

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ  
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ  
«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ  
«ВЫСШАЯ ШКОЛА ЭКОНОМИКИ»

На правах рукописи

Хвостов Владислав Александрович

**РОЛЬ СТАТИСТИЧЕСКИХ СВОЙСТВ РАСПРЕДЕЛЕНИЙ ЗРИТЕЛЬНЫХ  
ПРИЗНАКОВ В ВОСПРИЯТИИ И КАТЕГОРИЗАЦИИ  
МНОЖЕСТВЕННЫХ ОБЪЕКТОВ**

Резюме диссертации  
на соискание ученой степени  
кандидата психологических наук

Научный руководитель:  
Доктор психологических наук,  
Уточкин Игорь Сергеевич

Москва, 2022

## Оглавление

1. Введение.....	4
2. Параллельность и независимость зрительной обработки сводных статистик ансамблей .....	13
3. Сегментация множественных сочетаний признаков на основе их сводной статистики .....	21
4. Роль формы распределения в зрительном поиске по сочетанию признаков.....	29
5. Заключение .....	36
Список литературы .....	38

### **Статьи, выбранные для защиты:**

1. Khvostov, V.A., & Utochkin, I.S. (2019). Independent and parallel visual processing of ensemble statistics: Evidence from dual tasks. *Journal of Vision*, 19 (8): 3, 1-18. DOI: 10.1167/19.9.3  
*(Независимая и параллельная зрительная обработка статистик ансамблей: Доказательства, полученные с помощью двойной задачи).*
2. Utochkin, I.S., Khvostov, V.A., & Stakina Y.M. (2018). Continuous to discrete: Ensemble-based segmentation in the perception of multiple feature conjunctions. *Cognition*, 179, 178-191. DOI: 10.1016/j.cognition.2018.06.016  
*(Из непрерывности в дискретность: сегментация множественных сочетаний признаков, основанная на статистиках ансамблей).*
3. Utochkin, I.S., Khvostov, V.A., & Wolfe J.M. (2020). Categorical grouping is not required for guided conjunction search. *Journal of Vision*, 20(8): 30, 1-22. DOI: 10.1167/jov.20.8.30  
*(Управляемый поиск по сочетанию признаков не требует категориальной группировки).*

### **Исследование, связанное с темой диссертации, также описано в статье:**

4. Khvostov, V.A., Lukashevich, A.O., & Utochkin, I.S. (2021). Spatially intermixed objects of different categories are parsed automatically. *Scientific Reports*, 11: 377, 1-8. DOI: 10.1038/s41598-020-79828-4  
*(Пространственно перемешанные объекты разных типов категоризируются автоматически).*

Диссертация подготовлена на базе научно-учебной лаборатории когнитивных исследований НИУ «Высшая Школа Экономики».

## 1. Введение

### 1.1. Проблема исследования

Каждую секунду зрительная система сталкивается с большим количеством разнообразных объектов. Суровые ограничения сфокусированного внимания и рабочей памяти (Cowan, 2001; Luck & Vogel, 1997; Pylyshyn & Storm, 1988) препятствуют детальной обработке всех объектов (Wolfe et al., 2011). Однако люди обычно не испытывают трудностей с одновременным восприятием множества объектов. Одним из возможных решений этого парадокса восприятия является идея о том, что зрительная система репрезентирует некоторую общую статистическую информацию о наборах объектов без сохранения и репрезентации информации о каждом отдельном объекте (Alvarez, 2011; Ariely, 2001). Это может быть осуществлено с помощью вычисления статистических моментов для распределений признаков группы объектов (Whitney & Yamanashi Leib, 2018) – сводных статистик ансамблей. Было показано, что испытуемые могут извлекать первый и второй статистические моменты, что соответствует среднему значению зрительного признака (Alvarez & Oliva, 2008; Ariely, 2001; Bauer, 2009; Chong & Treisman, 2003, 2005b, 2005a) и его разбросу/дисперсии (Dakin & Watt, 1997; Morgan et al., 2008; Norman et al., 2015; Solomon et al., 2011). Также, испытуемые могут оценивать приблизительное количество объектов в сцене (Burr & Ross, 2008; Chong & Evans, 2011; Halberda, Sires, & Feigenson, 2006).

Большое количество зрительных признаков может быть репрезентировано в виде сводных статистик ансамблей: ориентация (Alvarez & Oliva, 2009; Dakin & Watt, 1997; Parkes, Lund, Angelucci, Solomon, & Morgan, 2001), цвет (Gardelle & Summerfield, 2011; Maule & Franklin, 2015), размер (Ariely, 2001; Chong & Treisman, 2003, 2005b, 2005a), движение (Watamaniuk & Duchon, 1992), даже эмоциональные выражения лиц (Haberma n & Whitney, 2007) и многие другие (Florey, Clifford, Dakin, & Mareschal, 2016; Leib, Kosovicheva, & Whitney, 2016; Sweeny & Whitney, 2014). Это создаёт надёжную основу для управления

человеческим поведением в разных ситуациях. Сводные статистики ансамблей скорее напрямую репрезентируются перцептивной системой, нежели вычисляются с помощью более высоких когнитивных функций. Свидетельства в пользу этого утверждения получены на материале эффектов последствия селективной адаптации к статистическим свойствам (Burr & Ross, 2008; Corbett et al., 2012; Norman et al., 2015; Ying & Xu, 2017). Вычисление сводных статистик ансамблей происходит очень быстро (50-200 мс, Chong & Treisman, 2003; Whiting & Oriet, 2011) и часто происходит в условиях ограниченного или отсутствующего доступа сознания к индивидуальным объектам (Alvarez & Oliva, 2008; Ariely, 2001; Corbett & Oriet, 2011; Parkes et al., 2001).

Последние исследования показывают, что зрительная система может репрезентировать даже полное распределение признаков объектов, представленных в сцене (Chetverikov et al., 2016, 2017a, 2017b, 2017c; Kim & Chong, 2020; Oriet & Hozempa, 2016). Эти открытия позволяют предположить, что такие простые сводные статистики как среднее, разброс и численность не являются единственными репрезентациями ансамблевой информации. Скорее наоборот, ансамблевая репрезентация хранит крайне богатую информацию обо всей группе объектов, в то время как среднее, разброс и численность рассчитываются на основе этой богатой репрезентации распределения признака (Khvostov et al., 2021).

Как зрительная система использует эту богатую информацию о полном распределении признаков? Может ли она быть использована для параллельного и независимого извлечения разных сводных статистик ансамблей? Какую роль эта информация может играть в выполнении повседневных когнитивных задач, например, зрительном поиске или быстрой зрительной категоризации множественных объектов? Данная работа суммирует исследования, проведённые мной (совместно с коллабораторами) для изучения этих важных вопросов.

## **1.2. Цели исследования**

1) Протестировать, могут ли несколько ансамблевых статистик вычисляться параллельно.

2) Исследовать, осуществляются ли вычисления нескольких ансамблевых статистик одним общим когнитивным механизмом, или они вычисляются независимо, несколькими разными механизмами.

3) Выяснить роль формы распределений признаков в процессе быстрой зрительной категоризации и сегментации множественных объектов, заданных сочетанием нескольких признаков.

4) Выяснить роль формы распределений признаков в зрительном поиске по одному признаку и по их сочетанию.

### **1.3. Методология и теоретические основания работы**

Данная работа основывается и развивает следующие теоретические идеи:

1) группы объектов и сцены репрезентируются зрительной системой в сжатой форме *сводных статистик ансамблей* (Alvarez, 2011; Ariely, 2001; Chong & Treisman, 2003; Whitney & Yamanashi Leib, 2018).

2) зрительная система использует информацию о распределении признаков ансамбля для осуществления быстрой зрительной категоризации множественных объектов – *теория быстрой зрительной категоризации и сегментации, основанной на сводных статистиках ансамблей* (Utochkin, 2015).

3) процесс зрительного восприятия может быть концептуально разделён на «глубокую» и «поверхностную» обработку, например, процессы внимания и предвнимания (Neisser, 1967; *Feature integration theory*: Treisman & Gelade, 1980), или сфокусированного и распределённого внимания (Treisman, 2006), неселективный и селективный путь обработки информации об объектах и сценах (Wolfe et al., 2011).

4) зрительная система использует результат работы «поверхностных» процессов (пр., предвнимания) для того, чтобы эффективно направлять работу

«глубоких» процессов (пр., внимания) – теория управляемого поиска (Wolfe, 1994, 2021).

#### **1.4. Гипотезы**

1) Зрительная система извлекает разные сводные статистики о группе объектов независимо (с помощью разных когнитивных механизмов) и параллельно (без взаимной интерференции при одновременном протекании этих процессов).

2) Зрительная система использует форму распределений признаков в качестве опорной информации для быстрой зрительной категоризации и сегментации множественных объектов, заданных сочетанием нескольких признаков. Чем больше признаков имеют распределение, состоящее из нескольких пиков («сегментабельное» распределение), тем более успешным становится процесс категоризации.

3) Форма распределений признаков оказывает влияние на скорость зрительного поиска: наличие отдельных пиков в распределении («сегментабельное» распределение) улучшает выполнение задачи поиска.

#### **1.5. Методы исследования**

Все исследования в данной работе являются строго контролируруемыми психофизическими экспериментами. В различных исследованиях, были использованы разные задачи и парадигмы для сбора ответов. В Исследовании 1 была использована задача оценки сводных статистик ансамблей с помощью метода подравнивания, в Исследовании 2 – задача сегментации текстур вместе с двух-альтернативным вынужденным выбором (2AFC), в Исследовании 3 – зрительный поиск с методом «Да-Нет». Статистический анализ был проведён в программе JASP (JASP 0.9.0.1; JASP, Амстердам, Нидерланды) и R. Процедуры статистического анализа включали в себя серию дисперсионных анализов с повторными измерениями с последующими попарными t-тестами и корреляции

Пирсона.

## **1.6. Научная новизна**

Большинство предыдущих работ в области сводных статистик ансамблей были сфокусированы на исследовании вопросов об извлечении изолированных статистик (в основном, среднего значения) для одного признака (например, размера). Данная работа тестирует и развивает идею о том, что зрительная система имеет доступ к гораздо более богатой информации о полном распределении различных признаков объектов, а не просто к отдельным статистикам по одному признаку. Эта богатая информация используется в различных повседневных перцептивных задачах.

1) Мы исследовали как зрительная система справляется с ситуациями, в которых ей необходимо извлечь несколько сводных статистик для группы объектов одновременно. Нас интересовало два вопроса: (1) извлекает ли зрительная система разные статистики с помощью независимых когнитивных механизмов, или используется один общий механизм? (2) Может ли зрительная система извлекать несколько сводных статистик параллельно (без взаимной интерференции)? Мы разработали новую парадигму для сбора ответов – вариацию двойной задачи, применённую к вычислению ансамблевых статистик, которая позволила нам протестировать эти два связанных вопроса в одной задаче. Также, нами был разработан новый способ анализа данных для проб с двойной задачей, основанный на внутри-индивидуальной корреляции. Данный способ может прямо ответить являются ли ошибки при отчёте о двух статистиках скоррелированными между собой, что увеличивает мощность данного анализа по сравнению с предыдущими парадигмами. Используя эти методы, мы обнаружили новые убедительные доказательства в пользу того, что зрительная система вычисляет разные ансамблевые статистики независимо и параллельно.

2) Предыдущие работы показали, что распределение одного признака оказывает сильное влияние на выполнение различных задач, связанных с



восприятием множественных объектов (e.g., Chetverikov et al., 2016; Utochkin & Yurevich, 2016; Im et al., 2021). Нами впервые было протестировано, как зрительная система использует распределения нескольких признаков для таких задач. Мы обнаружили условия, в которых зрительная система может использовать несколько признаков одновременно для быстрой зрительной категоризации множественных объектов, заданных сочетанием этих признаков. Наши исследования показали, что успешная категоризация возможно только если оба распределения признаков имеют несколько отдельных пиков: например, только большие и только маленькие размеры, без промежуточных. Мы предложили новый механизм, который объясняет данные результаты, и получили первые эмпирические свидетельства в его пользу.

3) Впервые, нами было исследовано то, как распределения нескольких признаков влияют на работу внимания при зрительном поиске цели, заданной сочетанием признаков. Наше исследование обнаружило условия, когда они не оказывают никакого эффекта, и когда они сильно влияют на скорость и эффективность зрительного поиска. Мы обнаружили свидетельства в пользу того, что зрительная система может осуществлять эффективный зрительный поиск даже при отсутствии отдельных пиков в распределении признаков у дистракторов, то есть в отсутствии группировки среди дистракторов. Мы показали, что теория управляемого поиска (Wolfe, 1994, 2021) может идеально объяснять полученные результаты. Таким образом, наша работа прокладывает новую важную связь между теорией быстрой зрительной категоризации и сегментации на основе сводных статистик ансамблей (Utochkin, 2015) и классическими теориями зрительного поиска, в частности, теорией управляемого поиска (Wolfe, 1994, 2021). Используя теорию управляемого поиска, мы смогли предсказать очень контринтуитивный результат: при определённых характеристиках распределения, зрительный поиск по одному признаку может

быть более долгим и менее эффективным, чем зрительный поиск по сочетанию двух признаков. Впервые, мы протестировали и подтвердили данное предсказание.

### **1.7. Теоретическая значимость**

Данная работа вносит вклад в развитие целой группы современных важных теорий зрительного восприятия и внимания, таких как теория сводных статистик ансамблей (Alvarez, 2011; Whitney & Yamanashi Leib, 2018), теория быстрой зрительной категоризации и сегментации (Utochkin, 2015) и теория управляемого поиска (Wolfe, 1994, 2021). Также, она способствует улучшению нашего понимания архитектуры зрительных репрезентаций: какая информация о группе объектов доступна зрительной системе, и как она её использует.

### **1.8. Прикладная значимость**

Результаты данной работы могут быть использованы для оптимизации способа визуализации информации. Например, как следует пометить разные объекты, чтобы быть уверенным, что люди легко смогут сегментировать одну группу от другой (даже если объекты из разных групп перемешаны в пространстве) или найти интересующий объект среди остальных. Также, сводные статистики ансамблей – это область, связанная с математической статистикой, поэтому результаты могут быть использованы для улучшения преподавания математической статистики с помощью более интуитивно понятных примеров из зрительной статистики. Некоторые результаты уже используются в преподавании бакалаврских курсов «Когнитивная психология» и «Психология и нейрофизиология восприятия и внимания», в НИУ «Высшая школа экономики».

### **1.9. Положения, выносимые на защиту**

1) Несколько ансамблевых статистик могут вычисляться параллельно: то есть без взаимной интерференции. Извлечение различных статистик осуществляется с помощью разных когнитивных механизмов, которые осуществляют независимые друг от друга вычисления.

2) Зрительная система использует форму распределения признаков как

опорную информацию для осуществления быстрой зрительной категоризации и сегментации множественных объектов, заданных сочетанием нескольких признаков. Обязательным условием данной категоризации является «сегментабельность» обоих распределений (наличие нескольких отдельных пиков, соответствующих вероятным категориям). Возможный механизм для такой категоризации – стратегия половинного деления: выделение объектов из одной половины распределения первого признака и сравнение средних значений по второму признаку.

3) Зрительная система способна преодолевать отсутствие «сегментабельности» распределений дистракторов и осуществлять эффективный зрительный поиск известной цели, заданной сочетанием двух признаков. Однако зрительная система не может избежать влияния сегментабельности при зрительном поиске по одному признаку или поиске по сочетанию признаков, когда один из признаков цели неизвестен: скорость и эффективность поиска снижаются в случае несегментабельных распределений дистракторов.

### **1.10. Сбор данных**

Все три статьи, выбранные для защиты, описывают серии психофизических экспериментов. Данная работа включает в себя 12 лабораторных экспериментов, проведённых в НУЛ когнитивных исследований (НИУ ВШЭ, Москва, Россия) и Лаборатории зрительного внимания (Brigham and Women's Hospital and Harvard Medical School, Бостон, США). В общей сложности, более 200 испытуемых приняло участие в экспериментах.

### **1.11. Апробация работы**

Результаты работы были представлены в виде 5 устных и 6 постерных докладов на 8 международных и российских конференциях: встрече сообщества наук о зрении (Vision Sciences Society 2018, 2019), европейской конференции по зрительному восприятию (European Conference on Visual Perception 2018, 2019) и многих других.

## **1.12. Вклад автора**

Автор принимал непосредственное участие в проведении всех исследований, описанных в диссертации: разрабатывал идеи экспериментов, генерировал стимулы, программировал и проводил эксперименты, анализировал и интерпретировал данные, презентовал результаты на научных конференциях, а также писал тексты статей.

## **2. Параллельность и независимость зрительной обработки сводных статистик ансамблей**

Статья для защиты: Khvostov & Utochkin (2019)

Во введении мы описали множество исследований, показывающих, что испытуемые могут использовать информацию о различных признаках множественных объектов для получения информации о различных сводных статистиках: среднего значения признаков, их дисперсии/разброса и численности (например, Ariely, 2001; Burr & Ross, 2008; Chong & Treisman, 2003; Morgan et al., 2008). В большей части этих работ изучалась способность извлекать одну сводную статистику в изоляции от остальных (например, средний размер), что оставляет без ответа вопрос об функциональном взаимодействии разных статистик. В текущей работе (Khvostov & Utochkin, 2019) мы разделили этот вопрос на несколько. (1) Независимость: вычисляются ли разные ансамблевые статистики одним когнитивным механизмом («общий статистик»), или же они вычисляются независимыми механизмами? (2) Параллельность: насколько эффективно происходит координация нескольких сводных статистик для получения доступа к сознательному восприятию: могут ли они вычисляться параллельно, без затрат на разделение внимания между ними? Или вычисление двух статистик неизбежно приводит к взаимной интерференции между ними?

Оба вопроса в некоторой степени исследовались ранее с использованием двух основных подходов к изучению процессов восприятия и познания (Khvostov & Utochkin, 2019). Вопрос о независимости обычно изучается с помощью подхода, основанного на индивидуальных различиях (Huang et al., 2012; Underwood, 1975; Wilmer, 2008): исследователи оценивают кросс-индивидуальные корреляции между успешностью выполнения ряда задач группой испытуемых. Наличие корреляции свидетельствует в пользу общего источника шума (дисперсии), что подразумевает общий механизм, задействованный в обеих задачах. Отсутствие корреляции означает обратное: разные источники шума для

13

задач, и как следствие, разные когнитивные механизмы. Используя этот подход, Янг и соавторы (Yang et al., 2018) обнаружили отсутствие кросс-индивидуальных корреляций между успешностью выполнения задач на вычисление среднего и дисперсии (как для размера, так и для ориентации). Они пришли к выводу, что эти две статистики рассчитываются независимыми механизмами. Аналогичный подход использовался в исследовании Ли и коллег (Lee et al., 2016), которые измеряли точность оценок среднего размера кругов, их количества и общей площади. Их результаты показали, что средний размер и численность рассчитываются независимо друг от друга (тогда как общая площадь может быть вычислена на основе этих двух параметров).

Параллельность различных мыслительных операций часто исследуется с помощью парадигмы «пред/пост-подсказки» (Khvostov & Utochkin, 2019). Испытуемым предъявляется пара целей (объектов, наборов и т. д.), об одной из которых они должны сообщить какую-то информацию. В условии пред-подсказки испытуемые заранее знают, о какой цели они должны отчитаться, и могут задействовать все ресурсы внимания для обработки этой цели. Напротив, в условии пост-подсказки испытуемые не имеют предварительной информации, о какой цели им необходимо будет отчитаться. Они получают эту информацию только после исчезновения целей, поэтому вынуждены распределять внимание между двумя целями во время презентации. Если успешность выполнения при пред-подсказке выше, чем при пост-подсказке, это означает, что в какой-то момент обработки два процесса взаимодействуют и конкурируют за ограниченный ресурс внимания. Если успешность выполнения одинаковая в этих условиях, то можно сделать вывод, что два процесса, лежащие в основе обработки обеих целей, могут выполняться параллельно. Параллельность вычислений среднего и численности изучалась в исследовании Уточкина и Вострикова (Utochkin & Vostrikov, 2017). В эксперименте, в котором испытуемые должны были вычислить обе статистики для одного набора кругов, не было обнаружено

различий между условиями пред- и пост-подсказки. На основании этого авторы пришли к выводу, что численность и среднее значение могут вычисляться параллельно. Они также исследовали независимость этих расчетов, используя подход, основанный на индивидуальных различиях. Основываясь на отсутствии корреляций между испытуемыми, Уточкин и Востриков (Utochkin & Vostrikov, 2017) заключили, что расчеты среднего и численности выполняются независимыми механизмами.

В нашем исследовании мы хотели проверить выводы, полученные в прошлых исследованиях, а также усилить их путем проверки независимости и параллельности вычисления различных сводных статистик ансамблей с использованием нового, более точного метода — двойной задачи. Описанная ранее парадигма пред/пост-подсказки подразумевает, что испытуемые действительно пытаются разделить свое внимание между двумя статистиками в условии пост-подсказки (не используя при этом альтернативных стратегий). Мы разработали парадигму двойной задачи, которая позволяет не полагаться на это предположение и напрямую проверить, могут ли испытуемые вычислить две сводные статистики одновременно. Двойная задача требует от испытуемых отчитываться об обеих статистиках в каждой пробе (в отличие от условия пост-подсказки, в которой они должны сообщать только одну случайную выбранную статистику в каждой пробе). Парадигма двойной задачи также позволяет нам проверить независимость не только с помощью косвенных корреляций между испытуемыми, но и более прямо - с помощью корреляции между двумя ответами в одной и той же пробе. Если две статистики рассчитываются с помощью разных механизмов (независимо), точность двух ответов в одной и той же пробе не должна коррелировать друг с другом. В противоположном случае мы ожидаем получить положительную корреляцию: если в конкретной пробе «общий статистик» неточно извлек статистическую информацию, это должно оказывать одинаковое влияние на оба ответа в пробе (обратное также верно). Еще одна

причина использования данного корреляционного анализа связана с тестированием параллельности: он может выявить стратегию непараллельного распределения ресурсов. Если испытуемые не могут рассчитать обе статистики одновременно, они могут сосредоточиться на статистике № 1 в одной пробе и на статистике № 2 - в другой пробе. Если это так, то мы должны получить сильную отрицательную корреляцию между ответами в одной и той же пробе. Ещё одной важной целью исследования было расширение ранее полученных выводов. Ранее, отношения между разными парами статистик среднее-численность и среднее-диапазон изучались в разных исследованиях с использованием разных методов и стимулов. В данной работе мы тестировали обе пары в одном исследовании с использованием одного и того же метода и стимулов.

В эксперименте 1 ( $N = 24$ ) мы проверяли функциональную связь между восприятием среднего размера и численностью. Эксперимент состоял из нескольких блоков. В блоке с двойной задачей испытуемым на 500 мс предъявлялся набор кружков, число которых варьировалось от 7 до 36. После 200 мс пустого экрана они должны были сообщить средний размер кругов, а затем количество кругов (в другой половине проб порядок был обратным). Для отчёта о среднем размере они должны были отрегулировать зондовый круг с помощью колесика мыши. Для отчёта о численности – подравнять числовое значение зондового стимула. Для получения данных о базовом уровне успешности выполнения этих задач испытуемые также проходили два блока с одиночной задачей. Единственное различие между одиночной и двойной задачами заключалось в том, что в первом случае требовалось сообщать только одну заранее определённую статистику.

Основной зависимой переменной в данном эксперименте была точность подравнивания, то есть нормализованная абсолютная ошибка, которая рассчитывалась по следующей формуле:  $\text{Ошибка} = |\text{Ответ испытуемого} - \text{Правильный ответ}| / \text{Правильный ответ}$ . Анализ данных состоял из трёх частей. Во-

16



первых, мы исследовали независимость вычисления разных статистик с помощью вычисления корреляции между ошибками в двух отчётах в рамках одной пробе (в блоке с двойной задачей) для каждого испытуемого в отдельности. Для 19 из 20 испытуемых данной корреляции обнаружено не было. Мы также провели данный анализ для ошибок, вычисленных с учётом знака (чтобы проверить, не приводит ли недооценка количества к завышению среднего, как предсказывает формула среднего из обычной статистики): все 20 испытуемых показали отсутствие такой корреляции. Во-вторых, мы провели кросс-корреляционный анализ средних ошибок в отчетах о среднем размере и численности по всем испытуемым. Поскольку у нас было три измерения средней ошибки для каждой статистики, мы проводили три корреляционных анализа: между вычислением среднего размера и численностью в одиночных задачах, между первыми ответами в двойном задании и между вторыми ответами в двойном задании. Все три анализа не выявили наличия корреляции. Важно, что автокорреляции для каждой ансамблевой статистики, измеренной в разных условиях (например, средний размер в одиночной и двойной задаче) были очень высокими. Это свидетельствует о том, что параметры нашего эксперимента (размер выборки и т. д.) позволяют обнаружить корреляцию, если она существует. Кроме того, это свидетельствует в пользу надежности наших измерений. Таким образом, как на макроуровне (между испытуемыми), так и на микроуровне (между ответами из одной пробы для каждого испытуемого в отдельности) результаты показали, что средний размер и численность рассчитываются с помощью независимых когнитивных механизмов. Третий анализ был связан с проверкой параллелизма: мы сравнивали точность выполнения в двойной задаче с соответствующими измерениями в одиночных задачах. Данный анализ показал, что ошибка при первом отчёте в двойной задаче не отличалась от ошибки в одиночном задании (как для численности, так и для среднего размера). Точность второго ответа в двойной задаче была хуже, чем точность в одиночной задаче и точность первого ответа в двойной задаче.

Вероятно, это связано с процессами интерференции, происходящими в рабочей памяти на стадии припоминания. Итого, результаты показывают отсутствие значительных затрат на разделение внимания между двумя сводными статистиками, что означает, что средний размер и численность могут вычисляться параллельно.

Проведение эксперимента 2 ( $N = 19$ ) было необходимо для расширения результатов первого эксперимента на другую пару сводных статистик ансамблей: среднего размера и диапазона/дисперсии. Дизайн и процедура этих двух экспериментов были схожими, за исключением того, что отчет о количестве был заменен отчётом о диапазоне. Каждый дисплей всегда состоял из шестнадцати кругов, но средний размер и диапазон их размеров менялись от пробы к пробе. Для отчёта о диапазоне испытуемые должны были использовать другой набор из 16 кругов с фиксированным средним размером. Вращение колеса мыши увеличивало или уменьшало разнообразие набора кругов.

Нами было проведено три анализа для проверки независимости и параллельности расчетов среднего размера и диапазона (такие же, как и в предыдущем эксперименте). Как и в эксперименте 1, все испытуемые показали отсутствие корреляции между двумя ответами в двойной задаче (как для абсолютных ошибок, так и ошибок с учётом знака). Кросс-корреляции между средними ошибками для успешности вычисления среднего размера и диапазона были крайне низкими и незначимыми, в то время как автокорреляции для оценок одной и той же статистики в разных условиях были высокими. Эти результаты предоставляют убедительные доказательства в пользу независимости расчетов среднего размера и диапазона. Отметим, что полученное нами отсутствие корреляции теоретически однозначно свидетельствует в пользу независимых когнитивных механизмов, в то время как наличие корреляции могло бы иметь альтернативные объяснения: либо две задачи реализуются одним и тем же механизмом, либо общий механизм критически влияет на два разных механизма,

18

реализующих эти задачи. Анализ параллельности так же показал результаты, аналогичные эксперименту 1: средняя ошибка для второго ответа в двойном задании была хуже, чем для первого ответа или для одиночной задачи (вероятно, из-за интерференции в памяти из-за последовательного порядка отчётов). При этом средняя ошибка для первого ответа в двойной задаче была равна ошибке в одиночной. Отсутствие затрат на разделение внимания между расчетами среднего и диапазона позволяет сделать вывод, что эти два сводных показателя рассчитываются параллельно.

В целом, мы подтвердили результаты предыдущих работ (Lee et al., 2016; Utochkin & Vostrikov, 2017; Yang et al., 2018) и предоставили более убедительные доказательства в пользу независимости и параллельности вычислений нескольких сводных статистик ансамблей. Независимость вычислений позволяет сделать два важных вывода о восприятии ансамблей. Во-первых, разные источники шума при расчетах различных ансамблевых статистических данных могут рассматриваться как свидетельство того, что средний размер, численность и дисперсия могут быть рассчитаны с помощью различных когнитивных механизмов, которые могут относиться к различным (или, по крайней мере, частично непересекающимся) нейронным сетям. Например, независимость между механизмами вычисления численности и другими статистическими расчетами можно проиллюстрировать исследованиями, показывающими, что реализация функций, связанных с численностью, связана с активностью теменной коры (особенно внутритеменной борозды (Dehaene et al., 2003; Nieder & Dehaene, 2009)) и префронтальной коры (Nieder & Dehaene, 2009), в то время как обработка информации о статистике, связанной с формой (то есть не связанной с численностью) связана с активностью в парагиппокампальной области и латеральной затылочной области (Cant & Xu, 2012). Важно отметить, что утверждение о различных нейронных сетях является предположением и должно быть прямо проверено в отдельном нейрофизиологическом исследовании. Во-вторых, независимость означает, что

механизмы, реализующие вычисления разных статистик, не влияют друг на друга и не используют результаты работы друг друга для собственных вычислений. Данный факт может показаться контринтуитивным, потому что обычная статистика учит прямо противоположному. Например, чтобы для вычисление среднего нужно просуммировать все элементы и разделить на их количество, т. е. для вычисления среднего необходимо использовать результаты вычисления численности по определению (то же самое можно сказать и о вычислении дисперсии). Наше исследование показывает, что это не относится к вычислению зрительных статистик: каким-то образом зрительная система вычисляет средний размер группы объектов, не используя информацию о их количестве.

С этой точкой зрения согласуется и параллелизм вычислений разных статистик: расчёт разных статистик с помощью непересекающихся механизмов упрощает выполнение нескольких вычислений одновременно, уменьшая или предотвращая интерференцию между ними. Полученные нами результаты также согласуются с идеей о том, что при восприятии множественных объектов зрительная система строит достаточно богатую суммарную репрезентацию, отражающую распределение их признаков (Chetverikov et al., 2016, 2017a, 2017b, 2017c; Kim & Chong, 2020; Oriet & Hozempa, 2016). Эта богатая репрезентация может быть использована для извлечения различных сводных статистик ансамблей: среднего значения, численности, дисперсии. Как еще можно использовать эту репрезентацию? В следующих исследованиях мы изучали вопрос о том, как зрительная система использует информацию о распределении признаков для таких повседневных зрительных задач как быстрая зрительная категоризация и зрительный поиск.

### **3. Сегментация множественных сочетаний признаков на основе их сводной статистики**

Статья для защиты: Utochkin, Khvostov, & Stakina (2018)

Если зрительная система имеет доступ к полному распределению признаков объектов, предъявленных в сцене (Chetverikov et al., 2016, 2017a, 2017b, 2017c; Kim & Chong, 2020; Oriet & Hozempa, 2016), то эта репрезентация должна использоваться не только для прямой оценки сводных статистик ансамблей, но также и для выполнения повседневных когнитивных задач. Одна из идей заключается в том, что зрительная система использует информацию о распределении признаков как основу для быстрой зрительной категоризации групп объектов (Utochkin, 2015; Utochkin & Tiurina, 2014). Этот процесс можно проиллюстрировать следующим примером. Если вы смотрите футбольный матч, то сразу можете заметить наличие двух визуально отличающихся группы игроков: например, одна команда - в зеленом, а другая - в красном. Зрительная система легко производит эту быструю зрительную категоризацию, поскольку имеет доступ к распределению цвета объектов в зрительном поле. Это распределение имеет два четких пика (красный и зеленый), каждый из которых соответствует игрокам одной из команд. Такой тип распределения мы будем называть сегментабельным. Согласно теории быстрой сегментации и категоризации (Utochkin, 2015), такое распределение ведёт к успешной категоризации и сегментации множества объектов на несколько подмножеств. Другой тип распределения (называемый несегментабельным) состоит из одного широкого пика или является равномерным, что способствует восприятию всех элементов как принадлежащими к единому набору, а не разным категориальным подмножествам. Если вернуться к футбольному примеру, несегментабельное распределение может быть проиллюстрировано множеством игроков в футболках разного цвета: красного, желтого, оранжевого, зеленого и их оттенков. В этом случае мы не увидим сегментированных подмножеств — только одну

разношерстную группу игроков.

Теория быстрой сегментации и категоризации была проверена в исследовании с использованием задачи зрительного поиска, в котором испытуемые искали отличающуюся цель среди дистракторов с разным распределением признаков. Когда дистракторы имели сегментабельное распределение (например,  $0^\circ$ ,  $22.5^\circ$  и  $45^\circ$ ), цель (например,  $-45^\circ$ ) находилась медленнее, чем в случае несегментного распределения ( $0^\circ$ ,  $5^\circ$ ,  $10^\circ$ , ...,  $45^\circ$ ). Этот результат объяснялся тем, что в сегментабельном случае дистракторы группируются в несколько подмножеств, которые испытуемые должны последовательно проверять, чтобы отбросить их как нецелевые. В случае несегментабельного распределения все дистракторы воспринимались как принадлежащие к одной группе и могли быть отвергнуты сразу. Важность формы распределения для эксплицитной быстрой зрительной категоризации также была показана с использованием различных версий ансамблевых задач, требующих сводных статистических суждений для подмножеств (Im et al., 2021).

Во всех вышеупомянутых исследованиях изучалась роль распределения признаков в быстрой категоризации и сегментации по одному признаку. Однако в реальном восприятии множественные объекты редко демонстрируют вариации, группировку или сегментацию только по одному признаку. Часто объекты различаются по многим признакам, образуя индивидуальные сочетания. В текущем исследовании мы хотели проверить, играет ли форма распределения важную роль в процессе сегментации и категоризации множественных объектов, задаваемых комбинацией нескольких признаков (т. е. сочетаниями признаков).

Мы использовали задачу сегментации текстур, в которой испытуемым была предъявлена матрица  $8 \times 8$ , состоящая из белых линий разной длины и ориентации (всего 64). Мы осуществляли манипуляцию распределением этих двух признаков. Оно могло быть как сегментабельным: распределение состояло только из двух крайних значений признака (пример для ориентаций: линии, с ориентацией  $11^\circ$  и

86°) – так и несегментабельным: когда распределение состояло из двух крайних и множества промежуточных значений между ними (линии с ориентацией от 11° до 86° с шагом 5°). Поле из 64 ячеек было разделено пополам невидимым меридианом, который мог быть как горизонтальным, так и вертикальным. Каждая половина содержала линии с одинаковыми распределениями длины и ориентации (по отдельности), в то время как соединения этих элементов распределялись по-разному, обеспечивая различные уровни корреляции длины и ориентации. В одной половине текстуры ориентация и длина коррелировали положительно: более длинные линии были также более плоскими (и наоборот) — в другой половине они коррелировали отрицательно: более длинные линии имели также более крутую ориентацию. Мы предъявляли эти линии на 200 мс, после чего показывали маску (зашумленный набор перекрывающихся белых линий разной ориентации и длины) на 200 мс. Затем испытуемый должен был ответить, была ли граница между двумя группами линий с разным знаком корреляции горизонтальной (различались верхняя и нижняя половины) или вертикальной (различались левая и правая половины). Таким образом, мы проверяли, могут ли испытуемые быстро воспринимать линии с разными знаками корреляции как объекты, принадлежащих к разным группам, т.е. категоризировать их.

В первом эксперименте ( $N = 5$ , опытные испытуемые) мы сравнивали успешность выполнения задачи сегментации текстур в двух условиях: сегментабельном (распределения ориентации и длины были сегментабельными) и несегментабельном (оба распределения были несегментабельными). Кроме того, мы варьировали коэффициент корреляции длины и ориентации в текстурах — он изменялся от  $-1.00$  до  $1.00$  с шагом  $0.25$ . Это было сделано с помощью изменения пропорции каждого сочетания признаков. Например, корреляция, равная  $-0,25$ , обеспечивалась тем, что  $5/8$  длинных линий были более крутыми, а  $3/8$  длинных линий более плоскими. Мы строили психометрические функции (ось абсцисс — коэффициент корреляции, ось ординат — доля правильных ответов) для каждого

испытуемого, отдельно для сегментабельного и несегментабельного условий. Для этого были использованы нормальные кумулятивные функции плотности вероятности. Мы анализировали  $\sigma^2$  — дисперсию нормального распределения, характеризующую различимость стимулов. Основным результатом эксперимента было то, что все пять испытуемых показали гораздо меньшую  $\sigma^2$  (лучшее различение) в сегментабельном условии по сравнению с несегментабельным. Это означает, что наличие отчетливых пиков в распределениях ориентации и длины действительно обеспечивает лучшую категоризацию: испытуемые могут более четко видеть границу между двумя областями с разными типами объектов.

В эксперименте 2 ( $N = 21$ ) мы хотели более подробно изучить влияние сегментабельности на успешность быстрой зрительной категоризации. Во-первых, мы протестировали несколько длительностей предъявления стимула (от 100 до 500 мс), чтобы выявить, какие процессы стоят за этой сегментацией текстуры: процессы, связанные с медленным локальным сфокусированным вниманием (Myczek & Simons, 2008) или с быстрым глобальным распределенным вниманием (Chong & Treisman, 2003). Во-вторых, мы хотели узнать, как эффект сегментабельности возникает из более базовых свойств сегментабельности отдельных распределений признаков. Достаточно ли сегментабельности одного признака, или же этот эффект возникает только в том случае, если сегментабельны оба распределения? Таким образом, мы манипулировали сегментабельностью распределения каждого признака по отдельности и, вместо двух условий сегментабельности, их было четыре: «оба» (распределения ориентации и длины были сегментабельными), «ориентация» (ориентация была сегментабельной, длина — несегментабельной), «длина» (длина была сегментабельной, ориентация — несегментабельной), «ни одно» (оба распределения были несегментабельными). В этом эксперименте мы использовали только две, крайние корреляции (1,0 и -1,0) и вычисляли меру чувствительности к ориентации границы текстуры ( $d'$ , как в теории обнаружения



сигнала). Результаты показали, что эффект сегментабельности возникает довольно рано (200 мс) и не меняется при более длительном времени предъявления. Это позволяет заключить, что процессы, связанные с обработкой нескольких объектов медленным сфокусированным вниманием, играют маленькую роль в этом эффекте. Этот результат больше согласуется с быстрыми процессами распределенного внимания, что соответствует наиболее популярному взгляду на сводные статистики ансамблей (например, Chong & Treisman, 2005a). Второй важный результат заключается в том, что только условие, в котором распределения длины и ориентации были сегментабельными, обеспечивало довольно хорошую успешность выполнения задания ( $d'$  около 0,8-0,9), тогда как в остальных условиях успешность была намного хуже ( $d'$  ниже 0,3). Это позволяет заключить, что сегментабельности только одного признака недостаточно, чтобы обеспечить хорошую основу для категоризации множественных соединений. Мы вернёмся к этому результату в эксперименте 4.

Эксперимент 3 ( $N = 23$ ) был посвящен более прямой проверке того, отражает ли эффект сегментабельности работу глобальных процессов, связанных с восприятием ансамблей или локальных процессов сфокусированного внимания. Вместо того, чтобы анализировать все предъявленные объекты, испытуемые могли выполнять задачу сегментации текстур, используя стратегию, основанную исключительно на работе сфокусированного внимания (хотя это и противоречит результатам эксперимента 2, в которых эффект сегментабельности проявлялся уже на 200 мс). Данная стратегия состоит в том, чтобы сравнить два соседних элемента из разных половин вдоль одной из границ (например, горизонтальной) — если они «согласуются» (например, две длинные вертикальные линии или одна длинная вертикальная и одна короткая горизонтальная линии), испытуемому следует ответить, что граница противоположная (т.е. вертикальна), если они не «согласуются» — граница горизонтальная. Чтобы проверить действительно ли испытуемые используют эту стратегию, мы сравнили успешность выполнения

задачи в условии, в котором мы предъявляли полную текстуру («полнотекстурное» условие: так же, как и в предыдущем эксперименте), и в условии, в котором мы предъявляли только линии вдоль горизонтального и вертикального меридианов матрицы («приграничное» условие). Если испытуемые используют сфокусированное внимание для выполнения задачи, их успешность не должна снижаться в приграничном условии, поскольку они используют только два элемента вблизи одной из границ. Но если в выполнении задания задействованы глобальные процессы, производительность в полнотекстурном условии должна быть выше, так как мы предоставляем испытуемым больше статистической информации. Результаты показали преимущество полнотекстурного над приграничным условием при всех условиях сегментабельности, что является еще одним аргументом в пользу объяснения эффекта сегментабельности глобальными процессами распределенного внимания, а не локальными процессами, связанными со сфокусированным вниманием.

Если сфокусированное внимание не играет существенной роли в быстрой категоризации множественных соединений, как испытуемые выполняют эту задачу? Мы предполагаем, что испытуемые задействуют глобальные механизмы распределенного внимания для реализации стратегии, которую мы называем «половинным разделением»: испытуемые пытаются выбрать все объекты из одной половины первого распределения признаков, а затем сравнить группы выбранных объектов по второму признаку, чтобы найти различие в среднем значении признака. Например, испытуемый может выделить подмножество всех длинных линий и проверить, где есть разница между средними ориентациями. Эта глобальная ансамблевая стратегия может быть довольно быстрой и создавать впечатление различий между разными половинами текстуры без необходимости вычисления полных корреляций. Сегментация признаков может облегчить этот процесс двумя способами. Во-первых, за счет упрощения процесса селекции: намного проще выделить все длинные линии, если они имеют одинаковую длину,

а все линии, которые нужно игнорировать – одинаково короткие. Во-вторых, облегчение процесса сегментации текстур достигается за счет увеличения разницы средних и уменьшения вариативности внутри групп по второму признаку. В сегментабельном случае разница средних ориентаций между двумя участками составляет  $75^\circ$  (дисперсия равна 0, поскольку все линии имеют одинаковую ориентацию), в несегментабельном случае разница средних равна  $40^\circ$  (а дисперсия составляет  $12^\circ$ ). Эти факты можно использовать для объяснения результатов эксперимента 2, согласно которым сегментабельности только одного признака недостаточно для существенного повышения успешности выполнения задачи быстрой категоризации.

Эксперимент 4 (N=16) был посвящен проверке гипотезы о половинном разделении. Мы искусственно моделировали ситуацию идеальной селекции по одному признаку и предъявляли испытуемым текстуры из эксперимента 2, удалив из исходных стимулов половину элементов с признаками, взятыми из одной половины распределения ориентации или длины. Таким образом, сегментация текстур превращалась в простое сравнение участков по одному признаку. Будет ли успешность выполнения задания выше без необходимости половинного деления и селекции по первому признаку? И повысится ли в этом случае успешность при сегментабельности второго признака? Результаты показали, что успешность сегментации была намного выше, чем во всех предыдущих экспериментах ( $d' = 1,12$  в эксперименте 4 против  $0,345$  в эксперименте 2 при времени предъявления равном 200 мс). Кроме того, мы обнаружили эффект сегментабельности для второго признака: если половинное деление было сделано на основе длины (представлены только более длинные линии), сегментабельность ориентации резко улучшила успешность выполнения задания (и наоборот). Таким образом, результаты этого эксперимента свидетельствуют в пользу стратегии половинного деления. Две причины плохой сегментации текстур, предсказанные этой стратегией, были эмпирически проверены. Во-

первых, у испытуемых действительно возникают проблемы с корректным выделением объектов из одной половины распределения первого признака: когда этот отбор сделан искусственно или когда этот признак сегментабельный, успешность выполнения сильно повышается. Во-вторых, сравнение средних значений по второму признаку также может быть проблемой, так как сегментабельность этого признака повышает успешность сегментации.

Четыре эксперимента в рамках данного исследования показали, что зрительная система может использовать информацию о форме распределения для быстрой зрительной категоризации множественных объектов. Это может происходить, даже когда группы объектов заданы сочетанием нескольких признаков. Испытуемые, вероятно, используют стратегию половинного разделения, чтобы выбрать группу объектов из одной половины распределения первого признака и сравнить эти выбранные объекты по второму признаку. Такое свойство распределения, как сегментабельность (наличие нескольких пиков с большим промежутком между ними), делает этот сложный процесс возможным, поскольку облегчает отбор объектов по первому признаку и увеличивает разницу средних (и уменьшает внутригрупповую дисперсию) по второму признаку.

#### 4. Роль формы распределения в зрительном поиске по сочетанию признаков

Статья для защиты: Utochkin, Khvostov, & Wolfe (2020).

Данная работа является логическим продолжением предыдущего исследования. В данном исследовании мы хотели проверить, действует ли сегментабельность распределений двух признаков аналогичным образом в отношении другой важной повседневной перцептивной задачи — зрительного поиска. Как и сегментации текстур, зрительный поиск требует одновременной обработки множественных объектов для определения присутствия цели (объекта с предопределенными характеристиками или отличающегося ото всех объекта) среди отвлекающих стимулов (дистракторов).

В эксперименте 1 (N=12) испытуемым предъявлялось множество линий разной ориентации (от очень пологих до очень крутых) и цветов (от красного до синего). Их задача заключалась в том, чтобы как можно быстрее найти предопределённую заранее цель, заданную сочетанием двух признаков (например, красную пологую линию) среди дистракторов (например, красных крутых и синих пологих линий). Как и в предыдущем исследовании, наша основная экспериментальная манипуляция была связана с формой распределения. Оно могло быть либо несегментабельным (ориентация в диапазоне от  $10^\circ$  до  $80^\circ$  с шагом  $10^\circ$  или цветовые оттенки из цветового круга CIE Lab от  $270^\circ$  (синий) до  $360^\circ$  (красный) с шагом  $12-13^\circ$ ), либо сегментабельным (ориентации: только  $10^\circ$  и  $80^\circ$ , цвета: только  $270^\circ$  и  $360^\circ$ ). Мы манипулировали сегментабельностью каждого признака, что формировало четыре условия сегментабельности: «оба» (оба признака сегментабельны), «ориентация» (ориентация сегментабельна, цвет несегментабелен), «цвет» (ориентация несегментабелен, цвет сегментабелен) и «ни один» (обе признака несегментабельны). Как и в любом исследовании зрительного поиска, мы также варьировали количество объектов (9 или 17) и присутствие цели на экране (присутствует или отсутствует). Мы строили графики функции количества объектов  $\times$  RT. Наклон графика говорит нам, насколько

29

эффективным было управление вниманием (сколько миллисекунд испытуемый тратит на обработку каждого дополнительного объекта на экране). Наклон может варьироваться от нулевого значения (что означает, что поиск был близок к параллельному: внимание мгновенно направлялось к целевому местоположению вне зависимости от количества дистракторов) до очень высоких значений (что означает, что поиск был полностью последовательным: испытуемые должны осматривать элемент за элементом до тех пор, пока не обнаружится цель). Результаты показали полное отсутствие эффекта сегментабельности. Во всех условиях сегментабельности испытуемые показали достаточно эффективный поиск: наклоны функций зависимости времени реакции от количество объектов были равны 9-14 мс/элемент, что довольно близко к параллельному поиску.

Эксперимент 2 ( $N = 15$ ) был направлен на то, чтобы реплицировать этот результат, используя другую пару признаков, взятую из предыдущего исследования: длину и ориентацию. Мы использовали тот же план эксперимента, что и в эксперименте 1, и получили аналогичные результаты: отсутствие эффекта сегментабельности признаков и очень эффективный поиск во всех условиях (наклоны: 0–5 мс/элемент).

Учитывая наш предыдущий вывод о значительном влиянии сегментабельности на распознавание текстуры (исследование 2), данный результат выглядит удивительным. Но отсутствие эффекта сегментабельности при поиске по сочетанию признаков можно объяснить с помощью теории управляемого поиска (Wolfe, 1994, 2021). Эта модель предполагает, что зрительная система имеет доступ к различным ретинотопическим картам признаков (цвета, ориентации, размера и т.д.). Каждое место на такой карте признаков имеет активацию, пропорциональную сходству признака объекта в этом месте с целевым признаком. Суммируя активации из нескольких карт разных признаков, зрительная система строит карту приоритета внимания, которая используется для управления вниманием при зрительном поиске —

избирательное внимание сначала направляется в место с наивысшей активацией. Таким образом, модель управляемого поиска может объяснить эффективность наших поисков из первых двух экспериментов: внимание направляется одновременно на красный цвет и на крутую ориентацию при поиске красной линии с крутой ориентацией. Поскольку все дистракторы на дисплее либо крутые, либо красные, только цель одновременно и крутая, и красная. Эта двойная доза активации одинаково привлекает внимание к целевому местоположению как в сегментабельных, так и в несегментабельных условиях. В последнем случае цель получает двойную дозу активации, а дистракторы — меньшую, поскольку по мере усиления одного признака (например, «крутизны») другой («краснота») ослабевает (из-за противоположного направления корреляции между цветом и ориентацией для цели и дистракторов). Эта модель может объяснить, почему стимулы, которые создают сильный эффект сегментабельности в случае сегментации текстуры, не производят того же самого эффекта для зрительного поиска по сочетанию признаков. Кроме того, эта модель предсказывает, что сегментабельность будет играть важную роль в поиске по сочетанию признаков, если испытуемым не известны все целевые признаки.

В эксперименте 3 ( $N = 13$ ) испытуемые выполняли так называемую задачу поиска внутри подмножества в дисплеях из эксперимента 1. Единственное отличие состояло в том, что цель не была задана заранее, поэтому испытуемые искали уникальную ориентацию внутри подмножества элементов известного цвета (например, уникальную ориентацию в красном подмножестве). Это означает, что цель (как и распределение дистракторов) менялась от пробы к пробе: в половине проб целью была красная крутая линия (дистракторами были красные пологие и синие крутые линии), а в другой половине целью могла быть красная пологая линия (дистракторами были красные крутые и синие пологие линии). Как и в предыдущих экспериментах, у нас было четыре условия сегментабельности («оба», «ориентация», «цвет», «нет») и еще одно условие для сравнения —

стандартный поиск по сочетанию признаков (условие «оба» из эксперимента 1). Результаты снова показали, что поиск во всех условиях был достаточно эффективными (наклоны равнялись 5-14 мс/элемент). Но, в отличие от предыдущих экспериментов, сегментабельность оказала значительное влияние на среднее время реакции: испытуемые были намного медленнее (~200-300 мс) в тех условиях, когда одно или оба распределения признаков были несегментабельными (условия «цвет», «ориентация» и «ни одно»). Важно отметить, что этот эффект был похожим на эффект в исследовании 2 (эксперимент 2), где сегментация была успешной только в условии с сегментабельными распределениями обоих признаков. Наше объяснение этого эффекта состоит в том, что испытуемые тратят дополнительное время на то, чтобы выяснить, какой является ориентация цели в конкретной пробе: это гораздо сложнее сделать, если хотя бы один признак несегментабельный. Если распределение цвета несегментабельно, то сложнее выделить цветовое подмножество для поиска внутри него уникальной ориентации. Если распределение ориентации несегментабельно, то уникальный объект больше похож на дистракторы из целевого подмножества по цвету. Но как только испытуемый выясняет целевую ориентацию, задача становится типичным поиском по сочетанию признаков, в котором испытуемые могут одинаково эффективно распределять внимание в сегментабельных и несегментабельных условиях. Следовательно, мы не обнаружили изменения наклона функции количества объектов  $\times$  времени реакции, но наблюдали значительное влияние сегментабельности на средние значения времени реакции. Таким образом, мы смогли обнаружить эффект сегментабельности в задаче зрительного поиска, в которой роль предварительной категоризации высока.

В эксперименте 4 (N=14) проверялось другое предсказание модели управляемого поиска в отношении сегментабельности. Отсутствие эффекта сегментабельности в экспериментах 1-2 объяснялось тем, что карта приоритета



внимания суммирует паттерны активации двух карт признаков и дает аналогичный результат для сегментабельных и несегментабельных случаев. Но этот похожий паттерн на карте приоритета внимания возникает из совсем разных активаций на картах признаков. В условиях сегментабельности все места на картах признаков либо сильно активированы (объекты имеют точно такой же признак, что и цель), либо слабо активированы (объекты имеют очень отличное от целевого значения признак). Напротив, в несегментабельных условиях обе карты признаков показывают градиент активаций от высокого к низкому. Таким образом, модель управляемого поиска может предсказать довольно контринтуитивный результат: поиск по одному признаку в несегментабельном условии (когда цель определяется либо уникальным цветом, либо только уникальной ориентацией) может быть медленнее и менее эффективным, чем поиск по сочетанию признаков для тех же стимулов. Это предсказание служит хорошей проверкой наших объяснений, потому что оно вступает в противоречие с классическим результатом, полученным Трейсман и Геладом (Treisman & Gelade, 1980), о том, что поиск по одному признаку выполняется быстрее, чем соответствующий поиск по сочетанию признаков. Мы использовали чуть отличающиеся стимулы, однако процедура и план эксперимента были полностью аналогичны эксперименту 1. Из четырех условий сегментабельности из эксперимента 1, мы взяли только условия «оба» и «ни один» и сравнили их с поиском по каждому признаку в отдельности (по ориентации и цвету), в несегментабельном условии. Во всех условиях испытуемые искали одну и ту же цель (например, белую вертикальную линию). При поиске по ориентации испытуемые должны были искать среди разноориентированных белых линий (несегментабельное распределение), при поиске по цвету испытуемым приходилось искать среди вертикальных линий разного цвета (от белого до красного, несегментабельное распределение), при поиске по сочетанию признаков (в сегментабельном условии) им приходилось искать среди красных

вертикальных и белых горизонтальных линий; при поиске по сочетанию признаков (в несегментабельном условии) они должны были искать среди линий с разными ориентациями и цветами. Результаты показали, что поиск по сочетанию признаков (как в сегментабельном, так и в несегментабельном случае) был намного быстрее и эффективнее, чем поиск по одному признаку (особенно плохо испытуемые справлялись с поиском по ориентации), что является сильным эмпирическим свидетельством в пользу нашего объяснения всего набора экспериментов.

Эксперимент 5 ( $N = 26$ ) был посвящен репликации основных результатов этой серии экспериментов с большей выборкой испытуемых (чтобы показать, что отсутствие эффекта сегментабельности не было связано с недостатком мощности) и расширению применимости сделанных выводов на более плотные, «текстуроподобные» дисплеи с большим количеством объектов. Экспериментальный план был похож на эксперимент 1, за исключением добавления двух дополнительных условий количества объектов на экране (всего 4 размера набора: 9, 17, 33, 65). Результаты этого эксперимента полностью повторяют наши результаты предыдущих экспериментов. Когда испытуемые знают оба признака цели, они демонстрировали очень эффективный поиск (7-10 мс/элемент) во всех условиях сегментабельности без какой-либо разницы между условиями.

В данной серии экспериментов мы исследовали то, как зрительная система использует информацию о форме распределений признаков для выполнения зрительного поиска цели, заданной сочетанием этих признаков. Когда оба целевых признака известны испытуемым, зрительная система может преодолеть отсутствие четких категориальных групп дистракторов и эффективно находить цель, используя параллельно несколько карт активации признаков. Но если испытуемым неизвестен хотя бы один из целевых признаков, отсутствие сегментабельности значительно снижает скорость поиска. В других ситуациях,

когда зрительная система не может использовать карты активации нескольких признаков (например, при поиске по одному признаку), мы продемонстрировали, что сегментальность распределения признаков также является обязательным условием осуществления быстрого эффективного поиска.

## 5. Заключение

Главный вопрос данной диссертационной работы заключался в том, как зрительная система может использовать богатую репрезентацию распределений признаков ансамблей. В результате нашего исследования мы пришли к следующим выводам.

Во-первых, зрительная система может эффективно извлекать различные сводные статистики ансамблей из этой богатой репрезентации. Мы получили убедительные доказательства в пользу того, что она можно делать параллельно, т.е. без интерференции между вычислениями нескольких разных статистик. В соответствии с этим различные ансамблевые статистики вычисляются с помощью независимых когнитивных (и, вероятно, нейронных) механизмов, что означает, что вычисления каждой статистики выполняется независимо от других.

Во-вторых, форма распределения признаков может быть использована как основа для быстрой зрительной категоризации и сегментации даже в таком сложном случае, когда подмножества заданы разными сочетаниями нескольких признаков. Испытуемые, вероятно, используют стратегию половинного деления: они выделяют группу объектов из одной половины распределения первого признака и сравнивают эти объекты по второму признаку. Сегментабельность обоих распределений признаков является необходимым условием успешного выполнения такой задачи, поскольку облегчает выделение объектов по первому признаку и увеличивает средний контраст по второму признаку. Этот эффект сегментабельности отражает работу глобального распределенного внимания, поскольку возникает довольно рано, а успешность выполнения задачи выше при предъявлении полной текстуры по сравнению с предъявлением только локальных элементов вблизи границы.

В-третьих, задача зрительного поиска по сочетанию признаков состоит в более сложных отношениях с сегментабельностью. Когда испытуемые знают оба целевых признака, зрительная система может преодолеть отсутствие

сегментабельности распределений дистракторов и эффективно найти цель, используя параллельно несколько карт активации признаков. Но если хотя бы один признак неизвестен, несегментабельность распределения резко снижает скорость поиска. В других ситуациях, когда зрительная система не может использовать карты активации нескольких признаков (например, при поиске по одному признаку), сегментабельность распределения признаков также является необходимым условием для быстрого эффективного поиска.

Следует отметить, что выводы об использовании богатой ансамблевой репрезентации распределений признаков могут быть обобщены на множество различных зрительных признаков. В своих экспериментах мы не полагались слепо на предположение, что ансамблевое восприятие работает одинаково для всех признаков, а проверяли это с использованием разных признаков: размера, цвета и ориентации. Более того, в каждом исследовании был по крайней мере один эксперимент с парой признаков ориентация-размер (или только размер в исследовании 1), поэтому разные результаты между исследованиями нельзя объяснить выбором разных признаков. Таким образом, выводы, полученные в наших экспериментах, представляются надёжными, воспроизводимыми и генерализируемыми для целого диапазона зрительных признаков.

## Список литературы

- Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, 15(3), 122–131. <https://doi.org/10.1016/j.tics.2011.01.003>
- Alvarez, G. A., & Oliva, A. (2008). The Representation of Simple Ensemble Visual Features Outside the Focus of Attention. *Psychological Science*, 19(4), 392–398. <https://doi.org/10.1111/j.1467-9280.2008.02098.x>
- Alvarez, G. A., & Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proceedings of the National Academy of Sciences*, 106(18), 7345–7350. <https://doi.org/10.1073/pnas.0808981106>
- Ariely, D. (2001). Seeing Sets: Representation by Statistical Properties. *Psychological Science*, 12(2), 157–162.
- Bauer, B. (2009). Does Stevens's Power Law for Brightness Extend to Perceptual Brightness Averaging? *The Psychological Record*, 59(2), 171–185. <https://doi.org/10.1007/BF03395657>
- Burr, D., & Ross, J. (2008). A Visual Sense of Number. *Current Biology*, 18(6), 425–428. <https://doi.org/10.1016/j.cub.2008.02.052>
- Cant, J. S., & Xu, Y. (2012). Object Ensemble Processing in Human Anterior-Medial Ventral Visual Cortex. *Journal of Neuroscience*, 32(22), 7685–7700. <https://doi.org/10.1523/JNEUROSCI.3325-11.2012>
- Chetverikov, A., Campana, G., & Kristjánsson, Á. (2016). Building ensemble representations: How the shape of preceding distractor distributions affects visual search. *Cognition*, 153, 196–210. <https://doi.org/10.1016/j.cognition.2016.04.018>
- Chetverikov, A., Campana, G., & Kristjánsson, Á. (2017a). Chapter 5 - Learning features in a complex and changing environment: A distribution-based framework for visual attention and vision in general. In C. J. Howard (Ed.), *Progress in Brain Research* (Vol. 236, pp. 97–120). Elsevier. <https://doi.org/10.1016/bs.pbr.2017.07.001>

- Chetverikov, A., Campana, G., & Kristjánsson, Á. (2017b). Representing Color Ensembles. *Psychological Science*, 28(10), 1510–1517. <https://doi.org/10.1177/0956797617713787>
- Chetverikov, A., Campana, G., & Kristjánsson, Á. (2017c). Set size manipulations reveal the boundary conditions of perceptual ensemble learning. *Vision Research*, 140, 144–156. <https://doi.org/10.1016/j.visres.2017.08.003>
- Chong, S. C., & Evans, K. K. (2011). Distributed versus focused attention (count vs estimate): Distributed versus focused attention. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(6), 634–638. <https://doi.org/10.1002/wcs.136>
- Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, 12.
- Chong, S. C., & Treisman, A. (2005a). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics*, 67(1), 1–13. <https://doi.org/10.3758/BF03195009>
- Chong, S. C., & Treisman, A. (2005b). Statistical processing: Computing the average size in perceptual groups. *Vision Research*, 45(7), 891–900. <https://doi.org/10.1016/j.visres.2004.10.004>
- Corbett, J. E., & Oriet, C. (2011). The whole is indeed more than the sum of its parts: Perceptual averaging in the absence of individual item representation. *Acta Psychologica*, 138(2), 289–301. <https://doi.org/10.1016/j.actpsy.2011.08.002>
- Corbett, J. E., Wurnitsch, N., Schwartz, A., & Whitney, D. (2012). An aftereffect of adaptation to mean size. *Visual Cognition*, 20(2), 211–231. <https://doi.org/10.1080/13506285.2012.657261>
- Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, 1(1), 42–45. <https://doi.org/10.20982/tqmp.01.1.p042>

- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87–114. <https://doi.org/10.1017/S0140525X01003922>
- Dakin, S. C., & Watt, R. J. (1997). The computation of orientation statistics from visual texture. *Vision Research*, 37(22), 3181–3192. [https://doi.org/10.1016/S0042-6989\(97\)00133-8](https://doi.org/10.1016/S0042-6989(97)00133-8)
- Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). THREE PARIETAL CIRCUITS FOR NUMBER PROCESSING. *Cognitive Neuropsychology*, 20(3–6), 487–506. <https://doi.org/10.1080/02643290244000239>
- Florey, J., Clifford, C. W. G., Dakin, S., & Mareschal, I. (2016). Spatial limitations in averaging social cues. *Scientific Reports*, 6(1). <https://doi.org/10.1038/srep32210>
- Gardelle, V. de, & Summerfield, C. (2011). Robust averaging during perceptual judgment. *Proceedings of the National Academy of Sciences of the United States of America*, 108(32), 13341–13346.
- Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*, 17(17), R751–R753. <https://doi.org/10.1016/j.cub.2007.06.039>
- Halberda, J., Sires, S. F., & Feigenson, L. (2006). Multiple Spatially Overlapping Sets Can Be Enumerated in Parallel. *Psychological Science*, 17(7), 572–576. <https://doi.org/10.1111/j.1467-9280.2006.01746.x>
- Huang, L., Mo, L., & Li, Y. (2012). Measuring the interrelations among multiple paradigms of visual attention: An individual differences approach. *Journal of Experimental Psychology: Human Perception and Performance*, 38(2), 414–428. <https://doi.org/10.1037/a0026314>
- Im, H. Y., Tiurina, N. A., & Utochkin, I. S. (2021). An explicit investigation of the roles that feature distributions play in rapid visual categorization. *Attention, Perception, & Psychophysics*, 83(3), 1050–1069. <https://doi.org/10.3758/s13414-020-02046-7>



- Khvostov V., Lukashevich A., Utochkin I. S. (2021). Spatially intermixed objects of different categories are parsed automatically. *Scientific Reports*, 11, 377, 1-8. <https://doi.org/10.1038/s41598-020-79828-4>.
- Khvostov, V. A., & Utochkin, I. S. (2019). Independent and parallel visual processing of ensemble statistics: Evidence from dual tasks. *Journal of Vision*, 19(9), 3. <https://doi.org/10.1167/19.9.3>
- Kim, M., & Chong, S. C. (2020). The visual system does not compute a single mean but summarizes a distribution. *Journal of Experimental Psychology: Human Perception and Performance*, 46(9), 1013–1028. <https://doi.org/10.1037/xhp0000804>
- Lee, H., Baek, J., & Chong, S. C. (2016). Perceived magnitude of visual displays: Area, numerosity, and mean size. *Journal of Vision*, 16(3), 12. <https://doi.org/10.1167/16.3.12>
- Leib, A. Y., Kosovicheva, A., & Whitney, D. (2016). Fast ensemble representations for abstract visual impressions. *Nature Communications*, 7(1). <https://doi.org/10.1038/ncomms13186>
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390(6657), 279–281. <https://doi.org/10.1038/36846>
- Maule, J., & Franklin, A. (2015). Effects of ensemble complexity and perceptual similarity on rapid averaging of hue. *Journal of Vision*, 15(4), 6. <https://doi.org/10.1167/15.4.6>
- Morgan, M., Chubb, C., & Solomon, J. A. (2008). A “dipper” function for texture discrimination based on orientation variance. *Journal of Vision*, 8(11), 9–9. <https://doi.org/10.1167/8.11.9>
- Myczek, K., & Simons, D. J. (2008). Better than average: Alternatives to statistical summary representations for rapid judgments of average size. *Perception & Psychophysics*, 70(5), 772–788. <https://doi.org/10.3758/PP.70.5.772>
- Neisser, U. (1967). *Cognitive psychology*. Appleton-Century-Crofts.

- Nieder, A., & Dehaene, S. (2009). Representation of Number in the Brain. *Annual Review of Neuroscience*, 32(1), 185–208. <https://doi.org/10.1146/annurev.neuro.051508.135550>
- Norman, L. J., Heywood, C. A., & Kentridge, R. W. (2015). Direct encoding of orientation variance in the visual system. *Journal of Vision*, 15(4), 3. <https://doi.org/10.1167/15.4.3>
- Oriet, C., & Hozempa, K. (2016). Incidental statistical summary representation over time. *Journal of Vision*, 16(3), 3–3. <https://doi.org/10.1167/16.3.3>
- Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, 4(7), 739–744. <https://doi.org/10.1038/89532>
- Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision*, 3(3), 179–197. <https://doi.org/10.1163/156856888x00122>
- Solomon, J. A., Morgan, M., & Chubb, C. (2011). Efficiencies for the statistics of size discrimination. *Journal of Vision*, 11(12), 13–13. <https://doi.org/10.1167/11.12.13>
- Sweeny, T. D., & Whitney, D. (2014). Perceiving Crowd Attention: Ensemble Perception of a Crowd's Gaze. *Psychological Science*, 25(10), 1903–1913. <https://doi.org/10.1177/0956797614544510>
- Treisman, A. (2006). How the deployment of attention determines what we see. *Visual Cognition*, 14(4–8), 411–443. <https://doi.org/10.1080/13506280500195250>
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97–136. [https://doi.org/10.1016/0010-0285\(80\)90005-5](https://doi.org/10.1016/0010-0285(80)90005-5)
- Underwood, B. J. (1975). Individual differences as a crucible in theory construction. *American Psychologist*, 30(2), 128–134. <https://doi.org/10.1037/h0076759>
- Utochkin, I. S. (2015). Ensemble summary statistics as a basis for rapid visual categorization. *Journal of Vision*, 15(4), 8. <https://doi.org/10.1167/15.4.8>

- Utochkin, I. S., Khvostov, V. A., & Stakina, Y. M. (2018). Continuous to discrete: Ensemble-based segmentation in the perception of multiple feature conjunctions. *Cognition*, *179*, 178–191. <https://doi.org/10.1016/j.cognition.2018.06.016>
- Utochkin, I. S., Khvostov, V. A., & Wolfe, J. M. (2020). Categorical grouping is not required for guided conjunction search. *Journal of Vision*, *20*(8), 30. <https://doi.org/10.1167/jov.20.8.30>
- Utochkin, I. S., & Tiurina, N. A. (2014). Parallel averaging of size is possible but range-limited: A reply to Marchant, Simons, and De Fockert. *Acta Psychologica*, *146*, 7–18. <https://doi.org/10.1016/j.actpsy.2013.11.012>
- Utochkin, I. S., & Vostrikov, K. O. (2017). The numerosity and mean size of multiple objects are perceived independently and in parallel. *PLOS ONE*, *12*(9), e0185452. <https://doi.org/10.1371/journal.pone.0185452>
- Utochkin, I. S., & Yurevich, M. A. (2016). Similarity and heterogeneity effects in visual search are mediated by “segmentability”. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(7), 995–1007. <https://doi.org/10.1037/xhp0000203>
- Watamaniuk, S. N. J., & Duchon, A. (1992). The human visual system averages speed information. *Vision Research*, *32*(5), 931–941. [https://doi.org/10.1016/0042-6989\(92\)90036-I](https://doi.org/10.1016/0042-6989(92)90036-I)
- Whiting, B. F., & Oriet, C. (2011). Rapid averaging? Not so fast! *Psychonomic Bulletin & Review*, *18*(3), 484–489. <https://doi.org/10.3758/s13423-011-0071-3>
- Whitney, D., & Yamanashi Leib, A. (2018). Ensemble Perception. *Annual Review of Psychology*, *69*(1), 105–129. <https://doi.org/10.1146/annurev-psych-010416-044232>
- Wilmer, J. (2008). How to use individual differences to isolate functional organization, biology, and utility of visual functions; with illustrative proposals for stereopsis. *Spatial Vision*, *21*(6), 561–579. <https://doi.org/10.1163/156856808786451408>

- Wolfe, J. M. (1994). Guided Search 2.0 A revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238. <https://doi.org/10.3758/BF03200774>
- Wolfe, J. M. (2021). Guided Search 6.0: An updated model of visual search. *Psychonomic Bulletin & Review*, *28*(4), 1060–1092. <https://doi.org/10.3758/s13423-020-01859-9>
- Wolfe, J. M., Võ, M. L.-H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences*, *15*(2), 77–84. <https://doi.org/10.1016/j.tics.2010.12.001>
- Yang, Y., Tokita, M., & Ishiguchi, A. (2018). Is There a Common Summary Statistical Process for Representing the Mean and Variance? A Study Using Illustrations of Familiar Items. *I-Perception*, *9*(1), 204166951774729. <https://doi.org/10.1177/2041669517747297>
- Ying, H., & Xu, H. (2017). Adaptation reveals that facial expression averaging occurs during rapid serial presentation. *Journal of Vision*, *17*(1), 15–15. <https://doi.org/10.1167/17.1.15>