

## Метод извлечения из цифровой фотографии лица информации, достаточной для его идентификации

*Д.В. Явна, В.В. Бабенко*

*Южный федеральный университет, Ростов-на-Дону*

**Аннотация:** Предлагается подход к извлечению наиболее информативных для человека составляющих изображения, базирующийся на психофизиологических представлениях о зрительных механизмах группирования простых признаков – яркостных градиентов. В основе подхода лежит обработка изображения средствами программно реализованной модели «фильтрация-выпрямление-фильтрация». Выделение наиболее информативных областей фотоснимка основывается на поиске локальных максимумов активации выходов применяемой модели.

**Ключевые слова:** зрительные механизмы второго порядка, фильтрация изображений, области зрительного интереса, идентификация лиц, компьютерное моделирование.

### Введение

В данной работе предлагается подход к извлечению наиболее информативных для человека составляющих изображения, базирующийся на психофизиологически обоснованной модели зрительных механизмов группирования простых зрительных признаков.

Механизмы пространственной интеграции локальных (в пределах размеров рецептивного поля простого стриарного нейрона) зрительных признаков – градиентов яркости, варьирующих по ориентации, пространственной частоте и величине контраста, – получили в литературе название зрительных механизмов второго порядка. Раскрытию принципов их работы служит модель «фильтрация-выпрямление-фильтрация» (см. обзор [1]), представляющая начальные стадии обработки зрительной информации мозгом как последовательные этапы линейной фильтрации, разделённые промежуточной операцией выпрямления. Данная модель создавалась с учётом свойств ганглионарных клеток сетчатки, нейронов наружного коленчатого тела таламуса и простых нейронов проекционной коры, а также предполагаемых нервных клеток второго порядка. Такие клетки, избирательно реагирующие на стимул, в котором при постоянных яркости и

цвете пространственно варьируют контраст, пространственная частота и ориентация яркостных градиентов, были выявлены, например, в зоне МТ у макак-резусов [2], в 17 и 18 полях у кошек [3–5]. Модель также согласуется с множеством данных, полученных в психофизических экспериментах [1].

Зрительные механизмы второго порядка могут рассматриваться как обнаружители наиболее информативных областей в изображении [6]. С физической точки зрения эти области являются наиболее неоднородными: содержащиеся в них *простые признаки* (градиенты яркости) значительно варьируют по пространственной частоте, ориентации или контрасту. Уместно говорить, что амплитуда функции модуляции пространственной частоты, ориентации и/или контраста для этих областей максимальна. Можно полагать, что выделение таких областей является важной операцией, выполняемой мозгом при решении задачи идентификации объектов вообще и человеческого лица в частности, а содержимое этих областей является определяющим для его опознания.

### Моделирование

Используя компьютерную реализацию модели «фильтрация-выпрямление-фильтрация», мы попробовали показать, как может осуществляться поиск «областей интереса» в изображении лица, а также продемонстрировать роль информации, содержащейся в этих областях, в передаче его существенных (важных для идентификации) признаков.

Ранее нами была предложена модель группирования локальных признаков, специфичная к модуляциям контраста [7]. В текущей реализации модели механизм, обеспечивающий специфичность, не использовался из соображений целесообразности охвата возможно большего числа информативных признаков изображения. Так, модуляции контраста важны преимущественно для фигуρο-фоновой сегментации, модуляции ориентации

---

и пространственной частоты – для восприятия структуры поверхности и признаков глубины.

Сначала тестовые изображения, выровненные по средней яркости и контрасту, подвергались полосовой линейной фильтрации. Ядром фильтра была двумерная функция Габора, используемая в моделях стриарного нейрона ([9,10]). График значений такой функции показан на Рис. 1.

Параметры фильтров варьировались таким образом, чтобы достичь относительно полного ориентационного и пространственно-частотного охвата градиентов яркости в изображении: всего использовались 6 пиковых пространственных частот (от 0,5 до 16 циклов на изображение с увеличением на октаву на каждом шаге) и 6 предпочтительных ориентаций (от 0 до 150 град. с шагом в 30 град.), что соответствует представлениям об организации зрительных психофизических каналов у человека [11,12]. Эти представления определили и выбор полосы пропускания фильтров по ориентации и пространственной частоте. Кроме того, проводились аналогичные операции с ядрами со смещённой на 90 град. фазой.

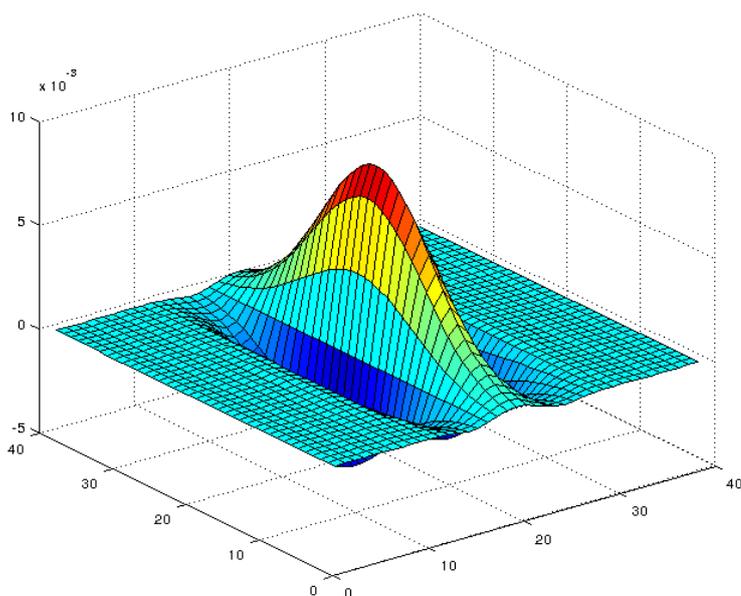


Рис. 1. Пример ядра фильтра Габора (по [8]). По оси абсцисс – пиксели, по оси ординат – значения функции Габора.

Таким образом, в результате моделирования первого этапа линейной фильтрации мы получили 72 отфильтрованных изображения, представлявших собой выходы 36 квадратурных пар габоровских фильтров. Результаты объединения выходов фильтров с разными ориентационными настройками для каждой пиковой пространственной частоты представлены на Рис. 2.

Этап нелинейных преобразований моделировался поэлементной операцией извлечения квадратного корня из сумм квадратов выходов фильтров, образующих квадратурную пару. Эта операция подобна той, что применяется при расчёте энергетического фильтра Габора [8, 13]. Психофизиологический смысл такого полноволнового выпрямления состоит в учёте как возбуждательных, так и тормозных ответов нейронов.



Рис. 2. Объединённые по всем ориентациям выходы фильтров Габора с различными частотными настройками: *а* – для частоты 4 цикла/изображение, *б* – 8, *в* – 16, *г* – 32, *д* – 64, *е* – 128 циклов/изображение.

На следующем этапе полученные 36 наборов значений выходов энергетического фильтра подвергались линейной фильтрации с использованием фильтров Габора, пиковая пространственная частота которых была в восемь раз ниже, чем у соответствующих фильтров первой стадии [14]. Так, если для получения данного набора значений на первом этапе использовался фильтр, настроенный на частоту 8 циклов/изображение,

на втором этапе применялся фильтр, настроенный на 1 цикл/изображение. Ориентационные предпочтения фильтров на обоих этапах совпадали.

После операции фильтрации осуществлялось объединение выходов фильтров второго этапа с разными ориентационными настройками для каждой из пиковых пространственных частот. В результирующую матрицу выходов фильтров записывались максимальные значения из возможных шести, соответствующих выходам ориентационно-избирательных фильтров второго этапа. Таким образом, были сформированы 6 пространственно-частотных «срезов» изображения, содержащих информацию об амплитуде модуляции простых признаков по контрасту, ориентации и пространственной частоте, для каждой пары соотношений пространственных частот несущей и огибающей.

Следующей операцией был поиск локальных максимумов на полученных «срезах». Координаты найденных максимумов рассматривались как центры концентрических областей, диаметры которых соответствовали половине длины периода функции модуляции, на частоту которой был настроен используемый для получения данного частотного «среза» фильтр второго этапа. Эти области рассматривались как «области интереса». Примеры таких областей показаны на Рис. 3.

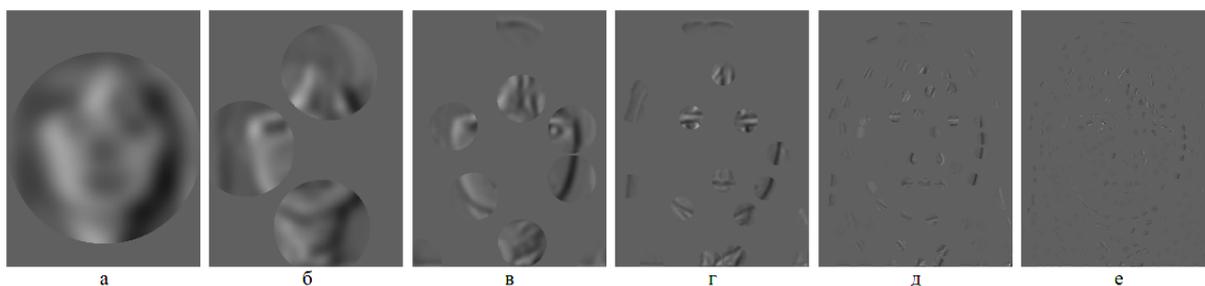


Рис. 3. Примеры «областей интереса», выявляемых в разных пространственно-частотных каналах второго порядка: *а* – для частоты 0,5 цикла/изображение, *б* – 1, *в* – 2, *г* – 4, *д* – 8, *е* – 16 цикл/изображение. В иллюстративных целях области заполнены отфильтрованным на первом

---

этапе изображением без ослабления от центра к периферии. Радиус областей составляет 2 предпочтительные длины волны фильтра первого порядка.

Для каждого соотношения частот несущей и огибающей выделенные «области интереса» заполнялись изображением, полученным на первом этапе линейной фильтрации. Области рассматривались как неоднородные: значения яркости пикселей от центра к периферии области ослаблялись по гауссиану, причём радиус области составлял 2 стандартных отклонения (Бабенко, Кульба 2002). Результаты показаны на Рис. 4.

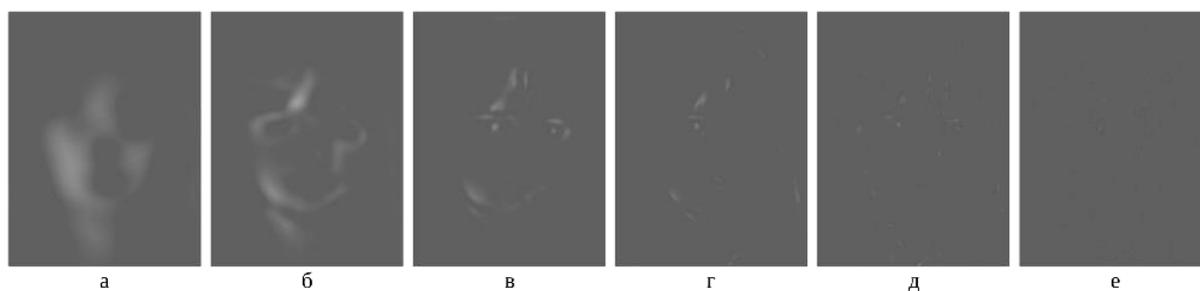


Рис. 4. Заполнение «областей интереса» изображением, полученным на первом этапе линейной фильтрации: *а* – для пиковой частоты фильтра второго этапа, настроенного на 0,5 цикла/изображение, *б* – 1, *в* – 2, *г* – 4, *д* – 8, *е* – 16 цикла/изображение. Пиковые частоты изображений, используемых для заполнения, см. в подписи к Рис. 2.

Частично восстановленные отфильтрованные изображения суммировались между собой. К низкочастотному изображению последовательно добавлялись более высокочастотные. Результаты представлены на Рис. 5.

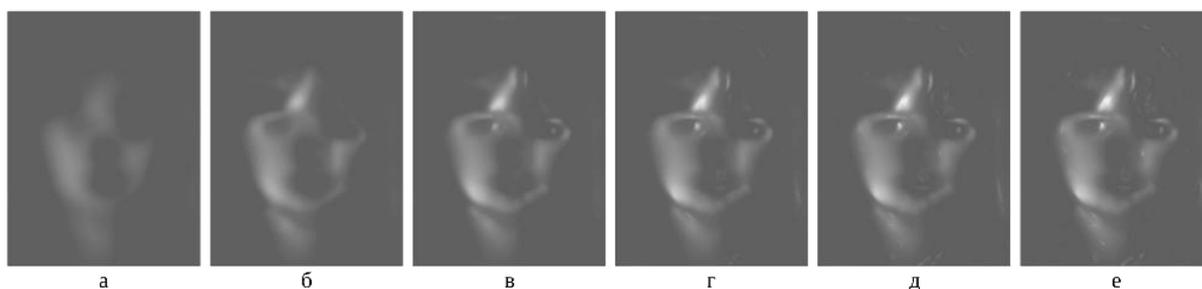


Рис. 5. Результаты инкрементного суммирования изображений, полученных путём заполнения «областей интереса»: *а* – низкочастотное изображение (см. Рис. 4 *а*), *б* – сумма изображения *а* и *б* из Рис. 4, *в* – сумма изображений *а*, *б* и *в* из Рис. 4 и т. д.

Из Рис. 5 видно, что низкочастотное изображение *а* предоставляет информацию об очертаниях головы, *б* – о половой принадлежности изображённого, *в* – о чертах его лица, что упрощает идентификацию показанного на картинке человека. Добавление более высокочастотных составляющих практически не вносит новых черт в формирующийся у наблюдателя перцептивный образ.

### Заключение

Очевидно, что абсолютное значение пространственной частоты, на которой представлена необходимая для идентификации лица информация, не является постоянной величиной: мы одинаково хорошо идентифицируем лица, видимые нами в широком диапазоне расстояний. Однако моделируемый механизм инвариантен к масштабу изображения [15]. На наш взгляд, организация такого механизма структурно близка к адаптивному многоканальному (многоскоростному) фильтру [16] с той особенностью, что настройка фильтрующей системы достигается преимущественно за счёт выбора оптимально работающих пространственно-частотных каналов, а не путём изменения весовых коэффициентов фильтров. Так, если угловые размеры объекта уменьшаются, то задача его обнаружения и идентификации будет решаться каналами, настроенными на более высокую частоту.

Значимым результатом проведённых расчётов следует признать также то, что идентификация лица может осуществляться посредством ограниченного числа психофизических пространственно-частотных каналов, задействуемых в зависимости от частоты, на которой обнаруживаются «зоны интереса». Это может быть использовано для оптимизации алгоритмов пространственно-частотного прореживания, применяемого, в частности, в общеизвестном способе сжатия с потерей качества Joint Photographic Expert Group (jpeg). Но главный с психофизиологической точки зрения результат заключается в том, что наиболее информативными областями изображения являются те, которые содержат наибольшее число изменений простых зрительных признаков – градиентов яркости.

**Работа выполнена при финансовой поддержке Минобрнауки России по теме № 213.01-11/2014-4 в рамках задания №2014/174 на выполнение государственных работ в сфере научной деятельности (базовая часть государственного задания).**

### Литература

1. Graham N.V. Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): Useful additions of the last 25 years. Vision Research. 2011. Vol. 51, №13. pp. 1397–1430.
  2. Albright T.D. Form-cue invariant motion processing in primate visual cortex // Science. 1992. Vol. 255, № 5048. pp. 1141–1143.
  3. Mareschal I., Baker C.L.Jr. Temporal and spatial response to second-order stimuli in cat area 18 // J. Neurophysiol. 1998. Vol. 80, № 6. pp. 2811–2823.
  4. Zhou Y.X., Baker C.L.Jr. A processing stream in mammalian visual cortex neurons for non-Fourier responses // Science. 1993. Vol. 261, № 5117. pp. 98–101.
-



5. Zhou Y.X., Baker C.L.Jr. Spatial properties of envelope-responsive cells in area 17 and 18 neurons of the cat // J. Neurophysiol. 1996. Vol. 75, № 3. pp. 1038–1050.
  6. Бабенко В.В., Кульба С.Н. Модель механизма зрительной сегментации // Сенсорные системы. 2002. Т. 16, № 3. С. 179-189.
  7. Явна Д.В. Компьютерное моделирование зрительных механизмов группирования, избирательных к пространственным модуляциям контраста // Инженерный вестник Дона. 2013. №4. URL: [ivdon.ru/ru/magazine/archive/n4y2013/2009](http://ivdon.ru/ru/magazine/archive/n4y2013/2009).
  8. Grigorescu C., Petkov N., Westenberg M.A. Contour detection based on non-classical receptive field inhibition // IEEE Trans. Image Processing. 2003. Vol. 12, № 7. pp. 729–739.
  9. Daugman J.G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters // J. Opt. Soc. Am. A. 1985. Vol. 2, № 7. pp. 1160–1169.
  10. Jones J.P, Palmer L.A. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex // J. Neurophysiol. 1987. Vol. 58, № 6. pp. 1233-1258.
  11. Wilson H.R., McFarlane D.K., Phillips G.C. Spatial frequency tuning of orientation selective units estimated by oblique masking // Vision Research. 1983. Vol. 23, № 9. pp. 873–882.
  12. Ellemberg D., Allen H.A., Hess R.F. Second-order spatial frequency and orientation channels in human vision // Vision Research. 2006. Vol. 46, № 17. pp. 2798–2803.
  13. Fogel I., Sagi D. Gabor filters as texture discriminator // Biol. Cybern. 1989. Vol. 61, № 2. pp. 103–113.
-

14. Бабенко В.В., Ермаков П.Н., Божинская М.А. Соотношение пространственно-частотных настроек зрительных фильтров первого и второго порядка // Психологический журнал. 2010. Т. 31, № 2. С. 48–57.

15. Бабенко В.В. Новый подход к вопросу о механизмах зрительного восприятия // Проблемы нейрокибернетики. Ростов-на-Дону: Издательство ростовского университета, 1989. С. 10–11.

16. Линович А.Ю. Метод гибкого формирования подсистемы частотно-временной локализации сигналов в задачах адаптивной фильтрации // Инженерный вестник Дона. 2014. №1. URL: <http://www.ivdon.ru/ru/magazine/archive/n1y2014/2289>

### References

1. Graham N.V. Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): Useful additions of the last 25 years. Vision Research. 2011. Vol. 51, №13. pp. 1397–1430.
  2. Albright T.D. Form-cue invariant motion processing in primate visual cortex. Science. 1992. Vol. 255, № 5048. pp. 1141–1143.
  3. Mareschal I., Baker C.L.Jr. Temporal and spatial response to second-order stimuli in cat area 18. J. Neurophysiol. 1998. Vol. 80, № 6. pp. 2811–2823.
  4. Zhou Y.X., Baker C.L.Jr. A processing stream in mammalian visual cortex neurons for non-Fourier responses. Science. 1993. Vol. 261, № 5117. pp. 98–101.
  5. Zhou Y.X., Baker C.L.Jr. Spatial properties of envelope-responsive cells in area 17 and 18 neurons of the cat. J. Neurophysiol. 1996. Vol. 75, № 3. pp. 1038–1050.
  6. Babenko V.V., Kul'ba S.N. Sensornye sistemy (Rus). 2002. Vol. 16, № 3. pp. 179-189.
  7. Yavna D.V. Inženernyj vestnik Dona (Rus). 2013. №4. URL: [ivdon.ru/ru/magazine/archive/n4y2013/2009](http://ivdon.ru/ru/magazine/archive/n4y2013/2009).
-



8. Grigorescu C., Petkov N., Westenberg M.A. Contour detection based on non-classical receptive field inhibition // IEEE Trans. Image Processing. 2003. Vol. 12, № 7. pp. 729–739.
9. Daugman J.G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J. Opt. Soc. Am. A. 1985. Vol. 2, № 7. pp. 1160–1169.
10. Jones J.P., Palmer L.A. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex // J. Neurophysiol. 1987. Vol. 58, № 6. pp. 1233–1258.
11. Wilson H.R., McFarlane D.K., Phillips G.C. Spatial frequency tuning of orientation selective units estimated by oblique masking // Vision Research. 1983. Vol. 23, № 9. pp. 873–882.
12. Ellemberg D., Allen H.A., Hess R.F. Second-order spatial frequency and orientation channels in human vision // Vision Research. 2006. Vol. 46, № 17. pp. 2798–2803.
13. Fogel I., Sagi D. Gabor filters as texture discriminator // Biol. Cybern. 1989. Vol. 61, № 2. pp. 103–113.
14. Babenko V.V., Ermakov P.N., Bozhinskaja M.A. Psikhologicheskii Zhurnal (Rus). 2010. Vol. 31, № 2. pp. 48–57.
15. Babenko V.V. Novyj podhod k voprosu o mehanizmah zritel'nogo vosprijatija [A new approach to the question of the mechanisms of visual perception]. Problemy nejrokibernetiki (Rus). Rostov-na-Donu: Izdatel'stvo rostovskogo universiteta, 1989. pp. 10–11.
16. Linovich A.Ju. Inženernyj vestnik Dona (Rus). 2014. №1. URL: [ivdon.ru/ru/magazine/archive/n1y2014/2289](http://ivdon.ru/ru/magazine/archive/n1y2014/2289).