

**Влияние пространственно-частотного содержания изображения лица на его
распознавание**

Алексеева Дарья Сергеевна

Студент (магистр)

Южный федеральный университет, Факультет психологии, Кафедра психофизиологии и
клинической психологии, Ростов-на-Дону, Россия

E-mail: susan05@list.ru

Лица относятся к особой категории объектов. К классу человеческих лиц мозг относит сложные стимулы, сочетающие фиксированный набор простых элементов (черт лица) и расстояний между ними, размеры которых могут варьироваться в определенных пределах. Человеческое лицо, как и любой объект реального мира, является поличастотным, то есть содержит и низко-, и средне-, и высокочастотную информацию. Зрительная система человека характеризуется наличием параллельных каналов, по каждому из которых передается информация в определенном пространственно-частотном диапазоне. Ранее было обнаружено 6 таких каналов с пиковыми значениями пропускания 0,5, 1, 2, 4, 8 и 16 циклов/град [3]. Позднее было доказано, что разные пространственные частоты осуществляют разный вклад в распознавание лиц [1, 2]. Наше исследование было связано с определением возможной очередности использования пространственных частот в распознавании лиц. Актуальность исследования состоит в возможности моделирования процесса распознавания лиц человеком для создания более надежной связи «человек-машина», объяснении ранних фаз процесса распознавания лиц, оптимизации алгоритмов сжатия изображений. Полученные данные могут быть использованы в криминалистике, рекламной индустрии, анализе произведений искусства, информационно-кибернетических отраслях. В эксперименте приняли участие 7 студентов и преподавателей факультета психологии Южного Федерального университета с нормальным либо корригируемым до нормального зрением. Перед испытуемыми ставилась задача выбора между стимулами-изображениями. Используемые изображения представляли собой выровненные по яркости и контрасту фотографии лиц, помещенные на равномерно серый фон. Они были подвергнуты полосовой пространственно-частотной Фурье-фильтрации с пиковыми частотами, выбранными исходя из настроек 6 психофизических частотных каналов, при ширине полосы пропускания в 2 октавы. Общее количество стимульных изображений составило 1288, а именно (1 неотфильтрованное + 6 отфильтрованных) * 184. Испытуемому предъявлялся целевой стимул — неотфильтрованное (эксперимент 1) либо отфильтрованное (эксперимент 2) лицо, затем демонстрировалась маска (для того, чтобы обеспечить фиксированную продолжительность обработки стимула) — оригинальное изображение лица, а после этого — четыре тестовых стимула, отфильтрованные (эксперимент 1) или неотфильтрованные (эксперимент 2). Один из тестовых стимулов был получен из той же фотографии, что и целевой, и от испытуемого требовалось его найти и обозначить свой выбор нажатием клавиши с соответствующей цифрой. Все 4 одновременно предъявляемых тестовых стимула изображали лица одного биологического пола. При этом и целевой, и тестовый стимулы были оба отфильтрованными или оба реальными лицами только в контрольных пробах (1/7 от общего числа). В остальных случаях, если в качестве целевого стимула выступало отфильтрованное по какой-либо из частот лицо, выбор осуществлялся из реальных лиц и наоборот. Время экспозиции целевого стимула случайно варьировалось в диапазоне от 2 до 10 кадровых развёрток с шагом в 2 кадровые развёртки (одна кадровая развёртка - 13,33 мс). Время предъявления маски было постоянным и составляло 15 кадровых развёрток. Каждый из испытуемых выполнил не менее 60 проб для каждого времени предъявления и каждой частотной полосы. Первые два эксперимента серии различались порядком

предъявления стимулов: в эксперименте 1 целевые стимулы были неотфильтрованными, а тестовые - реальными изображениями, в эксперименте 2 - наоборот. Задача эксперимента 2 относительно лиц, отфильтрованных по частотам 1-8 циклов/град, оказалась для испытуемых более сложной по сравнению с экспериментом 1. Если в эксперименте 1 распознавание для каждой из этих частот было столь же эффективным, сколь и для оригинальных лиц, то по результатам эксперимента 2 для различных частот наблюдалась различная точность. При фильтрации по 0,5 и 16 циклов/град задача в обоих экспериментах решалась менее эффективно по сравнению с другими частотами. Можно заметить, что в обоих экспериментах испытуемому предъявлялись одни и те же стимулы - отфильтрованные и неотфильтрованные лица. Изменялся только порядок предъявления. Мы предположили, что различия результатов могут быть связаны с увеличением алфавита целевых стимулов - количество отфильтрованных изображений в 6 раз превышало количество неотфильтрованных. Для проверки этой гипотезы мы провели эксперименты 3 и 4, в которых число вариантов целевых стимулов было сокращено до двух. В обоих экспериментах использовались неотфильтрованные лица и лица, отфильтрованные по частоте 1 цикл/град. Однако в эксперименте 3 оба целевых стимула предъявлялись в соотношении 1:1, а в эксперименте 4 отфильтрованных было в 6 раз больше (как в экспериментах 1 и 2). Уменьшение числа вариантов целевого стимула в эксперименте 3 не изменило соотношения в эффективности распознавания неотфильтрованных и отфильтрованных лиц (рис. 1). Результат эксперимента 3 фактически повторял результат эксперимента 2. Это указывает на то, что выдвинутая нами гипотеза не подтверждается и изменение результата в эксперименте 2 по сравнению с экспериментом 1 не может быть объяснено увеличением алфавита целевых стимулов. В эксперименте 4, который фактически воспроизводил oddball процедуру, результат оказался иным (рис.2). Эффективность распознавания неотфильтрованных лиц снизилась до уровня распознавания отфильтрованных изображений. Это может говорить о том, что в условиях данного эксперимента зрительная система испытуемых настраивается на восприятие входной информации в узком диапазоне частот (1 цикл/град). Когда же предъявляется неотфильтрованное (полочастотное) изображение, все равно используется только этот частотный канал. В результате количество информации, полученной из неотфильтрованного и отфильтрованного изображения выравнивается, а это уравнивает и эффективность их распознавания. В качестве возможного объяснения полученных нами результатов мы выдвигаем две гипотезы. Первая из них связана с тем, что отфильтрованное лицо не воспринимается мозгом как естественное изображение и потому запоминается хуже. Вторая исходит из особенностей работы кратковременной памяти - скорость стирания информации из кратковременной памяти постоянна, и если работает только один частотный канал, то входящей информации недостаточно, чтобы обеспечить распознавание с учетом этих потерь.

Источники и литература

- 1) Collin C. A., Therrien M., Martin C., Rainville S. (2006). Spatial frequency thresholds for face recognition when comparison faces are filtered and unfiltered // Perception & psychophysics. – 2006. – Т. 68. – №. 6. – С. 879-889.
- 2) Kwon M.Y., Legge G.E. Spatial-frequency cutoff requirements for pattern recognition in central and peripheral vision // Vision research. – 2011. – Т. 51. – №. 18. – С. 1995-2007.
- 3) Wilson H. R., Gelb D. J. Modified line-element theory for spatial-frequency and width discrimination // JOSA A. – 1984. – Т. 1. – №. 1. – С. 124-131.

Иллюстрации

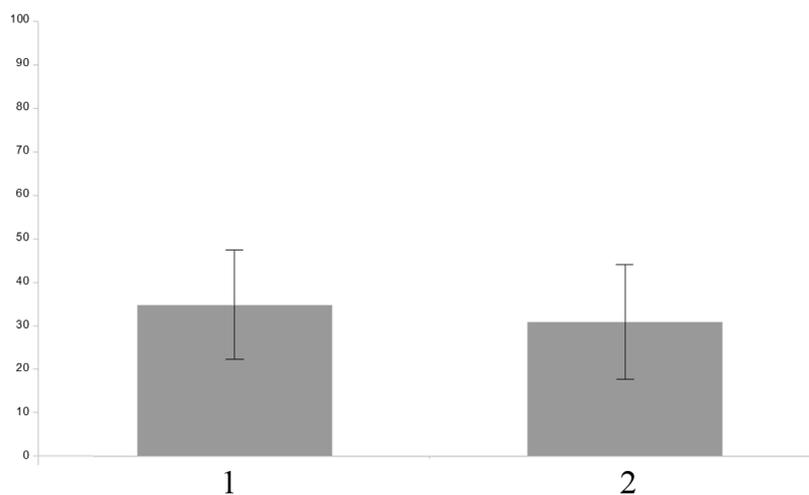


Рис. 1. Гистограмма сравнительной точности распознавания неотфильтрованных и отфильтрованных лиц в соотношении 1:6. По оси ординат – процент правильных ответов. 1 – доля правильных ответов испытуемого для неотфильтрованных лиц, 2 – для отфильтрованных. Вертикальными чертами обозначены доверительные интервалы.

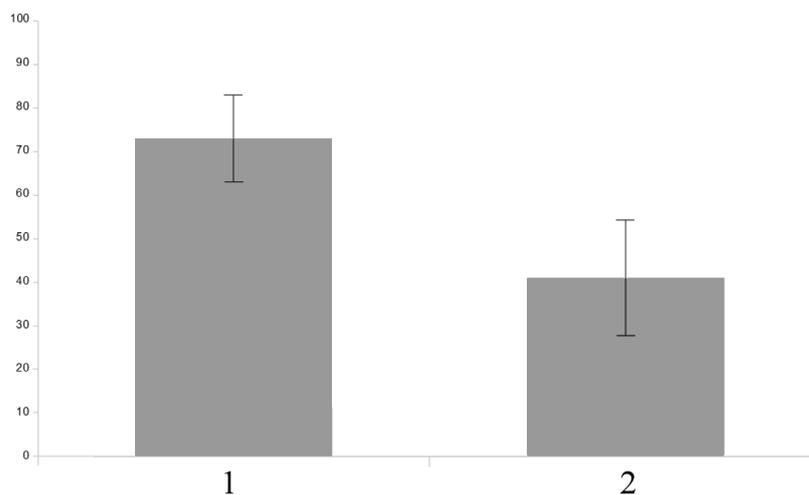


Рис. 2. Гистограмма сравнительной точности распознавания неотфильтрованных и отфильтрованных лиц в соотношении 1:1. По оси ординат – процент правильных ответов. 1 – доля правильных ответов испытуемого для неотфильтрованных лиц, 2 – для отфильтрованных. Вертикальными чертами обозначены доверительные интервалы.