

평균 크기 지각에 대한 계산 모형*

백 중 수¹⁾

정 상 철²⁾³⁾*

¹⁾연세대학교 글로벌융합기술원 ²⁾연세대학교 심리학과 ³⁾연세대학교 인지과학협동과정

인간의 시각 체계는 처리 용량의 한계를 극복하기 위해 복잡한 시각 자극의 중복되는 정보들을 요약해서 표상한다. 요약된 표상의 한 가지 예는 다양한 시각 자극의 평균 크기를 표상하는 것인데, 많은 선행 연구들을 통해 시각 체계의 평균 표상 능력은 매우 정확하다고 알려져 있다. 본 연구에서는 이와 같은 평균 크기 표상의 기제를 설명하는 계산 모형을 제안한다. 제시된 모형에서 시각 체계는 개별 자극들을 다소간의 잡음(초기 잡음)과 함께 부호화하고, 여러 자극들에서 부호화된 정보들을 통합한다. 시각 체계는 정보가 통합된 이후에도 의식 수준의 최종적인 표상이 형성되기 전까지 일정 수준의 잡음이 추가적으로 더해진다(후기 잡음). 본 모형의 타당도를 검증하기 위해 여러 크기의 자극들을 포함한 기준 자극 화면과 비교 자극 화면의 평균 크기를 비교하는 정신물리학 실험을 실시했다. 실험 결과, 크기 차이 변별의 민감도를 나타내는 역치는 자극 개수에 따라 감소했다. 다시 말해, 실험 참가자들은 자극 개수가 많아짐에 따라 그 자극들의 평균 크기를 더 정확하게 지각했는데, 이 결과는 후기 잡음을 포함한 모형으로 잘 설명된다. 본 모형은 개별 시각 정보들이 어떤 과정을 거쳐서 평균 표상을 형성하게 되는지를 보여주며, 본 연구에 사용된 실험 패러다임은 다양한 시각 속성의 평균 표상을 연구하는 데에 유용한 도구가 될 것으로 기대된다.

주제어 : 평균 표상, 지각, 계산모형, 초기 소음, 후기 소음, 중심극한정리, 신호처리이론

* 이 논문은 과학기술정보통신부 및 정보통신기술진흥센터의 ICT명품인재양성사업(IITP-2018-2017-0-01015)과 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(NRF-2016R1A2B4016171).

† 교신저자 : 정상철, 연세대학교 심리학과 및 인지과학협동과정, (03722) 서울특별시 서대문구 연세로 50
E-mail: scchong@yonsei.ac.kr

제한된 처리 용량을 가진 인간의 시각 체계는 매순간 주어지는 많은 시각적 입력을 효율적으로 처리해야 한다. 시각 체계는 모든 세부 정보들을 처리하기 보다는 복잡한 장면의 중복을 줄이기 위해 통계적 요약 정보를 이용해 세상을 표상한다(Alvarez, 2011; Ariely, 2001; Chong & Treisman, 2003). 일련의 연구들을 통해 시각 체계는 대상의 크기와 같은 시각 세부 특징(feature)의 평균을 매우 정확하게 계산한다는 것이 밝혀졌고(Chong & Treisman, 2003, 2005a, 2005b), 시각 정보에 대한 정확한 요약 능력은 많은 연구자들의 관심 대상이 되어 왔다. 이와 관련된 많은 연구들에서 여러 자극들의 평균 크기를 표상하는 능력은 평균의 대상이 되는 자극 개수와 상관없이 일정하다고 보고되어 왔는데(Allik, Toom, Raidvee, Averin, & Kreegipuu, 2013; Alvarez, 2011; Ariely, 2001; Chong & Treisman, 2005b), 다른 연구들에서는 자극 개수가 늘어남에 따라 평균 표상의 정밀도가 더 높아지는 것으로 나타났다(Haberman & Whitney, 2010; Lee, Baek, & Chong, 2016; Robitaille & Harris, 2011).

최근 제안된 ‘잡음과 선택 모형’은 왜 평균 표상 과정에서 자극 개수 효과가 나타나지 않는지에 관한 설명을 제공한다(Allik et al., 2013). 이 모형은 개별 자극들은 다소간의 잡음과 함께 부호화되며, 시각 체계는 모든 자극들을 처리하기 보다는 몇 개의 자극만을 선택해서 처리한다고 가정한다. 이 모형에서 시각 체계 내부의 잡음과 선택 과정은 개별 자극들의 정보가 통합되어지기 이전에 발생하기 때문에, 일부 자극은 평균 계산에 반영되지 못하고 이후 정보처리 과정에서 제거된다. 하

지만, Chong과 그의 동료들은 의도적으로 자극들 중 일부를 화면에 제시하지 않았을 때와 모든 자극들이 화면에 제시되었을 때에 자극들의 평균 표상의 정확도가 차이가 남을 밝힌바 있다(Chong, Joo, Emmanouil, & Treisman, 2008). 이 결과는 ‘잡음과 선택 모형’으로는 설명되지 못하며, 제시된 모든 자극들이 정보 통합 과정에 포함됨을 시사한다. 또한 ‘잡음과 선택 모형’은 내적 잡음이 정보 통합 과정 이전의 개별적인 자극 표상에만 발생하는 것으로 가정하지만, 평균 크기에 대한 표상이 형성된 이후에도 발생할 수 있다. 즉, 추가적인 잡음이 지각적인 처리 과정 이후에 발생해 자극 수준과 상관없이 관찰자의 수행에 영향을 미치게 된다. 심리학 실험에서 이와 같은 ‘후기 잡음(late noise)’은 자극 강도가 충분히 강해도 관찰자들의 수행률이 일정 수준 이상으로 올라가지 못하는 - 혹은 역치가 일정 수준 이하로 내려가지 못하는 - 현상으로 표현되며, 다양한 실험 패러다임과 자극을 사용한 연구들에서 흔히 관찰된다. 시각 과정의 속성과 기제를 밝히기 위해 개발된 많은 관찰자 모형들은 이를 반영하기 위해 후기 잡음을 포함한다(Lu & Doshier, 1998, 1999, 2008; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001). 평균 크기 표상이 자극 개수에 상관없이 정확하다는 Chong과 Treisman(2005b)의 연구 결과는 후기 잡음의 효과에 기인했을 가능성이 있다.

본 연구에서는, 여러 개의 시각 자극들의 평균을 표상하는 과정에 대한 관찰자 모형을 개발했다. 이 모형은 ‘잡음과 선택 모형’에서 고려되지 않았던 ‘후기 잡음’을 모형에 포함함으로써, 왜 기존 연구들에서 자극 개수 효과

가 관찰되지 않았는지에 대한 설명을 제공한다. 관찰자가 기준 자극 세트와 비교 자극 세트를 비교해 자극들의 평균 크기가 더 큰 것을 선택해야 하는 2지 강제 선택 과제(2 Alternative Forced Choice; 2AFC) 실험 결과를 통해 제안된 모형의 타당도를 검증했다.

모형

본 연구에서는, 시각 정보 처리 과정에서의 잡음, 정보 통합, 의사 결정 등의 일련의 인지 처리 과정을 그 구성 요소로 하는 관찰자 모형(observer model)을 개발하였다. 이 모형은 지각적 잡음을 반영하기 위해 초기 잡음과 후기 잡음을 가정했고, 자극 개수에 따른 잡음 상쇄 정도를 반영하기 위해 중심극한정리를 적용했다. 또한, 의사 결정 과정을 모형화하기 위해 신호탐지이론을 포함한다(Green & Swets,

1966; Macmillan & Creelman, 1991). 본 모형의 개념적 도식은 Figure 1에, 수리적 분석 모형은 부록에 수록되어 있다.

초기 잡음 시각 체계의 정보 처리 과정은 오차를 포함하는데, 이러한 오차는 수용기의 표집 오류, 무선적인 신경 반응 등 다양한 내재적 잡음에 의해 발생한다. 본 모형은 시각 자극들이 개별적으로 부호화되고, 부호화된 각각의 크기 정보에 무선적인 내적 잡음이 추가되는 것을 가정한다. 구체적으로, 각각의 자극은 물리적 자극 강도의 지수 함수에 비례하는 특정 강도의 내적 반응을 유발하게 된다(Foley & Legge, 1981; Nachmias, 1981; Nachmias & Sansbury, 1974). 예를 들어, 동그라미는 그 면적의 0.76제곱에 비례하는 크기로 지각된다(Teghtsoonian, 1965). 이런 지각된 크기와 그에 비례하는 강도의 초기 잡음이 합쳐져서 개별

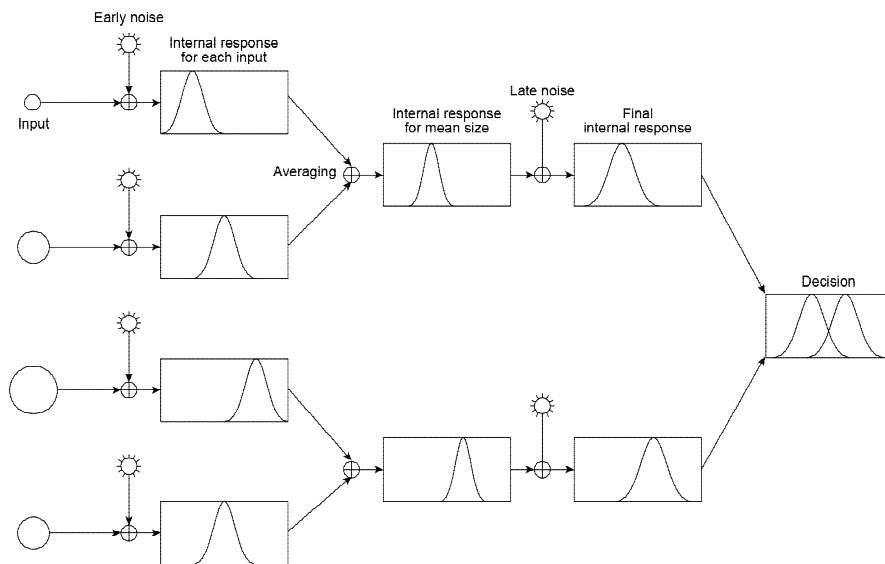


Figure 1. Internal processes for averaging sizes.

자극에 대한 내적 반응을 산출한다.

정보 통합 시각 체계는 시각장에 제시된 여러 자극들을 요약하기 위해 여러 개의 내적 반응들을 통합하게 된다. 평균 계산 과정에서 개별 자극에 대한 무선적인 잡음은 서로 간에 상쇄되기 때문에, 많은 자극들에 대한 평균 크기 표상은 적은 자극들의 평균 크기 표상보다 더 정확하게 된다.

후기 잡음 시각 체계의 내적 오차는 정보 처리 과정 전반에 걸쳐서 발생한다. 따라서, 본 모형은 정보 통합 과정 이후의 평균 표상에 일정 수준의 내적 잡음이 추가됨을 가정한다. 많은 시각 과제에서 민감도를 측정하는 역치는 일정 수준 이하로 내려가지 않는 경우가 많은데, 이와 같은 역치의 한계는 후기 잡음에 기인하는 것으로 설명되곤 한다(Burgess, Wagner, Jennings, & Barlow, 1981; Eckstein, Ahumada, & Watson, 1997; Heeley & Buchanan-Smith, 1996; Lu & Doshier, 1999; Morgan, Ward, & Hole, 1990; Parkes et al., 2001; Pelli, 1985).

의사 결정 관찰자가 기준 자극 세트와 비교 자극 세트를 비교해 각 세트에 있는 자극들의 평균 크기가 더 큰 것을 선택해야 할 때, 신호탐지이론은 각각의 자극 세트에 대한 내적 표상의 정규 분포를 가정한다. 때 시행에서 관찰자는 두 개의 내적 분포에서 나온 무선 변수들을 비교하고, 이 중 더 큰 것을 유발한 자극 세트가 더 큰 평균 크기를 가지고 있다고 반응한다. 시각민감도를 나타내는 d' 은 신

호 대 잡음비로 정의되는데, 이것은 두 자극 세트에 대해 형성된 평균 표상 분포의 평균 차이와 표준편차에 따라 달라진다. 다시 말해, 두 자극 세트 간의 평균 크기의 차이가 클수록 더 큰 내적 반응의 차이를 나타내게 되며, 이는 실험 결과에서 더 높은 민감도로 표현된다. 또한, 정확한(precise) 내적 반응은 평균 크기 비교 과제에서 더 높은 민감도를 나타내게 된다.

모형 예측 Figure 2는 제시된 모형이 예측하는 자극 개수에 따른 평균 크기 차이의 역치를 보여준다. 초기 잡음만을 가정하는 모형(왼쪽)과 초기 및 후기 잡음 모두를 가정하는 모형(오른쪽) 모두 자극 개수가 커짐에 따라 역치가 작아질 것을 예측한다. 초기 잡음만을 가정하는 모형에서는 역치가 자극 개수에 따라 로그-로그 축에서 -0.5의 기울기를 가질 것으로 예측된다. 이는 중심극한정리를 통해 유도되는 자극 개수와 내적 표상의 변량 사이의 \sqrt{n} 관계를 나타내는 것이다. 후기 잡음도 포함하는 모형에서는 역치와 자극 개수 관계의 기울기가 -0.5보다 완만하고(크고), 특히 자극 개수가 증가함에 따라 점점 더 완만해진다. 결과적으로 큰 자극 개수 조건(예를 들면, 32 혹은 64)에서는 자극 개수 간의 역치 차이가 매우 작은 것으로 나타나며, 모형 예측이 아닌 경험적 실험 데이터에서는 이들 간의 유의미한 차이를 찾기 어려울 수도 있다(Allik et al., 2013; Alvarez, 2011; Ariely, 2001; Chong & Treisman, 2005b).

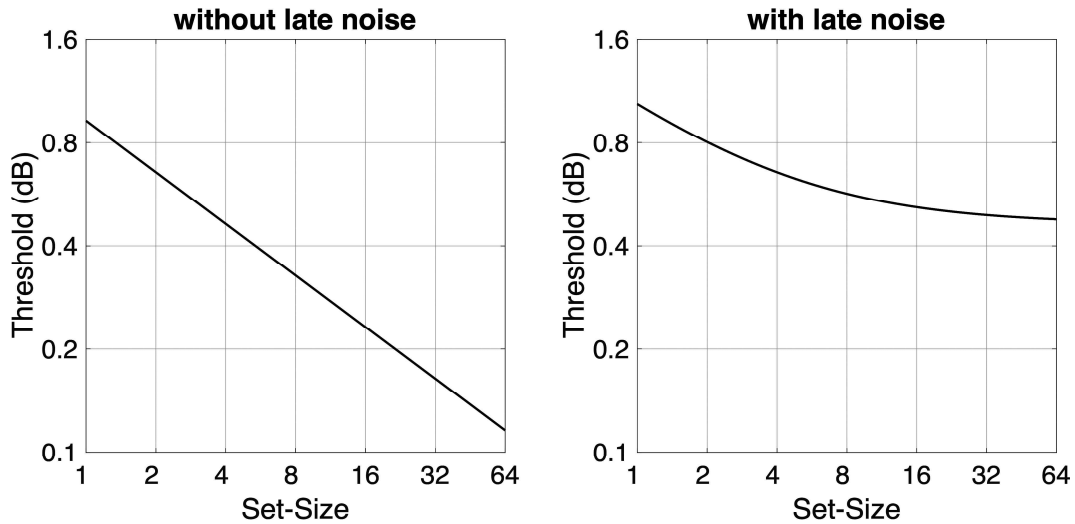


Figure 2. Predictions by the model without late noise, and the model with late noise. Thresholds, corresponding to percent choice of test (pc) was 79.4%, were plotted as a function of set-size on a log-log axis.

실 험

제안된 계산 모형의 타당도를 검증하기 위해 정신물리학 실험을 실시하였다. 매 시행에서 여러 자극들로 이루어진 두 개의 화면이 연속적으로 제시되었고, 실험 참가자들은 이 두 개의 화면들 중 어떤 화면이 더 큰 평균 자극 크기를 가지고 있는지 판단했다. 실험 결과에서 관찰된 자극 개수에 따른 역치의 변화와 모형에서 예측하는 역치 차이를 비교해 모형의 타당도를 검증했다.

방 법

참가자 연세대학교 대학원생 8명이 실험에 참가하였다. 모든 참가자들은 교정시력이 정상이었으며, 실험의 가설에 대해 사전에 알지 못했다. 본 실험은 연세대학교 IRB의 승인 하

에 이루어졌고, 참가자들은 실험 참가 동의서를 작성한 후 실험에 참여하였다.

기구 실험은 IBM 호환 PC와 21인치 모니터(HP P1230)를 이용해 실시되었다. 자극은 MATLAB과 Psychtoolbox(Brainard, 1997; Pelli & Zhang, 1991)를 통해 통제되었으며, 1600×1200 픽셀 해상도의 화면에 제시되었다. 참가자와 모니터간의 거리는 이마-턱 받침대를 통해 60cm로 유지되었다.

자극 화면 중앙으로부터 시각도 6.49° 거리의 가상의 원 위에 등간격으로 위치한 16개의 자리에 동그라미 자극들이 제시되었다. 화면에 제시되는 동그라미들의 위치는 이 가상의 16개 자리들 중에서 무선적으로 선택되었으며, 그 후 선택된 위치에 상하좌우로 ±0.80° 범위 이내의 무선적 변화를 추가하였다. 실험에 사

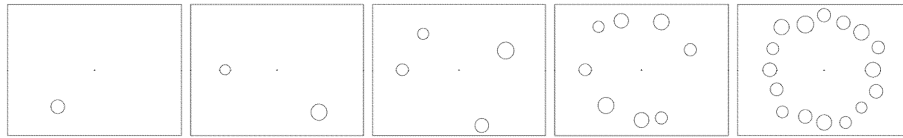


Figure 3. Examples of stimulus display. Set-size 1, 2, 4, 8, and 16.

용된 자극의 예가 Figure 3에 제시되어 있다. 자극은 회색의 배경 화면에 검은색 실선의 동그라미들이 제시되었다. 매 시행에서 기준 자극 세트로 제시된 동그라미들의 평균 크기는 시각도 1.45-1.83°에서 무선적으로 선택되었고, 비교 자극 세트로 제시된 동그라미들의 평균 크기는 계단법(Staircase procedure)에 의해서 결정되었다. 기준 자극의 평균 크기와 비교 자극의 평균 크기 차이는 데시벨 (dB) 단위로 계산되었다. 데시벨은 두 수치간 비율의 상용로그 값($20 \times \log_{10} \frac{y}{x}$)으로, 에너지의 강도 비율을 표현하는 단위로 광범위하게 사용되고 있다. 0 dB는 두 수치간 비율이 1:1, 1 dB는 1:1.12, 2 dB는 1:1.26임을 의미한다. 평균 크기는 각 자극 면적의 0.76제곱에 대한 산술 평균으로 계산되었다(Chong & Treisman, 2003; Teghtsoonian, 1965). 기준 자극과 비교 자극의 세트 안 개별 자극들의 크기는 주어진 목표 평균 크기의 $\pm 26\%$ 범위 이내의 균일 분포(uniform distribution)에서 무선적으로 선택되었다. 자극 개수가 1개일 때에는 자극의 크기가 목표 평균 크기로 제시되었다.

설계 본 연구에서는 5개의 자극 개수(1, 2, 4, 8, 16)가 사용되었다. 각 자극 개수별로 80시행이 실시되었으며, 자극 개수는 무선적인 순서로 제시되었다. 각 참가자는 약 30분정도

걸리는 실험을 이틀에 걸쳐 두 번 반복해서 실시했다.

절차 시행 절차가 Figure 4에 예시되어 있다. 매 시행은 화면 중앙에 제시되는 응시점과 함께 시작된다. 응시점 제시 500ms 후, 두 개의 자극 화면이 각각 순서대로 500ms동안 제시되는데, 그 사이에는 500ms동안 빈 화면이 제시된다. 기준 자극과 비교 자극은 첫 번째 혹은 두 번째 자극 화면에 제시된다. 반응 편향이 작다고 알려진 2지 강제 선택 과제(2AFC)에서도 피험자들은 첫 번째 혹은 두 번째 화면을, 아니면 기준 자극 혹은 비교 자극을 더 많이 선택하는 편향을 나타낼 수 있다(Klein, 2001; Lesmes, Lu, Baek, Tran, Doshier, & Albright 2015; Yeshurun, Carrasco, & Maloney, 2008). 특정 화면에 대한 반응 편향을 최소화하기 위해, 기준 자극과 비교 자극이 나타나는 순서가 무선적으로 선택되었고, 실험참가자는 어떤 화면이 기준 자극인지 사전에 알지 못하도록 조작했다. 즉, 어떤 시행에서는 기준 자극이 첫 번째 화면에 나타나고, 다른 시행에서는 두 번째 화면에 나타나기 때문에, 참가자들은 어떤 화면이 기준 자극인지를 사전에 알지 못한 채로, 첫 번째 혹은 두 번째 화면에서 평균 자극 크기가 더 큰지를 보고했다. 또한 특정 평균(예, 원주 기반 평균)으로의 반응 편향을

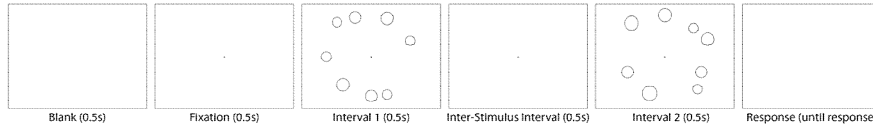


Figure 4. Illustration of experimental procedure.

막기 위해, 반응에 대한 피드백도 제시되지 않았다(Bauer, 2009).

매 시행에서 사용될 비교 자극의 평균 크기는 각 자극 개수 조건별로 3-up 1-down 계단법에 의해 결정되었다. 참가자가 비교 자극의 평균 크기가 기준 자극보다 더 크다고 세 번 연속해서 응답하면 다음 시행에서 비교 자극의 평균 크기가 줄어들고, 비교 자극이 기준 자극보다 더 작다고 한 번 응답하면 다음 시행에서 비교 자극의 평균 크기가 더 커졌다. 이 절차는 비교 자극이 기준 자극보다 더 크다고 응답하는 확률이 79.4%에 해당되는 역치로 수렴한다(Levitt, 1971). 다시 말해, 비교 자극의 평균 크기가 어느 정도가 되었을 때에, 관찰자가 79.4%의 확률로 비교 자극이 기준 자극보다 더 크다고 하는지를 측정한다. 계단법에서 (1) 지난 시행에서 자극 강도가 강해지거나 유지되다가 이번 시행에서 자극 강도가 약해지는 경우와 (2) 지난 시행에서 자극 강도가 약해지거나 유지되다가 이번 시행에서 강해지는 경우를 전환점(reversal)이라고 한다. 각 자극 개수 조건별로 전환점들의 자극 강도를 평균 내 역치를 계산하였다. 조건별 80개의 시행에서, 전환점이 홀수개일 경우에는 처음 세 개 이후의 전환점들을, 전환점이 짝수개일 경우에는 처음 네 개 이후의 전환점들을 평균 냈다(Lu & Doshier, 2013). 계단법이 더 빠르고 정확하게 역치를 측정할 수 있도록, 실험 전

반을 거쳐 단계 크기(step size)를 점진적으로 조정했다(Levitt, 1971; Lu & Doshier, 2013). 즉, 각 자극 개수 조건별로 계단법 절차는 0.8 dB의 단계 크기로 시작해, 첫 번째 전환점 이후에는 0.4 dB, 세 번째 전환점 이후에는 0.2 dB, 일곱 번째 전환점 이후에는 0.1 dB, 그리고 열다섯 번째 전환점 이후에는 0.05 dB의 간격으로 평균 크기가 조정되었다. 계단법을 포함한 적응형 실험 절차들(adaptive procedure)은 첫 시행에서의 자극 강도에 많은 영향을 받는 것으로 알려져 있다. 이와 같은 시작점의 영향을 최소화하기 위해, 각 자극 개수 조건별 계단 절차의 첫 시행에서 비교 자극과 기준 자극 평균 크기 차이를 0.8 ± 0.4 dB 범위 내에서 무선적으로 선정했다.

결 과

자극 개수 효과 계단절차법으로 측정된 역치가 자극 개수 조건에 따라 달라지는지를 확인했다. 반복 측정된 역치에 대한 일원배치 분산분석 결과, 자극 개수 효과는 유의미한 것으로 나타났다($F(4,28)=7.519$, $p<.001$, $\eta^2=0.518$). 사후 검증 결과, 자극 개수가 1인 조건과 8인 조건($p=0.043$), 그리고 자극 개수가 2인 조건과 8인 조건 사이($p=0.039$)에서 유의미한 차이가 관찰되었다. 즉, 평균 차이의 변별 역치는 자극 개수에 따라 차이가 나는 것으로

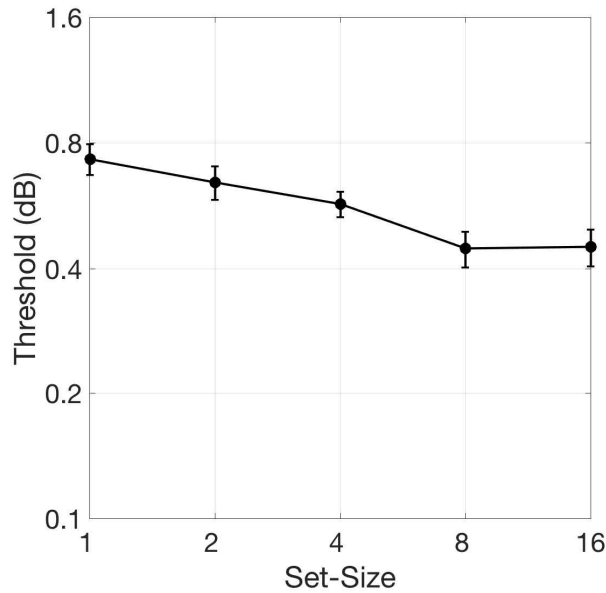


Figure 5. Threshold as a function of set-size. Thresholds corresponding to $pc=79.4\%$ decreased with set-size. Error bars represent standard errors of the mean (SEM).

나타났다. Figure 5에 나타나듯이, 역치는 자극 개수가 많아짐에 따라 줄어들다가 자극 개수 8개 이후로는 비슷한 수준을 나타낸다. 이 결과는, 실험 참가자들은 자극 개수가 많아짐에 따라 그 자극들의 평균 크기를 더 정밀하게 (precise) 지각했음을 의미한다.

모형 검증 후기 잡음을 가정하지 않은 모형에 따르면, 역치는 자극 개수가 늘어남에 따라 지속적으로 줄어든다(Figure 2a). 반면, 후기 잡음을 가정한 모형은 자극 개수가 커짐에 따라 역치가 점진적으로 낮아지다가, 그 감소는 점진적으로 둔화되어 역치는 일정 수준으로 수렴하게 된다(Figure 2b). 두 개의 모형들을 검증하기 위해, 최소자승화(least square) 추정 기법을 통해, 초기 잡음만을 포함한 경감 모

형(공식 A1과 A8)을 실험에서 측정된 역치에 피팅하였는데, 이 모형은 역치에 대해 매우 빈약한 설명력을 보여주고 있다($r^2=-0.808$). 반면, 초기 잡음뿐 아니라 후기 잡음도 포함한 전체 모형(공식 A2과 A8)을 역치에 피팅한 결과, 이 모형은 경감 모형에 비해 더 나은 모형적합도를 나타냈다($r^2=0.928$). 모형에서 포함하고 있는 파라미터들에 대한 최적추정치는 표1에 요약되어 있고, 그에 따른 모형 예측치는 Figure 6에 정리되어 있다.

제안된 모형들을 평가하기 위해, 경감 모형과 전체 모형의 모형적합도를 비교했다. 두 모형은 중첩(전체 모형이 경감 모형을 수리적으로 포함)되어 있는데, 이 경우에 파라미터 개수가 더 많은 전체 모형은 더 적은 경감 모형보다 항상 더 나은 적합도(r^2)를 나타낸다.

모형적합도 비교는 이러한 파라미터 개수로 인한 적합도 향상을 감안했을 때에도 더 나은 설명력을 갖는지, 즉 후기 잡음이 주어진 데이터를 설명하는데 유의미한 요소인지를 검증하는 과정이다. k_{full} 개의 변수를 포함한 전체 모형과 $k_{reduced}$ 개의 변수를 포함한 경감 모형의 비교를 위해 F 검증을 실시했는데, 이에 사용된 F 값은 다음과 같이 정의된다(Wannacott & Wannacott, 1981).

$$F(df_1, df_2) = \frac{(r_{full}^2 - r_{reduced}^2) / df_1}{(1 - r_{full}^2) / df_2}$$

여기서 $df_1 = k_{full} - k_{reduced}$ 이고, $df_2 = N - k_{full}$ 이며, N 은 데이터 포인트의 개수(이 연구에서는 자극 개수 조건의 수인 5)를 의미한다. 검증 결과, 모형에 후기 잡음이 포함되었을 때 모형의 적합도는 유의미하게 증가했다($F(1, 3) = 72.007$, $p = 0.003$). 이 결과는 후기 잡음이 평균 크기 표상 과제를 수행하는 관찰자의 수행률을 설

Table 1. The best fitting parameters

	σ_1	σ_2	r^2
Full Model	0.537	0.372	0.928
Reduced Model	0.777	-	-0.808 ¹⁾

¹⁾ r^2 is negative when the model fits data worse than a horizontal line. Results show the reduced model explains data very poorly.

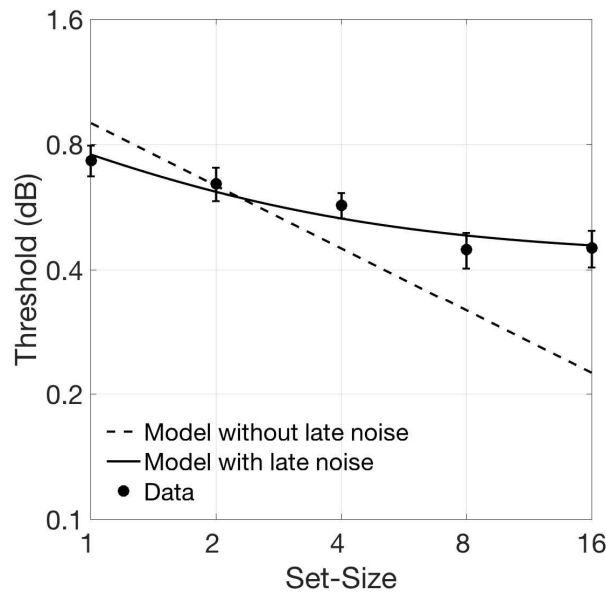


Figure 6. Psychophysical data and best fitting models. Error bars represent standard errors of the mean (SEM).

명하는 데 유의미한 요소임을 의미한다.

논 의

본 연구에서는 평균 크기 표상에 관한 계산 모형을 제시했다. 이 모형에서 개별 자극들은 초기 잡음과 함께 부호화되며, 부호화된 개별 크기 정보들은 통합되어져서 평균 크기에 관한 내적 표상이 형성되고, 추가적으로 후기 잡음이 더해지게 된다. 모형 검증을 위한 정신물리학 실험에서, 본 연구에서는 화면에 제시되는 자극 개수를 변화시켜가면서 자극들의 평균 크기를 변별하는 과제를 실시했다. 후기 잡음을 고려하지 않은 계산 모형은 관찰자들의 실험 결과를 잘 설명하지 못했지만, 후기 잡음을 포함했을 때에는 관찰자들의 평균 크기 표상에 관한 역치를 매우 잘 설명했다.

본 모형이 가정한 잡음은 신호 탐지기부터 의사 결정 및 반응 단계를 아우르는 시각 체계 전반에서 생기는 무선적인 오차들을 모형화한 것이다. 이와 같은 잡음은 정보 통합 과정 이전뿐만 아니라 이후에도 생길 수 있다고 보는 것이 타당하다. 후기 잡음을 포함한 모형은 매우 많은 개수의 자극들이 제시된다 하더라도 평균 크기 표상에 관한 역치가 특정 수준 이하로 떨어지지 않을 것으로 예측한다. 다시 말해, 자극 개수가 커질수록 표상은 점차 정밀해지지만, 자극 개수가 무한대에 가까워지더라도 정밀도는 0이 되지 않고 약간의 오차를 포함한다. 대신, 이와 같은 역치 감소는 자극 개수가 커질수록 둔화되어, 역치는 특정 수준으로 수렴하게 된다.

후기 잡음으로 인한 역치 감소의 둔화는 그

동안 평균 크기 표상에 관한 연구들에서의 자극 개수 효과에 대해 왜 상반된 결과들이 나오게 되었는지에 대한 설명을 제공해준다. 본 연구 결과에 의하면, 지난 연구들에서 자극 개수 효과를 발견하지 못한 이유는 자극 개수 효과가 잘 표현되지 않는 영역을 주로 측정했기 때문인 것으로 보인다. 본 연구에서 자극 개수 효과는 자극 개수 2에서 8사이에서 가장 분명하게 나타났지만, 자극 개수 효과를 발견한 연구들은 주로 이보다 큰 자극 개수를 이용해 실험을 진행했다. 예를 들면, Chong과 Treisman(2005b)은 자극 개수 8과 16에 집중해 데이터를 수집했다. 평균 크기 표상에 관한 자극 개수 효과가 작은 자극 개수에서 분명하게 나타남을 감안할 때(Figure 2), 비교적 큰 자극 개수들을 이용한 연구들에서는 통계적으로 유의미한 자극 개수 효과를 확인하기 쉽지 않았음을 의미한다.

본 연구는 평균 크기 표상에 관한 계산 모형을 제안하고, 정신 물리학 실험을 통해 이를 검증했다. 제안된 계산 모형은 실험 데이터를 매우 잘 예측할 뿐만 아니라, 이전 연구들에서 발견된 자극 개수 효과에 관한 상반된 결과도 잘 설명한다. 본 연구에서는 적응형 절차(adaptive procedure)인 계단법을 사용해 각 자극 개수 조건별 크기 변별 역치를 측정했다. 적응형 절차를 사용했기 때문에, 불필요한 데이터 수집을 최소화하고 실험에 소요되는 시간을 줄일 수 있었다. 이와 같은 실험 절차는 평균 표상을 연구하는 추후 연구들에서 유용하게 사용될 수 있을 것으로 기대된다.

본 연구에서는 시각 자극의 여러 속성들 중 평균 크기 표상의 형성에만 초점을 맞추었다.

하지만, 시각 체계는 이 외에도 자극의 방향(Dakin & Watt, 1997), 색상(Maule, Witzel, & Franklin, 2014), 운동 방향(Williams & Sekuler, 1984), 운동 속도(Watamaniuk & Duchon, 1992) 등 다양한 기본적 시각 속성뿐 아니라 선호도(Leib, Kosovicheva, & Whitney, 2016), 성별(Haberman & Whitney, 2007), 시선(Sweeny & Whitney, 2014), 얼굴 표정(Haberman & Whitney, 2007, 2009) 등의 상위 시각 속성에 대해서도 정확한 평균 표상을 형성한다는 것이 보고되었다. 추후 연구를 통해, 본 연구에서 제시된 모형이 다른 시각적 속성의 평균 표상에도 적용될 수 있는지 연구되어야 할 것이다. 뿐만 아니라, 정보 처리 과정에 관여하는 여러 가지 다른 인지 기제들 - 예를 들면, 선택적 주의의 효과 - 을 포함한 확장 모형의 개발이 필요할 것으로 보인다.

References

- Allik, J., Toom, M., Raidvee, A., Averin, K., & Kreegipuu, K. (2013). An almost general theory of mean size perception. *Vision Research*, 83, 25-39.
- Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, 15(3), 122-31.
- Ariely, D. (2001). Seeing Sets: Representation by Statistical Properties. *Psychological Science*, 12(2), 157-162.
- Bauer, B. (2009). The danger of trial-by-trial knowledge of results in perceptual averaging studies. *Attention, Perception & Psychophysics*, 71(3), 655-65.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433-436.
- Burgess, A. E., Wagner, R. F., Jennings, R. J., & Barlow, H. B. (1981). Efficiency of human visual signal discrimination. *Science*, 214(4516), 93-4.
- Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, 43(4), 393-404.
- Chong, S. C., & Treisman, A. (2005a). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics*, 67(1), 1-13.
- Chong, S. C., & Treisman, A. (2005b). Statistical processing: computing the average size in perceptual groups. *Vision Research*, 45(7), 891-900.
- Chong, S. C., Joo, S. J., Emmanouil, T.-A., & Treisman, A. (2008). Statistical processing: not so implausible after all. *Perception & Psychophysics*, 70(7), 1327-34.
- Dakin, S. C., & Watt, R. J. (1997). The computation of orientation statistics from visual texture. *Vision Research*, 37(22), 3181-3192.
- Eckstein, M. P., Ahumada, A. J., & Watson, A. B. (1997). Visual signal detection in structured backgrounds. II. Effects of contrast gain control, background variations, and white noise. *Journal of the Optical Society of America, A, Optics, Image Science, and Vision*, 14(9), 2406-19.

- Foley, J. M., & Legge, G. E. (1981). Contrast detection and near-threshold discrimination in human vision. *Vision Research*, *21*(7), 1041-53.
- Fouriez, G., Rubinfeld, S., & Capstick, G. (2008). Visual statistical decisions. *Perception & Psychophysics*, *70*(3), 456-464.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology : CB*, *17*(17), R751-3.
- Haberman, J., & Whitney, D. (2010). The visual system discounts emotional deviants when extracting average expression. *Attention, Perception & Psychophysics*, *72*(7), 1825-38.
- Haberman, J., & Whitney, D. (2009). Seeing the mean: ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 718-34.
- Heeley, D. W., & Buchanan-Smith, H. M. (1996). Mechanisms specialized for the perception of image geometry. *Vision Research*, *36*(22), 3607-27.
- Klein, S. A. (2001). Measuring, estimating, and understanding the psychometric function: a commentary. *Perception and Psychophysics*, *63*, 1421-1455.
- Lee, H., Baek, J., & Chong, S. C. (2016). Perceived magnitude of visual displays: Area, numerosity, and mean size. *Journal of Vision*, *16*(3), 12.
- Leib, A. Y., Kosovicheva, A., & Whitney, D. (2016). Fast ensemble representations for abstract visual impressions. *Nature Communications*, *7*, 13186.
- Lesmes, L. A., Lu, Z.-L., Baek, J., Tran, N., Doshier, B. A., & Albright, T. D. (2015). Developing Bayesian adaptive methods for estimating sensitivity thresholds (d') in Yes-No and forced-choice tasks. *Frontiers in Psychology*, *6*, 1070.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2), 467-477.
- Lu, Z.-L., & Doshier, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Research*, *38*(9), 1183-98.
- Lu, Z.-L., & Doshier, B. A. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *Journal of the Optical Society of America A*, *16*(3), 764.
- Lu, Z.-L., & Doshier, B. A. (2008). Characterizing observers using external noise and observer models: assessing internal representations with external noise. *Psychological Review*, *115*(1), 44-82.
- Lu, Z.-L., & Doshier, B. A. (2013). *Visual Psychophysics: From Laboratory to Theory*. Cambridge: The MIT Press.
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection Theory: A User's Guide*. New York: Cambridge University Press.
- Maule, J., Witzel, C., & Franklin, A. (2014). Getting the gist of multiple hues: metric and

- categorical effects on ensemble perception of hue. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, 31(4), A93-102.
- Morgan, M. J., Ward, R. M., & Hole, G. J. (1990). Evidence for positional coding in hyperacuity. *Journal of the Optical Society of America. A, Optics and Image Science*, 7(2), 297-304.
- Nachmias, J. (1981). On the psychometric function for contrast detection. *Vision Research*, 21(2), 215-23.
- Nachmias, J., & Sansbury, R. V. (1974). Grating contrast: discrimination may be better than detection. *Vision Research*, 14(10), 1039-42.
- Parkes, L., Lund, J., Angelucci, a, Solomon, J. A., & Morgan, M. J. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, 4(7), 739-44.
- Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America. A, Optics and Image Science*, 2(9), 1508-32.
- Pelli, D. G., & Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research*, 31(7-8), 1337-50.
- Robitaille, N., & Harris, I. M. (2011). When more is less: extraction of summary statistics benefits from larger sets. *Journal of Vision*, 11(12), 1-8.
- Sweeny, T. D., & Whitney, D. (2014). Perceiving crowd attention: ensemble perception of a crowd's gaze. *Psychological Science*, 25(10), 1903-13.
- Teghtsoonian, M. (1965). The Judgment of Size. *The American Journal of Psychology*, 78(3), 392.
- Wannacott, T. H., & Wonnacott, R. J. (1981). *Regression a second course in statistics*. New York: Wiley.
- Watamaniuk, S. N., & Duchon, A. (1992). The human visual system averages speed information. *Vision Research*, 32(5), 931-41.
- Williams, D. W., & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research*, 24(1), 55-62.
- Yeshurun, Y., Carrasco, M., & Maloney, L. T. (2008). Bias and sensitivity in two-interval forced choice procedures: tests of the difference model. *Vision Research*, 48, 1837-1851.
- 1 차원고접수 : 2018. 01. 25
수정원고접수 : 2018. 04. 20
최종게재결정 : 2018. 04. 20

A computational model of mean size perception

Jongsoo Baek¹⁾

Sang Chul Chong²⁾³⁾

¹⁾Yonsei Institute of Convergence Technology, Yonsei University

²⁾Department of Psychology, Yonsei University

³⁾Graduate Program in Cognitive Science, Yonsei University

Human visual system represents a statistical summary of a complex visual input to overcome its limited capacity. An example of summary representation is the average size perception, which has been known to be accurate and precise. In the current study, we developed and validated a computational model of mean size perception. In this model, we assumed that the visual system encodes individual sizes with early noise, then integrates the noisy size information from multiple inputs. Finally, the integrated size information is added by late noise. The suggested model was validated with a psychophysical experiment, in which the standard and the test displays included multiple circles with different sizes and observers were asked to report which display had larger mean size. The psychophysical data was well accounted by the model with late noise: threshold for mean size discrimination was decreased with set-size, but the decrement of threshold was decelerated in large set-sizes. The proposed model allows us to understand underlying mechanism of mean size perception, and the experimental paradigm used in the current study is expected to be a useful tool for studying ensemble perception of various visual properties.

Key words : mean size perception, computational model, noisy percept, early noise, late noise, central limit theorem, signal detection theory

부 록

평균 크기 표상에 대한 분석 모형

초기 잡음. 시각 체계상의 내적 잡음들은 특정 변량을 가지고 있는 무선 변수이기 때문에, 어떤 자극에 대한 내적 반응은 매번 달라질 수 있다. 따라서, 이 단계에서 내적 반응 x 는 정규분포를 따르는 무선 변수 $g(x, \mu, \sigma_1)$ 로 기술될 수 있다.

정보 통합. 중심극한정리에 의해, n 개의 자극에 대한 평균 크기 표상은 n 이 증가함에 따라 표준편차가 줄어드는 정규분포상의 무선 변수가 된다. 내적 표상의 표준편차는 자극 개수 n 의 제곱근의 역함수가 된다. σ_1 을 개별 자극들에 대한 초기 잡음의 표준편차라고 하면, n 개의 자극들의 평균 크기 표상의 표준편차는 공식 1과 같이 정리된다.

$$\sigma = \frac{\sigma_1}{\sqrt{n}} \dots \text{공식 A1}$$

후기 잡음. σ_2 를 후기 잡음의 표준편차라고 할 때, 후기 잡음을 포함한 내적 평균 표상의 표준편차는 다음과 같이 정리될 수 있다.

$$\sigma = \sqrt{\frac{\sigma_1^2}{n} + \sigma_2^2} \dots \text{공식 A2}$$

의사 결정. 신호탐지이론은 시각 체계가 표준 자극 세트와 비교 자극 세트 대한 내적 표상의 정규 분포, $g(x_s, \mu_s, \sigma_s)$ 와 $g(x_t, \mu_t, \sigma_t)$ 를 만든다고 가정한다. 시각민감도를 나타내는 d' 은 신호대잡음비로 정의된다.

$$d' = \frac{\mu_t - \mu_s}{\sqrt{\frac{\sigma_t^2 + \sigma_s^2}{2}}} \dots \text{공식 A3}$$

이 때, 두 자극 세트가 동일한 개수의 자극을 포함하고 있으면, 즉 $\sigma = \sigma_s = \sigma_t$ 이면, 공식 A3은 다음과 같이 단순화될 수 있다.

$$d' = \frac{\mu_{diff}}{\sigma} \dots \text{공식 A4}$$

여기서 μ_{diff} 는 두 자극 세트 간의 물리적 평균 크기 차이, 즉, $\mu_{diff} = \mu_c - \mu_s$ 이다. 공식 A2와 공식 A4를 통해, d' 은 아래와 같이 재정리될 수 있다.

$$d' = \frac{\mu_{diff}}{\sqrt{\frac{\sigma_1^2}{n} + \sigma_2^2}} \dots \text{공식 A5}$$

모형 예측. 신호탐지이론에서 관찰자가 비교 자극 세트가 기준 자극 세트에 비해서 그 평균 크기가 더 크다고 반응할 비율 pc 는 민감도와 응답 편향에 의해서 결정되어진다. 응답 편향이 매우 작다고 알려진 2지 강제 선택 과제(2AFC)에서 응답자의 반응 비율 pc 는 아래의 공식으로 계산될 수 있다.

$$pc = \Phi\left(\frac{d'}{\sqrt{2}}\right) = \Phi\left(\frac{\mu_{diff}}{\sigma\sqrt{2}}\right) \dots \text{공식 A6}$$

여기서 $\Phi(x)$ 는 표준누적정규분포곡선을 나타낸다. 특정한 반응 비율 pc (예를 들면, 71%, 76%, 79% 혹은 84%)에 해당하는 역치는 아래의 공식으로도 계산되어질 수 있다.

$$\theta(pc_i) = G^{-1}(pc_i, 0, \sigma\sqrt{2}) \dots \text{공식 A7}$$

여기서 G^{-1} 은 평균이 0이고 표준편차가 $\sigma\sqrt{2}$ 인 역 누적정규분포함수를 의미한다.