략 2초 정도의 시간간격을 두고 x, y, 그리고 z 축으로 64 x 64 x 34 차원을 가지는 복셀들에서 측정한다. 보통, 머리 움직임 보정, 표준 템플릿으로의 변환, 그리고 스무딩(smoothing) 등의 전처리(preprocessing)의 분석 과정을 거친다. 우리는 피츠버그대학에서 개최하는 심리 상태 추측 컨테스트인 PBAIC(Pittsburgh Brain Activity Interpretation Competition)의 2007 data에 하이퍼그래프 분석법을 적용해 보았다.

하이퍼그래프는 여러 개의 노드(node)를 포함하는 하이퍼에지들로 이루어진다[4]. 기존의 그래프 이론이

두 노드간의 연결성에 관심을 가졌다면 하이퍼그래프는 높은 차원(order)의 상관관계를 중요시하는 연구에서 쓰여진다[5][6]. 더 나아가서, 하이퍼네트워크는 클릭(clique)으로 하이퍼에지들을 만들고 주어진 데이터와 가까운 하이퍼에지의 수를 증폭함으로써 강도(weight)를 증가시키는 연산을 수행한다[7]. 하이퍼네트워크를 가지고 대뇌피질두께 데이터로 남녀를 분류하는 연구가 있었다[8].

우리는 칼리브해의 군도에서 일어난 전쟁데이터를 바탕으로 다중적 중앙도(multidimensional centrality)를 적용한 논문에서 기술적 아이디어를 얻었다[6]. 이 논문에서는 시간과 사람과 장소에 대한 하이퍼그래프를 행렬(E)로 표시하고 이로부터 $E^T E$ 를 계산하여 인접 행렬을 구한다. 그 뒤 인접행렬을 대각화(diagonalization)하여 고유치와 고유 벡터를 구한다. 여기서 가장 큰 고유치의 고유벡터를 선택하며 이 고유벡터의 원소의 절대값이 중앙성(centrality)를 의미한다. 중앙성이 큰 값을 가지는 노드일수록 중요한 노드임을 알 수 있다.

13가지 감정 상태는 얼굴, 개, 과일, 무기, 건물 안, 건물 밖, 속도 등의 항목이며 1.75초마다의 스캔에 13가지 평가치(rating)이 주어진다. 처음 두 세션에는 평가치와 볼드 신호가 같이 주어지고 세번째 세션에서는 볼드 신호만 주어진다. 컨테스트는 처음 두 세션에서 다중 회귀(multiple regression)를 수행하고 세번째 세션에서의 볼드 신호로부터 13가지 항목의 감정 관련 값들을 추측하는 것이다.

우리는 PBAIC 2007 홈페이지에서 제공하는 매트랩 기반의 분석틀인 MVPA(MultiVariate Pattern Analysis)를 이용하였다. 즉, 한 항목과 상관관계가 높은 복셀들을 추출하고(최대 상관관계의 75% 이상 값을 가지는 복셀들) 이들을 하이퍼에지로 묶었다. 결국, 13개의 하이퍼에지로 이루어지는 하이퍼그래프가 생긴다. 이로부터 인접 행렬을 구하고 고유치와 고유벡터를 구하여 중앙도를 하이퍼그래프 노드들에 대하여 부여한다. 중앙도가 높은 복셀들을 찾은후 가까운 위치를 점유하는 복셀들의 집합들을 찾아 허브 모듈로 정의한다.

찾아진 중요한 복셀들이 정확하게 뇌의 어떤 영역에 해당하며 그 인지과학적인 해석이 어떠한지는 향후 연구(future work)으로 남겨둔다. 우리의 연구는 하이퍼그래프의 중앙도를 계산하여 허브 모듈을 찾는 방법론을 제시한다는데서 그 의미를 찾을 수 있다.

2. 연구 방법

PBAIC 홈페이지(<http://pbc.lrdc.pitt.edu/>)에서 PBAIC 2007는 전처리된 3명의 데이터를 제공한다. 우리는 이 중에서 두번째 데이터를 가져왔다. 이 때, 13가지

감정 상태의 평가치 데이터와 MVPA 프로그램 등도 함께 받아왔다. 13번째 피험자의 첫번째 세션에 대하여 각각의 감정 항목들과 상관관계가 높은 복셀들의 인덱스를 추출하였다. 이들 인덱스들로 하이퍼에지를 구성하고 매트랩의 *union*, 과 *ismember* 함수를 이용하여 하이퍼그래프 행렬을 만들었다. 이해하기 쉽게 변형한 그림이 그림 1에 주어져 있다. 복셀은 약 3만 5천개의 개수며 각 감정상태와 관계가 있는 복셀들은 6개부터 256개의 개수를 가진다. 메모리 문제로 선택된 복셀들의 인덱스를 그냥 쓸 수 없으며 *union* 명령으로 복셀들의 합집합(1206개의 원소)을 구하고, *ismember* 명령으로 합집합에서 각 감정 상태의 복셀들의 위치를 1로 표시한다. 결과는 간단히 표현하여 그림 1의 행렬 E 와 비슷하다. 이 행렬 E 로부터 인접 행렬(adjacency matrix) $E^T E$ 를 구하고 이를 대각화하여 고유치와 고유벡터를 구한다. 가장 큰 값의 고유치의 고유벡터를 선택하여 원소를 내림차순으로 정렬한다. 그러면 선택된 복셀에 대하여 중앙성을 중요한 순서대로 부여한 것이 된다. 계산은 64비트 리눅스 머신에서 자유소프트웨어인 영의 매트랩 대용 프로그램인 옥타브(octave)를 이용하였다.

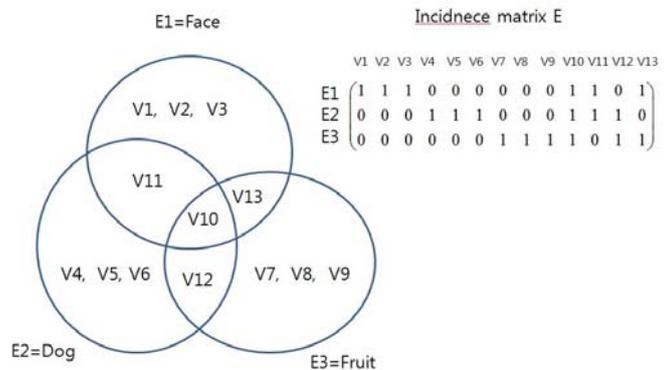


그림 1. 하이퍼그래프와 그 행렬표현. 행렬의 첫째 행에는 얼굴관련 복셀이, 둘째 행에는 개와 관련된 복셀이, 그리고 세번째 행에는 과일과 상관관계가 높은 복셀들이 표시되어 있다. 위 그림은 13가지의 감정 상태의 복셀들의 하이퍼그래프를 이해하기 쉽게 3가지의 하이퍼에지로 그린 것이다.

3. 연구 결과

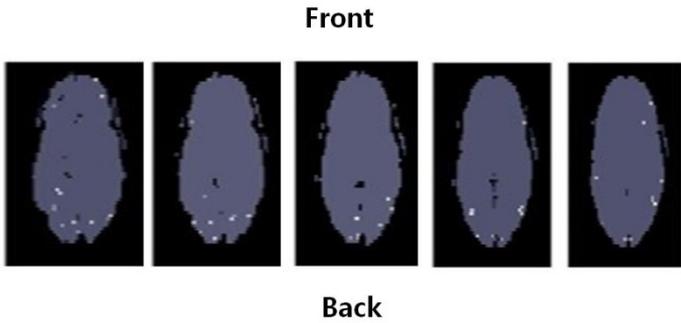


그림 2. 몸(body)과 상관관계가 높은 복셀들의 표시. 그림의 위가 뇌의 앞을, 그리고 그림의 아래가 뇌의 뒤를 나타낸다. 왼쪽부터 오른쪽으로 그림들이 이동할수록 뇌의 아래에서 위로 올라가는 슬라이스(slice)를 나타낸다. 왼쪽 세 그림에서 뇌의 뒤쪽 시각영역에서 활성화된 복셀들이 나타남을 보인다. 또한 오른쪽 두 그림에서 오른쪽 앞쪽의 전전두엽(prefrontal)과 뒤쪽의 두정엽(parietal)이 활성화됨을 알 수 있다. 이는 작업기억(working memory)를 담당하는 영역들이며 특히 몸을 보는 것이 그림으로 인식되므로 오른쪽 반구를 사용한다고 생각된다[9].

그림 2에서 몸(body)과 상관관계가 높은 복셀들을 알 수 있다. 복셀들의 위치는 기존에 알려진 작업기억 관련 영역들과 일치하며 그래픽한 처리는 오른쪽 반구의 우세(dominance)라는 기존의 참고문헌들과 일치한다[9].

13개의 감정관련 상태들과 이들과 상관관계가 높은 복셀들의 개수를 표 1에 정리하였다.

표 1. 13가지 감정 상태와 그들과 각각 상관관계가 높은 복셀들의 개수.

인덱스	감정 상태	복셀의 개수
1	Annoyed angry	253
2	Arousal	101
3	Body	133
4	Dog visible	137
5	Dog	236
6	Faces	205
7	Fearful anxious	49
8	Fruits vegetables	98
9	Gender	55
10	Happy	58
11	Hits fruits	6
12	Hits people	53
13	Hits weapon	48

선택된 복셀들의 합집합의 원소수는 1206개 였다. 이중 상위 26개의 합집합에서의 복셀 인덱스와 실제 복셀 인덱스, 그리고 중앙성(centrality)는 표 2과 같다.

표 2. 첫째열은 1206개의 원소를 가지는 합집합에서의 복셀 인덱스를 중앙도가 큰 순서부터 기록한 것이다. 둘째열은 14만여개의 전체 복셀에서의 인덱스를 나타낸 것이다.

인덱스(합집합에서)	실제 인덱스	중앙도
51	909	0.0738
1153	33607	0.0704
1177	34519	0.0704
1191	34890	0.0704
1204	35169	0.0704
786	18643	0.0704
835	20125	0.0704
878	21864	0.0704
1194	34931	0.0704
27	454	0.0704
133	4027	0.0704
165	4810	0.0704
291	6595	0.0704
427	8822	0.0704
431	8852	0.0704
567	12442	0.0704
611	13441	0.0704
665	15203	0.0704
793	18724	0.0704
1180	34582	0.0640
29	534	0.0640
430	8833	0.0640
127	3909	0.0640
1206	35217	0.0607
772	17864	0.0607
788	18679	0.0607

표 1의 상위 11개의 복셀을 세개의 모듈로 묶을 수 있었다. 즉 모듈 1(51, 27, 133), 모듈 2(786, 835, 878), 그리고 모듈 3(1153, 1177, 1191, 1194, 1204)의 세개의 모듈로 묶어진다. 모듈들은 100, 800, 그리고 1200의 인덱스 근처의 복셀들을 각각의 모듈들로 잡았다. 모듈 원소들은 합집합에서의 인덱스로 표시되었다. 합집합의 인덱스가 오름차순으로 정렬된 원래의 이미지 데이터의 복셀 인덱스에서 추출된 만큼 물리적인 복셀들의 위치역시 근접하리라고 추측된다.

그림 3에는 1206개의 중요성(centrality)들이 내림차순으로 정렬된 결과의 그래프가 제시되어 있다.

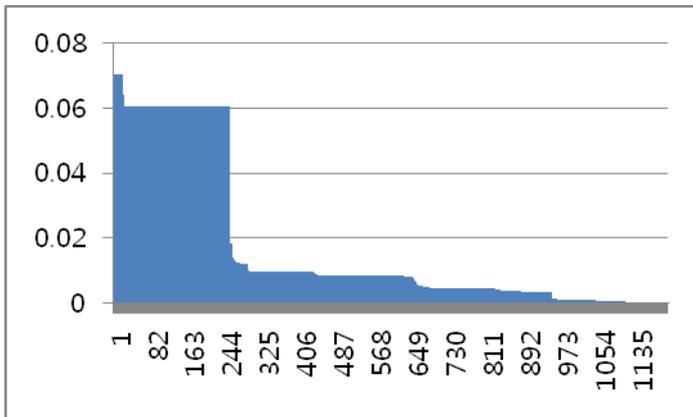


그림 3. 합집합의 1206개 복셀의 중앙성을 내림차순으로 정렬한 결과. 스케일 프리(scale free)에서 볼 수 있는 긴꼬리가 보인다.

4. 논의

일반적으로 뇌의 구조적 그리고 기능적 연결성(connectivity) 모두 작은 세상 네트워크(small-world network)을 따른다고 알려져 있다. 이 작은 세상 네트워크는 국소적으로 중심이 되는 지역적 허브(provincial hub)와 각 모듈을 연결하는 연결 허브(connector hub)로 이루어진다[10]. 이 구조는 뇌에서만뿐만 아니라 인적 네트워크와 컴퓨터 통신망 등의 여러가지 분야의 네트워크에서 보편적으로 나타난다[11].

본 연구는 이러한 작은 세상 구조를 가지는 뇌에서 감정관련 핵심 영역들을 찾는 방법론을 제시한다는 점에서 그 의미가 있다. 작은 세상 구조를 설명하는데 하이퍼그래프가 장점을 가진다는 보고[12]도 있다. 또한 하이퍼네트워크로 분류문제를 풀기도 하였다[8]. 이론적으로 분류문제와 군집화(clustering)문제를 하이퍼그래프로 다룬 연구도 있었다[5]. 고유치와 고유벡터로부터 중앙성(centrality)를 구하는 사회 구조(social network) 분야의 하이퍼그래프 연구[6]를 fMRI 데이터에 적용한 연구는 본 연구가 처음이라고 하겠다.

고유치가 최대인 고유벡터를 선택한 것은 가장 주요한 성분을 선택한 점에서, 주성분 분석(principal component analysis, PCA)방법과 유사하다 하겠다. 이때, 고유벡터의 원소의 크기는 인접 행렬에 참여하고 기여한 정도를 반영하며 각 복셀의 중요성(importance)을 나타낸다고 볼 수 있겠다.

실험 결과의 표 2에서 볼 수 있듯이 상위 11개의 복셀을 크게 3가지의 모듈로 나타낼 수 있었다. 100 근처 복셀들과 800 근처 복셀들, 그리고 1200 근처의

복셀들이 그들이다. 이는 감정처리가 국소적인 영역들을 중심으로 처리되며 이들이 작은 세상을 이룬다고 해석이 가능하다. 뇌신경과학의 큰 주제중의 하나는 분리(segregation)과 통합(integration)의 문제이다[13]. 작은 세상 네트워크는 뇌연결구조와 기능의 분리와 통합을 효율적으로 수행하는 적절한 네트워크 형태라고 하겠다.

마지막으로 제약조건과 하이퍼에지에 관해 논의해 보겠다. 여러 노드를 하나로 묶는 하이퍼에지는 하나의 제약조건을 의미한다고 할 수 있다. 여러 개의 제약조건들이 있을 때 모든 제약조건을 만족하는 해(solution)을 구하는 문제는 인공지능의 중요한 문제이다. 문제의 예로서는, 여러가지 병을 각각의 제약조건으로 보고 노드를 치료약으로 볼 때, 몇가지 병에 동시에 걸린 환자가 어떤 약들을 복용하여야 치료 효과가 좋을지 예측하는 문제가 있겠다. 마찬가지로 본 연구에서 13가지의 감정 상태 관련 복셀들 중에서 여러 감정에 동시에 포함되는 허브역할을 하는 중요한 복셀을 찾아내는 연구는 중요하다. 이러한 연구는 앞으로 감정 조절(control)의 원리를 설명하는 좀 더 중요한 연구로 발전할 수 있다고 생각된다.

5. 감사의 글

이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2010-0018950). 이 논문은 교육과학기술부의 재원으로 국가연구재단의 지원을 받아 수행된 연구(314-2008-1-D00377, Xtran)이며, 지식경제부 및 한국산업기술평가관리원의 IT산업원천기술개발사업(KI002138, 차세대 맞춤형 서비스를 위한 기계학습 기반 멀티모달 복합 정보 추출 및 추천기술 개발, MARS), 지식경제부 산업원천기술개발사업(10035348, 모바일 플랫폼 기반 계획 및 학습 인지 모델 프레임워크 기술 개발) 및 교육과학기술부의 BK21-IT사업에 의해 일부 지원되었음. 논의에 참여해주신 하정우씨께 감사드립니다.

참고문헌

- [1] K.A. Norman, S.M. Polyn, G.J. Detre, J.V. Haxby "Beyond mind-reading: multi-voxel pattern analysis of fMRI data," *Trends in cognitive sciences*, vol.10, pp.424-430, 2006.
- [2] D.W. McRobbie, E.A. Moore, M.J. Graves, M.R. Prince *MRI from picture to proton*, Cambridge university press, 2007.
- [3] S. Huettel, A.W. Song, G. McCarthy *Functional magnetic resonance imaging*. Sinauer Associates, 2004.
- [4] C. Berge *Hypergraphs: combinatorics of finite sets*,

North-Holland press, 1989.

- [5] D. Zhou, J. Huang, B. Schölkopf "Learning with hypergraphs: clustering, classification, and embedding," *Proc. of Advances in neural information processing systems 19*, edited by B. Schölkopf, J. Platt, T. Hoffman, MIT Press, pp.1601—1608, 2007.
- [6] P. Bonacich, A.C. Holdren, M. Johnston "Hyper-edges and multidimensional centrality," *Social networks*, vol.26, pp.189-203, 2004.
- [7] J.W. Ha, B.-T. Zhang "Evolutionary hypernetwork models for higher order pattern recognition on real-valued feature data without discretization," *Journal of the Korea Information Science Society: Software and Applications*, vol.37, pp.120-128, 2010. (in Korean)
- [8] J.-W. Ha, J.W. Jang, D.-H. Kang, W.H. Jung, J.S. Kwon, B.-T. Zhang "Gender classification with cortical thickness measurement from magnetic resonance imaging by using a feature selection method based on evolutionary hypernetworks," *Proc. of the 18th international conference on Fuzzy Systems*, IEEE press, pp.41-46, 2009.
- [9] H. Koshino, R. Kana, T.A. Keller, V.L. Cherkassky, N.J. Minshew, M.A. Just "fMRI investigation of working memory for faces in autism: visual coding and underconnectivity with frontal areas," *Cerebral cortex*, vol.18, pp.289-300, 2008.
- [10] O. Sporns, C.J. Honey, R. Kötter "Identification and classification of hubs in brain networks," *Plos one*, vol.2, pp.e1049, 2007.
- [11] A.-L. Barabasi "Scale-free networks: a decade and beyond," *Science*, vol.325, pp.412-413, 2009.
- [12] E. Estrada, J.A. Rodríguez-Velázquez "Subgraph centrality and clustering in complex hyper-networks," *Physica A* vol.364, pp.581-594, 2006.
- [13] G. Tononi, O. Sporns, G.M. Edelman "A measure for brain complexity: relating functional segregation and integration in the nervous system," *PNAS*, vol.91, pp.5033-5037, 1994.