УДК 621.397.3: 536.3

ПРИМЕНЕНИЕ ОБОБЩЕННЫХ ЭТАЛОННЫХ ФУНКЦИЙ В ЕСТЕСТВЕННЫХ И ИСКУССТВЕННЫХ ЗРИТЕЛЬНЫХ СИСТЕМАХ

© **2007 г.** **В. Р. Луцив**, канд. техн. наук

НПК "Государственный оптический институт им. С.И. Вавилова", Санкт-Петербург

Разработан метод структурной декомпозиции изображений путем их локального сопоставления со специальными обобщенными эталонами. Примененный подход, заимствованный в естественных зрительных системах, позволил представить с единых позиций ряд разделов теории обработки изображений и был успешно использован для решения разнообразных задач технического зрения.

Коды OCIS: 100.2980, 3000.3020.

Поступила в редакцию 29.03.2007.

1. Введение

Появление реально действующих систем технического зрения можно датировать серединой прошлого века, когда хорошо формализованная теория распознавания образов начала практически воплощаться в конкретных кибернетических системах. Сначала источником алгоритмов распознавания образов была статистическая теория связи. Методы согласованной фильтрации и корреляции положили начало созданию формально оптимальных пространственно неинвариантных или частично инвариантных объектно-независимых алгоритмов (см., например, [1, 2]), имевших ограниченные возможности сопоставления изображений. В это же время получили достаточно широкое распространение эффективные объектнозависимые системы узкого применения. Поначалу они также использовали формальную теорию распознавания образов, унаследованную от статистической теории связи (см., например, [3]), однако вскоре появляются эвристические решения, основанные на моделировании элементов человеческого восприятия в том виде, как оно понималось в то время инженерами и программистами (см., например, [4, 5]). К этому периоду можно отнести также первые успехи реализации алгоритмами технического зрения обнаруженных в живых зрительных системах (ЖЗС) методов преобразования сигналов [6]. Но действительно эффективным такое заимствование стало, когда нейрофизиологи "поднялись" от тестирования микрореакций отдельных нейронов зрительной коры до исследования нейронных ансамблей, "ответственных" за узнавание практически значимых компонентов зрительных образов (см., например, [7, 8]). Со временем на стыке формальных и эвристических подходов к автоматическому анализу изображений сформировался полезный инструментарий объектно-независимых алгоритмических и математических решений для выполнения распространенных частных операций обработки изображений (улучшения изображений, сегментации, выделения контуров и т. п.). Многие такие частные операции сформировались на основе эвристической интерпретации механизмов естественного зрения. Одним из ведущих направлений такой эвристической интерпретации можно считать структурное сопоставление зрительных образов. Как и большинство эвристических алгоритмов, методы структурной декомпозиции и сопоставления часто оказывались объектно-зависимыми, что сужало области их применения.

Автором этой статьи разработаны объектно-независимые методы структурного анализа, причем при структурной декомпозиции изображений используются детекторы признаков, аналогичные детекторам, обнаруженным в нейронных ансамблях ЖЗС, но отличающиеся от последних оптимально синтезированными обобщенными эталонными функциями (ОЭФ), с которыми локально сравниваются анализируемые изображения. Разработанные ОЭФ, описанию которых посвящена статья, оказались также пригодны для анализа текстуры, выделения контуров, формирования зон внимания (ЗВ), измерения параметров проективных преобразований, что свидетельствует о достаточной общности применяемого подхода.

2. Понятие объектно-независимого детектора структурных элементов

Упоминавшиеся в предыдущем разделе методы сопоставления изображений, основанные на их "глобальных" интегральных признаках, продемонстрировали высокую эффективность при нахождении точного соответствия анализируемого изображения и эталона, но они малопригодны в условиях естественной изменчивости наблюдаемых сцен. Тем не менее, аналоги корреляционного сопоставления часто встречаются в естественных и искусственных зрительных системах. Судя по всему, корреляционные алгоритмы являются основой наиболее фундаментальных, эволюционно сформировавшихся в

естественных зрительных системах механизмов: они эффективно реализуются базовой моделью нейрона, предложенной МакКаллоком и Питтсом [9], ими также можно описать многие поведенческие реакции простых организмов.

Одной из альтернатив корреляции является структурное сопоставление зрительных образов, основанное на распознавании и анализе примерного взаимного пространственного положения простых неизменных компонентов изображения. Такое сопоставление может быть малочувствительным к сложным взаимным геометрическим преобразованиям изображений, аддитивным и мультипликативным локальным изменениям яркости и текстуры [10–12]. Поэтому оно неэффективно при определении точного соответствия изображения и эталона и больше подходит для выяснения принадлежности зрительного образа к какому-либо достаточно широкому классу. Возможность использования структурных методов анализа в зрительных системах должна базироваться на надежных объектно-независимых способах выделения простых структурных составляющих образа, которые были бы инвариантны к характерным естественным преобразованиям изображений. Оказалось, что ЖЗС действительно могут выделять универсальные простые структурные составляющие анализируемых изображений. Современный уровень знания в области нейрофизиологии зрения не позволяет однозначно утверждать, каким способом выделяются такие универсальные структурные составляющие, однако гипотетический механизм, основанный на корреляции изображения с весовыми функциями специальной формы, оказывается очень привлекательной моделью. С одной стороны, такой механизм может эффективно реализоваться традиционной моделью нейрона, суммирующего входные сигналы с заданными синаптическими весами [9]. Некоторые нейрофизиологи предложили изящные нейронные модели (см., например, [8]), реализующие простые детекторы уже на низшем уровне обработки зрительной информации. С другой стороны, нейрофизиологические исследования зрительных систем [7] ясно показывают, что в зрительной системе человека на более высоких уровнях восприятия также имеются универсальные детекторы структурных элементов.

3. ОЭФ для измерения параметров пространственного положения локальных объектов

На мысль об использовании ОЭФ нас первоначально натолкнуло исследование Н.В. Завалишина и И.Б. Мучника [7], в котором они установили, что при разглядывании контурных изображений точки

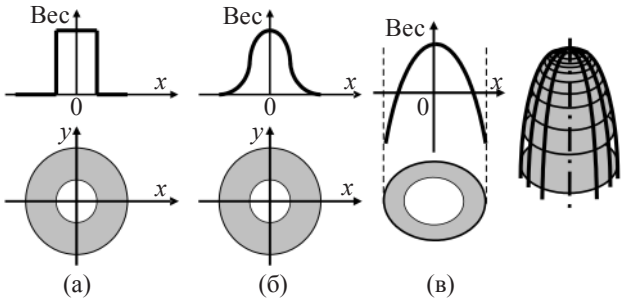


Рис. 1. Информационные функции для выделения структурных компонентов изображений. а, б – функции, предложенные Н.В. Завалишиным и И.Б. Мучником, в – ОЭФ, позволяющая измерять параметры аффинного преобразования.

саккадических фиксаций располагаются на биссектрисах углов, перекрестиях и кончиках линий, а также вдоль самих линий, если их концы находятся вне поля зрения. Результаты экспериментов они объяснили несложной математической моделью – поиском экстремумов разности изображения со специальными локально определенными двумерными “информационными функциями” (ИФ). Несложно показать, что аналогичные результаты могли бы быть получены при замене операции вычитания корреляцией. Примеры таких функций приведены на рис. 1а и 1б. Положение точек фиксации было предложено использовать для структурного описания и распознавания изображений. Кстати, современные исследования (см., например, [13]) подтверждают, что в ЖЗС действительно имеются детекторы отрезков прямых и их перекрестий.

Для решения проблем технического зрения необходимы ИФ, которые не только “притягивались” бы к указанным особенностям контуров, но позволяли бы измерять на изображениях параметры пространственного положения локальных объектов произвольной формы. Для этого несущественные локальные особенности идентифицируемых объектов не должны взаимодействовать с ИФ. Нам удалось показать [14], что необходимым требованиям удовлетворяет ИФ, имеющая в локальном окне форму эллиптического параболоида (см. рис. 1в). Благодаря своему очень низкочастотному спектральному составу и специфической форме такая функция (мы назвали ее ОЭФ) позволяет измерить положение центра тяжести изображения любого пространственно локального выделенного из фона объекта, а затем полностью нормализовать его изображение относительно полной группы аффинных преобразований (привести изображение к эталонному виду) [14, 15]. При этом все измеряемые параметры аффинного преобразования выражаются в аналитическом виде, аналогичном, например, формулам для нахождения центра тяжести.

В [15] показано также, что с помощью этой ОЭФ можно определить параметры аффинного преобразования и для нелокализованных контурных изображений, исследованных в [7], а в [16] нам удалось найти такой способ управления ЗВ (содержимым изображения в локальном окне, на котором задана ОЭФ), при котором параметры полного аффинного преобразования могут корректно определяться для изображений любых пространственно компактных объектов (объектов, полностью включенных в ЗВ), даже если они не выделены из фона. Наконец, в [17] с помощью этой ОЭФ удалось определить и компенсировать параметры геометрического преобразования изображений пространственно локализованных объектов, описываемого полной проективной группой на плоскости. Пример такой компенсации приведен на рис. 2, на котором изображения (а) и (б) получены путем применения к нормализованному изображению (в) различающихся неизвестных проективных преобразований. Разработанный алгоритм компенсации произвольного проективного преобразования вновь приводит искаженные изображения (а) и (б) к одному и тому же эталонному виду (в).

Таким образом, стало возможным распознавать изображения плоских объектов, сформированных при произвольном ракурсе наблюдения. Более того, оперируя содержимым локальных ЗВ, мы получили возможность определять параметры собственного и взаимного пространственного положения локальных объектов с однородными свойствами поверхности и распознавать такие объекты, что открыло путь для построения на их основе объектно-независимых структурных описаний трехмерных сцен [16]. С помощью построенных таким образом ЗВ можно также разделять изображения объемных объектов на изображения отдельных поверхностей [16] (см. рис. 3), что позволяет распространить описанный в разд. 4 аппарат структурного анализа изображений плоских объектов на случай трехмерных сцен.

Однако проективное преобразование нелинейно, поэтому нам не удалось единовременно измерять параметры проективной составляющей преобразования в аналитическом виде, например, как в формуле центра тяжести. Вместо этого искомые параметры могут итеративно определяться по градиенту смещения центра тяжести анализируемого локального объекта (градиент относительно параметров проективной составляющей преобразования) и, возможно, по градиенту его анизотропного по направлению масштабного преобразования [17]. Это создает интересные аналогии с методом определения ориентации поверхности по градиенту величины проективного преобразования ее текстуры [8], применяемым естественными зрительными системами. Согласно этому методу направление и величина наклона поверхности оцениваются по градиенту плотности расположения ячеек текстуры (текселов) – аналогу смещения (разбегания) их центров – или по градиенту сжатия (расширения) текселов в каком-либо направлении. Эти аналогии еще раз подчеркивают достаточно общий характер подходов, основанных на ОЭФ. Рис. 4 иллюстрирует возможность измерения наклона поверхности по градиенту параметров ее текстуры.

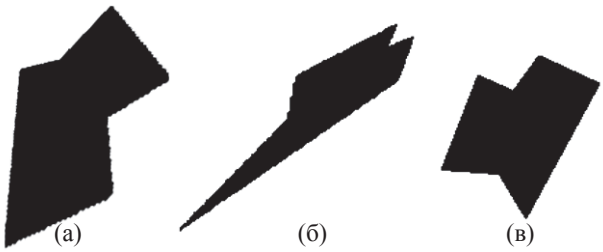


Рис. 2. Нормализация изображений локальных объектов относительно проективных преобразований с помощью ОЭФ. а, б – изображения одного и того же объекта для разных параметров преобразования, в – результат нормализации изображений а и б.

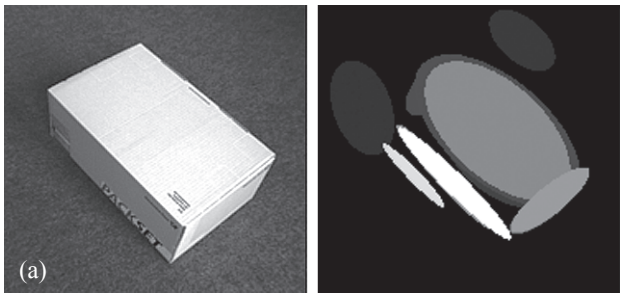


Рис. 3. Разделение изображения трехмерной сцены на ЗВ, соответствующие отдельным поверхностям, с помощью ОЭФ: а – исходное изображение, б – сформированные зоны внимания, соответствующие изображению а.

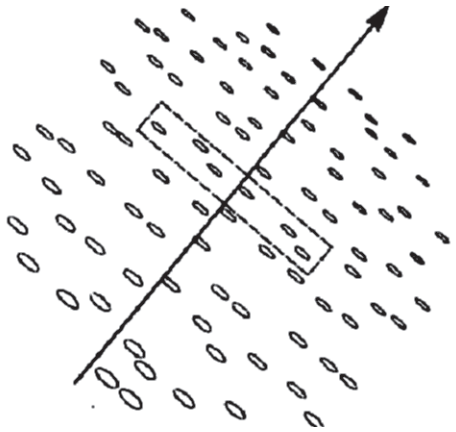


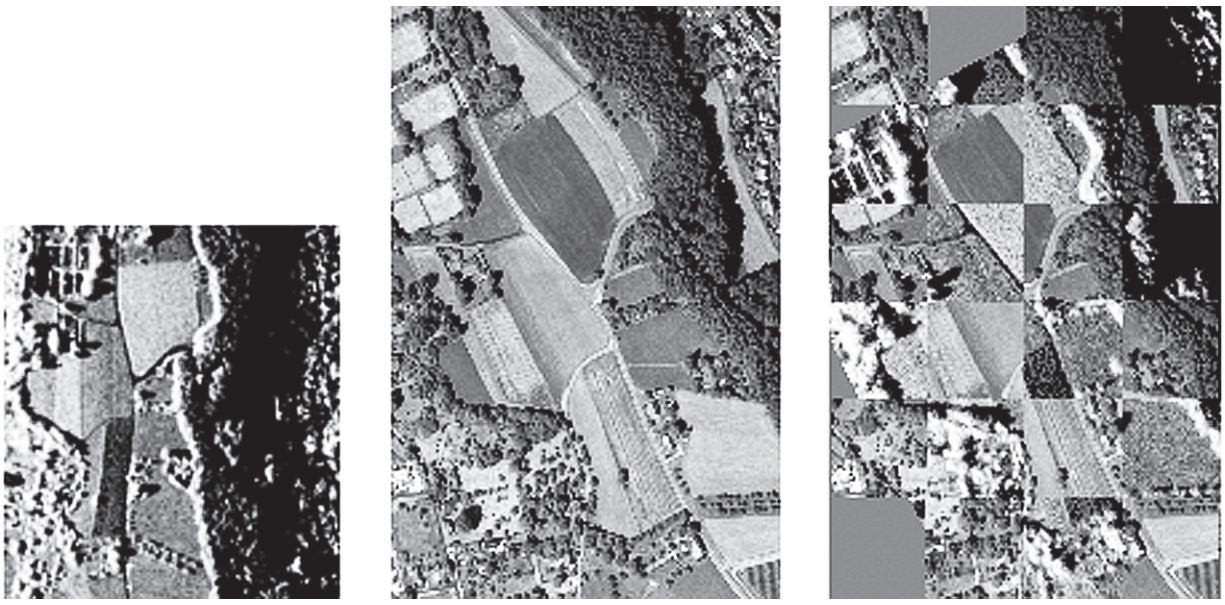
Рис. 4. Градиент смещения и размеров текселов, обусловленный локальным изменением проективного преобразования видимого изображения поверхности.

4. ОЭФ для обнаружения границ и измерения параметров их положения и ориентации

Пожалуй, чаще всего в числе компонентов изображений, подлежащих выделению, упоминаются границы, соответствующие резкому изменению наблюдаемых свойств поверхностей: их яркости, текстуры, пространственной глубины, параметров движения и т.п. После несложной обработки изображений таких границ получаются контурные изображения. Для выделенной границы должны измеряться ее пространственные координаты и ориентация. Границы и соответствующие им контуры являются источником наиболее устойчивых признаков, которые могут использоваться искусственными [10–12] и естественными [8] зрительными системами, а контурные структурные элементы, построенные на основе кривых второго порядка, отрезков прямых линий и углов между ними, являются инвариантами соответственно аффинных и проективных преобразований [10–12].

Используя вслед за нейрофизиологами Н.В. Завалишиным и И.Б. Мучником [7] структурные описания, основанные на таких инвариантных элементах, нам удалось достичь уникально надежного автоматического сопоставления изображений, имеющих очень сложные взаимные преобразования. В примере, приведенном на рис. 5, сопоставляются изображения (а) и (б), сформированные соответственно радиолокационным и оптическим датчиками, а результат (в) иллюстрируется путем мозаичного совмещения, при котором фрагменты сопоставленных изображений чередуются в порядке белых и черных клеток шахматной доски. Однако при этом применялись аффинные и проективные модели преобразований, заданные на плоскости, поэтому достигнутые результаты могли быть по-настоящему эффективны только для изображений плоских поверхностей. В то же время аппарат формирования ЗВ, описанный в разд. 3, позволяет разделять сцену на изображения достаточно гладких поверхностей, к каждому из которых может быть по отдельности применен представленный в текущем разделе метод контурного структурного анализа. Таким образом создана возможность структурного анализа трехмерных сцен.

Кстати, для выделения границ изображений объектов с целью построения контуров также используется корреляция изображений с ОЭФ. Широко применяемые для построения поля градиентов фильтры (посмотрите на горизонтальный фильтр Превитта $\begin{vmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{vmatrix}$) представляют хорошие примеры ОЭФ, позволяющих абстрагироваться от всех локальных особенностей, кроме направления и величины градиента. В [8] показано, что ЖЗС могут использовать детекторы границ, основанные на схожих ОЭФ.



(а) (б) (в)

Рис. 5. Пример сопоставления изображений по структурным описаниям их контуров. а, б – исходные изображения, в – результат сопоставления и мозаичного совмещения.

5. ОЭФ для измерения параметров текстур

В разд. 3 упоминалось использование параметров пространственного распределения и формы текселов при определении ориентации поверхности. Измерение указанных параметров как в статике, так и в их пространственной и временной динамике оказывается чрезвычайно важным для решения многих задач интерпретации изображений в ЖЗС [8].

Мелкие различия формы текселов обычно не сказываются качественно на конечном результате интерпретации; первостепенное значение имеют цвет, величина, направление и степень их удлинения [8]. Анализ литературы по нейрофизиологии зрения не позволяет однозначно определить, каким образом обнаруживаются текселы и идентифицируются параметры их формы в ЖЗС, однако встречается мнение (см., например, [18]), что в ЖЗС имеются детекторы текселов, основанные на хорошо известных рецептивных полях нейронов, напоминающих по форме "мексиканскую шляпу" и имеющих очень узкополосный низкочастотный пространственный спектр (см. рис. 6а). Мы реализовали такой подход к сегментации изображений по текстуре в искусственной нейронной сети (НС) [19]. Дискретная реализация детекторов текселов размерами в единицы пикселов привела к вырождению их рецептивных полей, как это показано на рис. 6б, что, однако, не помешало получить с их помощью неплохие результаты сегментации изображений по текстуре (см. рис. 6в и 6г). В указанной НС использованы только детекторы компактных текселов, показанные на рис. 6б, что ограничивало возможности НС в случаях, когда размеры детекторов заметно отличались от реальных размеров текселов. В свое время Вистнесс [20] достиг успеха, применяя дополнительно детекторы границ, использующие ОЭФ, подобные описанным в разд. 4.

Вырожденная весовая функция детектора, показанная на рис. 6б, мало отличается от разработанных нами ранее ОЭФ, ориентированных на идентификацию параметров пространственного положения произвольных компактных объектов (см. разд. 3, рис. 1в). Изображения естественных сцен имеют фрактальную природу, поэтому применение аналогичных ОЭФ (таких, как рассмотрены в разд. 3) на разных иерархических уровнях анализа изображений представляется вполне естественным. Как мы продемонстрировали в [21], образы, рассматриваемые как отдельные объекты, подлежащие распознаванию по форме при одном пространственном разрешении, становятся текселами при понижении разрешения изображений.

Если при анализе макрообъектов сцены (при данном пространственном разрешении или данном размере ЗВ) размер детектора заметно меньше размера подлежащего выделению компактного объекта или если объект чрезмерно вытянут в каком-либо направлении, детекторы края, описанные в разд. 4, органично дополняют детекторы компактных объектов, основанные на ОЭФ, описанных в разд. 3. В то же время, детекторы компактных объектов и детекторы края, использующие аналогичные ОЭФ, успешно дополняют друг друга и при анализе текстур, состоящих из микрообъектов заметно различающихся размеров.

Рассмотрим частный случай, когда вместо эллиптического параболоида рассматривается параболический цилиндр (такая цилиндрическая ОЭФ может применяться для определения координаты центра тяжести в одномерном пространстве [14]). Тогда

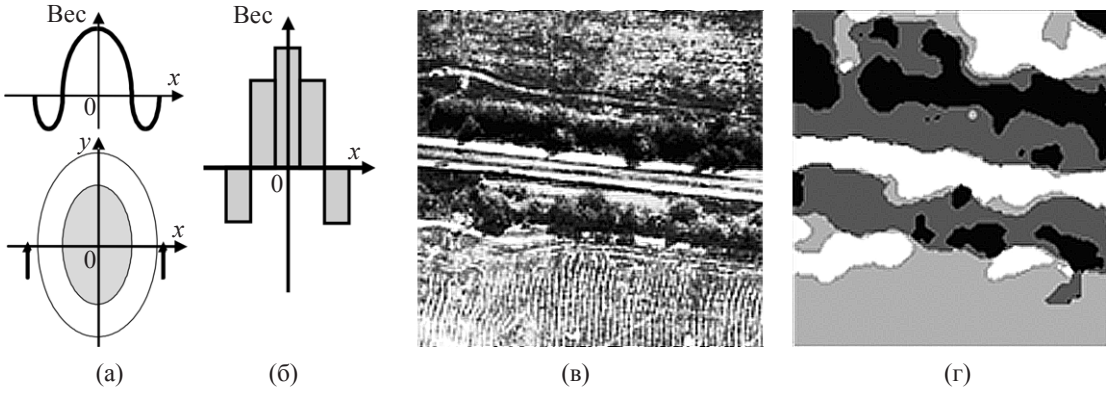


Рис. 6. Сегментация изображения на основе пространственного распределения его текселов. а – вероятная реализация детектора текселов на базе рецептивного поля нейрона, б – дискретное представление рецептивного поля (а), в – исходное изображение, г – результат сегментации изображения (в) по текстуре.

ОЭФ, рассмотренные в разд. 3 и 4, взаимно ортогональны при соответствующем ограничении областей их определения. Поэтому указанные пары ОЭФ изменяющегося пространственного масштаба [16] представляют собой своего рода псевдоортогональный базис (аналог вэйвлет) для декомпозиции изображений. В то же время, ОЭФ в форме эллиптического параболоида в ее общем виде может быть аналогичным образом поставлена в соответствие детекторам границ на основе фильтров Собела. Таким образом, пара ОЭФ, описанная в разд. 3 и 4, имеет универсальную значимость при структурной декомпозиции изображений.

Заключение

Представлен метод декомпозиции изображений путем их локального сопоставления со специальными обобщенными эталонами. Успешное его применение при решении разнообразных задач автоматического анализа изображений и аналогии, найденные в живых зрительных системах, дают основание полагать, что предлагаемый подход позволяет представить с единых позиций и дополнить ряд разделов теории обработки изображений.

Настоящее исследование выполнено при поддержке фонда “Научный потенциал” (Human Capital Foundation), грант № 26-03-31, договор № 80.

ЛИТЕРАТУРА

1. *Qin-sheng Chen, Defrise M., Deconinck F*. Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1994. V. 16. № 12. P. 1156–1167.
2. *Maitra S*. Moment invariants // Proc. IEEE. 1979. V. 67. № 4. P. 697–699.
3. *Ту Дж., Гонсалес Р*. Принципы распознавания образов / Пер. с англ. под ред. Ю.И. Журавлева. М.: Мир, 1978. 411 с.
4. Интегральные роботы. Сборник статей / Пер. с англ. под ред. Г.Е. Поздняка. В. 1. М.: Мир, 1973. 421 с.
5. Интегральные роботы. Сборник статей / Пер. с англ. под ред. Г.Е. Поздняка. В. 2. М.: Мир, 1975. 449 с.
6. *Розенблатт Ф*. Принципы нейродинамики (перцептрон и теория механизмов мозга). М.: Мир, 1965. 480 с.
7. *Завалишин Н.В., Мучник И.Б*. Модели зрительного восприятия и алгоритмы анализа изображений. М.: Наука, 1974. 344 с.
8. *Marr D.* Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. San Francisco, W.H. Freeman, 1982.
9. *McCulloch W.S., Pitts W.* A logical calculus of the ideas immanent in nervous activity // Bull. of Math. Biophys. V. 5. P. 115–133.
10. *Lutsiv V., Malyshev I., Pepelka V.* Automatic fusion of the multiple sensor and multiple season images // Proc. SPIE. 2001. V. 4380. P. 174–183.
11. *Lutsiv V., Malyshev I., Pepelka V., Potapov A*. The target independent algorithms for description and structural matching of aerospace photographs // Proc. SPIE. 2002. V. 4741. P. 351–362.
12. *Lutsiv V., Malyshev I., Potapov A.* Hierarchical structural matching algorithms for registration of aerospace images // Proc. SPIE. 2003. V. 5238. P. 164–175.
13. *Slavutskaya A.V. , Mikhailova E.S., Konishev V.A*. Human visual evoked potentials to bars and cross-like figures // Perception, 2006. V. 35 (Supplement. ECVP 2006. 29th European conference on visual perception, St.-Petersburg 20–25 August 2006. Abstracts). P. 99.
14. *Луцив В.Р.* Метод обработки изображений при управлении роботами в ГАП // ЭВМ в проектировании и производстве. Сб. статей. В. 2 / Под общ. ред. Г.В. Орловского. Л.: Машиностроение, 1985. С. 154–162.
15. *Луцив В.Р*. К вопросу структурного анализа изображений // Иконика. Книга VI. Труды ГОИ. 1992. Т. 79. В. 213. С. 65–71.
16. *Луцив В.Р*. Моделирование зон внимания в задачах автоматической декомпозиции и структурного анализа изображений // Оптический журнал. 2007. Т 74. № 4. С. 59–67.
17. Разработка методов оптимизационного системного проектирования механических конструкций, приводов, систем управления и адаптации роботизированных комплексов // Отчет по Г.Б. НИР Г-3-4-14-86 “Роботы”, промежуточный, № Г.Р. 01860016170. 1987. ЛИАП, Ленинград. 120 с.
18. *Julesz B*. Textrons. the elements of textural perception, and their interactions // Nature. V. 290. P. 91–97.
19. *Dolinov D.S., Zherebko A.K., Novikova T.A., Lutsiv V.R.* Using artificial neural networks in image processing problems // J. Opt. Technol. 1997. V. 64. № 2. P. 112–118.
20. *Vistnes R*. Texture models and image measures for segmentation // Proceedings: Image understanding workshop. 1988. P. 1005–1015.
21. *Lutsiv V., Malyshev I., Novikova T.* Hierarchical approaches to analysis of natural textures // Proc. SPIE. 2004. V. 5426. P. 144–154.