

## МОДЕЛЬНО-ОРИЕНТИРОВАННЫЙ ДЕСКРИПТОР ПОЛЯ ГРАДИЕНТА КАК УДОБНЫЙ АППАРАТ РАСПОЗНАВАНИЯ И АНАЛИЗА ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

Мясников В.В.

Институт систем обработки изображений РАН

### Аннотация

В работе предлагается новый дескриптор, используемый для описания цифрового изображения – модельно-ориентированный дескриптор поля градиента. Производные характеристики дескриптора, рассматриваемые как признаки цифрового изображения, позволяют эффективным образом решать задачи анализа, распознавания и поиска изображений. Представлены примеры решения таких задач с использованием предложенного дескриптора.

Ключевые слова: цифровые изображения, дескрипторы, признаки, анализ, распознавание, поиск.

### Введение

Локальные дескрипторы, вычисляемые для определенных зон интереса анализируемого изображения, – достаточно мощное современное средство решения широкого класса задач обработки, анализа, понимания и поиска цифровых изображений [1-6]. В настоящее время существует множество различных дескрипторов, которые условно можно разделить на следующие категории.

*Дескрипторы на основе вероятностных распределений.* Для этой категории дескрипторов фрагмент изображения используется для построения оценки некоторого распределения вероятностей. Наиболее простыми вариантами являются гистограмма распределения функции яркости в анализируемом фрагменте или гистограмма распределения расстояний между «особыми точками», найденными на этом фрагменте. Дескрипторы этой категории были предложены, в частности, Johnson и Hebert [7], Zabih и Woodfill [8] и другими авторами. Особую популярность за универсальность и качественные показатели в последнее время приобрели дескрипторы SIFT [9], HOG [3] и LESH [10], предложенные соответственно Lowe в 1999 году (SIFT – scale invariant feature transform), Dalal и Triggs в 2005 году (HOG-histogram of gradients) и Sarfraz и Hellwich в 2008 году (LESH – Local Energy based Shape Histogram). По принципу построения некоторые другие дескрипторы оказываются похожими на эти дескрипторы либо используют схожие подходы. В частности, геометрические гистограммы [11] и контекст формы [12] используют ту же идею, что и SIFT дескриптор, а GLOH (Gradient Location and Orientation Histogram) [2] – аналогичную дескриптору HOG.

*Дескрипторы на основе спектрального и спектрально-частотного представления.* Дескрипторы этого класса формируются как функции отсчетов спектрального или спектрально-частотного разложения анализируемого фрагмента изображения. Наиболее популярными способами построения дескрипторов этой категории являются использование фильтров Габора [4-6, 13] и вейвлетов [4-6, 14]. Другими популярными способами является использование Фурье-разложения [6], преобразований Радона

или Hough [4-6, 15, 16] степенных моментов или моментных инвариантов [4-6, 17-20] и др.

*Дифференциальные дескрипторы.* Дескрипторы этой категории рассчитываются как функции частных производных функции яркости анализируемого изображения. Такой способ описания анализируемого изображения использовали в своих работах Koenderink [21], Florack и соавторы [22], Freeman и Adelson [23] при построении управляемых фильтров (*steerable filters*), Baumberg [24], Schaffalitzky and Zisserman [25] и другие. В рамках рассматриваемой категории дескрипторов особо стоит отметить направление, связанное с использованием только фазовой составляющей градиента изображения – так называемого *поля направлений* [6, 26]. Использование только этой частичной информации позволяет достаточно эффективно анализировать особый тип изображений – *изображения со структурной избыточностью*, – к которым можно отнести широкий спектр медицинский изображений, микро- и наномасштабные изображения [6, 26].

Следует дополнительно отметить, что использование вероятностно-статистического аппарата для последней рассмотренной категории дескрипторов – дифференциальных дескрипторов – переводит их из третьей категории в первую, что подчеркивает (как было указано в начале обзора) условный характер представленной классификации. Более подробный анализ дескрипторов, их категорий, а также их аналитическое сравнение и сравнение их эффективности при решении прикладных задач могут быть найдены в работах [1, 2, 27] и других.

В настоящей работе предлагается новый тип дескриптора изображения, конструируемый на основе дифференциальных и вероятностных свойств локальной окрестности наблюдаемого изображения (комбинация первой и третьей рассмотренных категорий). Спецификой вводимого дескриптора является наличие априори заданного (или заранее определенного по обучающему множеству изображений) вероятностного распределения поля градиента, характеризующего *модель анализируемого изображения* и/или решаемую задачу. Значения компонент дескриптора для конкретного изображения (фрагмента) вычисляются как значения этой плотности вероятностей с аргументом в виде конкретного поля

градиента, то есть оказываются зависимы как от реализации (конкретного изображения), так и от модели (распределения вероятностей). Такая специфика расчёта дескриптора позволяет классифицировать его как *модельно-ориентированного* и использовать двумя способами. Во-первых, напрямую в решающем правиле классификатора, поскольку значение функции плотности на конкретной реализации – ключевая составляющая байесовского решающего правила [28]. Во-вторых, аналогично любым другим дескрипторам – в качестве численной характеристики некоторой локальной области изображения [2, 27, 29-30]. В этом случае для описания этой локальной области можно (и в ряде случаев – удобнее и численно проще) использовать не сами значения дескриптора, а некоторые их производные характеристики – *признаки дескриптора*.

Работа построена следующим образом. В первом разделе вводится предлагаемый *модельно-ориентированный дескриптор поля градиента* (MGFD – *model-based gradient field descriptor*). Во втором разделе предлагается ряд числовых *признаков* – производных характеристик, вычисляемых на основании значений введённого дескриптора. В третьем разделе представлена практически удобная для построения дескриптора модель градиентного поля с независимыми отсчётами, а также приведён ряд примеров с конкретными видами аналитических моделей. Показывается, что ряд известных способов обнаружения объектов на изображении (коррелятор, фазовая корреляция) оказываются частными случаями механизма расчёта признаков введённого дескриптора. Четвёртый раздел даёт краткое описание метода построения модели градиентного поля с зависимыми отсчётами. В пятом разделе предлагаются типовые способы построения алгоритмов обнаружения и распознавания на изображениях с использованием введённого дескриптора. В шестом разделе представлены примеры практических задач, решённых автором с использованием предложенного подхода. Наконец, в заключение работы приведены выводы, благодарности и список использованной литературы.

**Модельно-ориентированный дескриптор поля градиента**

Пусть  $f(t_1, t_2)$  – дифференцируемая функция яркости, определённая на области анализа  $\{(t_1, t_2): T_1^{\min} \leq t_1 \leq T_1^{\max}, T_2^{\min} \leq t_2 \leq T_2^{\max}\}$ . В качестве области анализа может выступать как область определения изображения целиком, так и фрагмент этого изображения. Пусть далее  $\Delta$  – шаг дискретизации непрерывного изображения, область анализа цифрового изображения определим в виде:

$$D = \left\{ (n_1, n_2) : 0 \leq n_i \leq \frac{T_i^{\max} - T_i^{\min}}{\Delta}, i = \overline{1,2} \right\},$$

и определим

$$g(n_1, n_2) \equiv \nabla f(n_1 \Delta + T_1^{\min}, n_2 \Delta + T_2^{\min}) = \left( \frac{\partial f / \partial t_1}{\partial f / \partial t_2} \right)_{\substack{t_1 = n_1 \Delta + T_1^{\min} \\ t_2 = n_2 \Delta + T_2^{\min}}}, (n_1, n_2) \in D$$

– цифровое поле градиента, в котором для каждого отсчёта анализируемого изображения хранится информация о модуле градиента  $|g(n_1, n_2)|$  и его ориентации по отношению к осям координат  $\arg(g(n_1, n_2))$ .

Для цифрового изображения градиентное поле может быть получено с использованием любого известного способа, например, с использованием *оператора Превитта* [4-6]. Обозначим далее анализируемый фрагмент цифрового изображения целиком как  $\bar{g} \equiv \{g(n_1, n_2)\}_{(n_1, n_2) \in D}$ .

Пусть далее  $\bar{G}$  – случайное градиентное поле с плотностью вероятностей  $p_{\bar{G}}(\cdot)$ , характеризующей *модель наблюдения* и/или решаемую задачу.

*Определение 1.* Модельно-ориентированным дескриптором поля градиента  $\bar{g}$  (дескриптором по отношению к модели  $\bar{G}$ ) называется вектор

$$\left( \begin{array}{c} p_{\bar{G}}(\bar{g}) \\ \min_{|\bar{g}|} p_{\bar{G}}(\bar{g}) \\ \max_{|\bar{g}|} p_{\bar{G}}(\bar{g}) \\ \min_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) \\ \max_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) \\ \min_{\arg \bar{g}} \min_{|\bar{g}|} p_{\bar{G}}(\bar{g}) \\ \max_{\arg \bar{g}} \min_{|\bar{g}|} p_{\bar{G}}(\bar{g}) \\ \max_{|\bar{g}|} \min_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) \\ \min_{\arg \bar{g}} \max_{|\bar{g}|} p_{\bar{G}}(\bar{g}) \\ \min_{|\bar{g}|} \max_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) \\ \max_{|\bar{g}|} \max_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) \end{array} \right). \tag{1}$$

Величина вида  $\min_{\arg \bar{g}} p_{\bar{G}}(\bar{g})$  определяется как значение плотности  $p_{\bar{G}}(\cdot)$  для экземпляра градиентного поля, в каждом отсчёте  $(n_1, n_2)$  которого модуль градиента совпадает со значением модуля градиента  $|g(n_1, n_2)|$  поля  $\bar{g}$ , а фазовые значения градиентов подобраны таким образом, чтобы минимизировать значение плотности  $p_{\bar{G}}(\cdot)$ . Остальные величины определяются аналогичным образом.

В соответствии с определением элементов дескриптора все они принимают значения в интервале

$$\left[ \min_{|\bar{g}|} \min_{\arg \bar{g}} p_{\bar{G}}(\bar{g}), \max_{|\bar{g}|} \max_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) \right],$$

причём  $\min_{|\bar{g}|} \min_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) \geq 0$ . Для них оказывается справедливым ряд очевидных соотношений, в частности:

$$\min_{|\bar{g}|} p_{\bar{G}}(\bar{g}) \leq p_{\bar{G}}(\bar{g}) \leq \max_{|\bar{g}|} p_{\bar{G}}(\bar{g}), \tag{2a}$$

$$\min_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) \leq p_{\bar{G}}(\bar{g}) \leq \max_{\arg \bar{g}} p_{\bar{G}}(\bar{g}), \tag{2б}$$

$$\min_{|\bar{g}|} \min_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) \leq \min_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) \leq \max_{|\bar{g}|} \min_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) \tag{2в}$$

и т.п.

На весь набор этих неравенств далее в тексте будем ссылаться как на соотношения (2).

**Признаки модельно-ориентированного дескриптора поля градиента**

Признаками анализируемого поля градиента могут являться производные характеристики введённого дескриптора. Учитывая соотношения (2), их удобно определить как относительное положение величины (например,  $p_{\bar{G}}(\bar{g})$ ) в интервале от её минимального до её максимального значения. Учитывая, что величины типа  $\max_{|\bar{g}|} \min_{\arg \bar{g}} p_{\bar{G}}(\bar{g})$  характеризуют исключительно модель наблюдения, а не анализируемую реализацию  $\bar{g}$  градиентного поля, получаем семь основных признаков следующего вида (здесь  $\phi: \mathbf{R}_+ \rightarrow \mathbf{R}$  – произвольная монотонно-возрастающая функция):

$$\begin{aligned} \vartheta_0 &= \frac{\phi(p_{\bar{G}}(\bar{g})) - \phi\left(\min_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right)}{\phi\left(\max_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right) - \phi\left(\min_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right)}, \\ \vartheta_1 &= \frac{\phi(p_{\bar{G}}(\bar{g})) - \phi\left(\min_{\arg \bar{g}} p_{\bar{G}}(\bar{g})\right)}{\phi\left(\max_{\arg \bar{g}} p_{\bar{G}}(\bar{g})\right) - \phi\left(\min_{\arg \bar{g}} p_{\bar{G}}(\bar{g})\right)}, \\ \vartheta_2 &= \frac{\phi(p_{\bar{G}}(\bar{g})) - \phi\left(\min_{\arg \bar{g}} \min_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right)}{\phi\left(\max_{|\bar{g}|} \max_{\arg \bar{g}} p_{\bar{G}}(\bar{g})\right) - \phi\left(\min_{\arg \bar{g}} \min_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right)}, \\ \vartheta_3 &= \frac{\phi\left(\min_{\arg \bar{g}} p_{\bar{G}}(\bar{g})\right) - \phi\left(\min_{\arg \bar{g}} \min_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right)}{\phi\left(\max_{|\bar{g}|} \min_{\arg \bar{g}} p_{\bar{G}}(\bar{g})\right) - \phi\left(\min_{\arg \bar{g}} \min_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right)}, \end{aligned} \tag{3a}$$

$$\begin{aligned} \vartheta_4 &= \frac{\phi\left(\max_{\arg \bar{g}} p_{\bar{G}}(\bar{g})\right) - \phi\left(\min_{\arg \bar{g}} \max_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right)}{\phi\left(\max_{|\bar{g}|} \max_{\arg \bar{g}} p_{\bar{G}}(\bar{g})\right) - \phi\left(\min_{\arg \bar{g}} \max_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right)}, \\ \vartheta_5 &= \frac{\phi\left(\min_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right) - \phi\left(\min_{\arg \bar{g}} \min_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right)}{\phi\left(\max_{\arg \bar{g}} \min_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right) - \phi\left(\min_{\arg \bar{g}} \min_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right)}, \\ \vartheta_6 &= \frac{\phi\left(\max_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right) - \phi\left(\min_{\arg \bar{g}} \max_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right)}{\phi\left(\max_{\arg \bar{g}} \max_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right) - \phi\left(\min_{\arg \bar{g}} \max_{|\bar{g}|} p_{\bar{G}}(\bar{g})\right)}. \end{aligned} \tag{3б}$$

Введённые признаки  $\vartheta_i$  характеризуют степень похожести анализируемого градиентного поля на потенциально-возможные реализации случайного поля  $\bar{G}$  (на модель). Большие значения означают большую похожесть, а меньшие – меньшую. При этом если  $\vartheta_0$ , например, определяет степень схожести поля модуля градиента, то  $\vartheta_1$  – степень схожести фазового рисунка и т.п. На всё множество этих признаков ниже будем ссылаться как на вектор  $\bar{\vartheta} \equiv (\vartheta_0, \vartheta_1, \dots, \vartheta_6)^T$ .

Выбор функции  $\phi$  существенным образом зависит от вида плотности  $p_{\bar{G}}(\cdot)$ . Более того, на практике вместо признаков  $\vartheta_i$  может быть удобным использование «вторичных» признаков вида  $\tilde{\vartheta}_i \equiv \kappa(\vartheta_i)$ , где  $\kappa: \mathbf{R}[0,1] \rightarrow \mathbf{R}$  – произвольная взаимно-однозначная функция. Следует, однако, отметить, что такая замена может уничтожить следующее очевидное, но очень полезное свойство исходных признаков.

*Предложение 1.*

$$\forall i = \overline{0,6} \quad 0 \leq \vartheta_i \leq 1.$$

Это свойство позволяет относительно просто решать задачи параметрической настройки решающих правил, входящих в состав алгоритмов обнаружения и распознавания на изображениях (см. два заключительных раздела настоящей работы).

**Модель градиентного поля с независимыми отсчётами**

Пусть отсчёты поля градиента независимы. Тогда плотность вероятностей градиентного поля, характеризующая его вероятностную модель, запишется в виде:

$$p_{\bar{G}}(\bar{g}) = \prod_{(n_1, n_2) \in D} p_{G(n_1, n_2)}(g(n_1, n_2)), \tag{4}$$

где  $p_{G(n_1, n_2)}(\cdot)$  – плотность вероятностей градиента в конкретном отсчёте  $(n_1, n_2)$  цифрового изображения. Используя различные способы задания плотности  $p_{G(n_1, n_2)}(\cdot)$ , можно получать различные группы при-

знаков (в том числе известные). Заметим, что в отличие от широко распространённой практики работы с градиентным полем в НОГ-дескрипторах [3, 27], функцию плотности  $p_{G(n_1, n_2)}(\cdot)$  предлагается задавать аналитически с точностью до некоторого числа неизвестных параметров. Параметрическая оценка плотности (оценка неизвестных параметров) может быть получена хорошо известными методом *моментов* или методом *максимального правдоподобия* [31].

*Замечание 1.* Дополнительным удобством модели с независимыми отсчётами является возможность упрощения вычислений в ситуации, когда рассчитываются дескрипторы по целому ряду областей, в том числе – перекрывающимся и/или являющимися объединением областей. В частности, для произвольных непересекающихся областей  $D_1$  и  $D_2$  справедливо соотношение:

$$p_G^{D_1 \cup D_2}(\cdot) = p_{G_1}^{D_1}(\cdot) \cdot p_{G_2}^{D_2}(\cdot).$$

Ниже мы рассмотрим несколько способов задания плотности вероятностей  $p_{G(n_1, n_2)}(\cdot)$ , которые приводят либо к известным результатам (признакам), либо были успешно использованы автором для решения реальных задач обнаружения и распознавания изображений. В последнем случае описание решённых прикладных задач представлено в заключительном разделе настоящей работы.

*Пример 1. Амплитудное соответствие*

Зададим плотность вероятностей для отсчётов градиентного поля в виде

$$p_{G(n_1, n_2)}(g(n_1, n_2)) = \begin{cases} a_{(n_1, n_2)} \cdot \exp\left[-\left(|g(n_1, n_2)| - \beta_{(n_1, n_2)}\right)^2\right], & 0 \leq |g(n_1, n_2)| \leq g_{\max}; \\ 0, & |g(n_1, n_2)| > g_{\max}. \end{cases}$$

Здесь  $a_{(n_1, n_2)}$  – масштабный множитель;  $\beta_{(n_1, n_2)}$  – величина, определяющая наиболее вероятное значение модуля градиента в соответствующем отсчёте изображения;  $g_{\max}$  – наибольшее значение для модуля градиента на изображении. Тогда

$$p_{\bar{G}}(\bar{g}) = A \exp\left(-\sum_{(n_1, n_2)} \left(|g(n_1, n_2)| - \beta_{(n_1, n_2)}\right)^2\right),$$

$$\max_{|\bar{g}|} p_{\bar{G}}(\bar{g}) = A, \quad \min_{|\bar{g}|} p_{\bar{G}}(\bar{g}) = A \exp(-\Delta_{\max}),$$

где

$$A = \prod_{(n_1, n_2)} a_{(n_1, n_2)},$$

$$\Delta_{\max} = \sum_{(n_1, n_2)} \left(\max(\beta_{(n_1, n_2)}, g_{\max}) - \beta_{(n_1, n_2)}\right)^2.$$

Полагая функцию  $\phi: \mathbf{R}_+ \rightarrow \mathbf{R}$  в виде

$$\phi(x) = \ln\left(\frac{x}{A}\right), \tag{5}$$

имеем

$$\vartheta_0 = 1 - \frac{\sum_{(t_1, t_2)} \left(|g(t_1, t_2)| - \beta_{(t_1, t_2)}\right)^2}{\Delta_{\max}}.$$

Здесь, например, удобно взять

$$\tilde{\vartheta}_0 = \sum_{(t_1, t_2)} \left(|g(t_1, t_2)| - \beta_{(t_1, t_2)}\right)^2,$$

и итоговый признак принимает легко интерпретируемый смысл: эта величина характеризует суммарное рассогласование модулей градиентов с величинами  $\beta_{(t_1, t_2)}$ , определяющими их наиболее вероятные значения. Здесь очевидна связь с обычным корреляционным способом обнаружения на изображении объекта по шаблону [6].

*Пример 2. Фазовое соответствие*

Пусть

$$p_{G(n_1, n_2)}(g(n_1, n_2)) = \begin{cases} a_{(n_1, n_2)} \cdot \exp\left[\cos(\arg g(n_1, n_2) - \varphi_{(n_1, n_2)}) + 1\right], & 0 \leq |g(n_1, n_2)| \leq g_{\max}; \\ 0, & |g(n_1, n_2)| > g_{\max}. \end{cases}$$

Параметры  $g_{\max}$  и  $a_{(n_1, n_2)}$  имеют тот же смысл, что и в первом примере; величина  $\varphi_{(n_1, n_2)}$  определяет наиболее вероятное значение фазы градиента  $\arg g(n_1, n_2)$  в соответствующем отсчёте изображения. Тогда

$$p_{\bar{G}}(\bar{g}) = A \exp\left(\sum_{(n_1, n_2)} \left[\cos(\arg g(n_1, n_2) - \varphi_{(n_1, n_2)}) + 1\right]\right),$$

$$\max_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) = A \exp(2|D|), \quad \min_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) = A.$$

Здесь  $|D| = \sum_{(n_1, n_2)} 1$  – мощность области  $D$ . Полагая

функцию  $\phi$  в виде (5), имеем:

$$\vartheta_1 = \frac{1}{|D|} \sum_{(n_1, n_2)} \frac{1}{2} \left[\cos(\arg g(n_1, n_2) - \varphi_{(n_1, n_2)}) + 1\right]$$

и признак также приобретает легко интерпретируемый смысл: эта величина характеризует среднее согласование (по фазе) анализируемого поля градиента с полем наиболее вероятных направлений  $\varphi_{(n_1, n_2)}$ . Данная величина удобна при анализе полей направления [6, 26] в качестве показателя, определяющего их фазовую согласованность: чем больше величина  $\vartheta_1$ , тем выше степень согласованности.

*Пример 3. Амплитудная величина фазового соответствия*

Зададим плотность вероятностей для отсчётов градиентного поля в виде

$$p_G(n_1, n_2)(g(n_1, n_2)) = \begin{cases} a(n_1, n_2) \cdot \exp \left[ |g(n_1, n_2)| \cdot \cos \left( \arg g(n_1, n_2) - \varphi(n_1, n_2) \right) \right], & 0 \leq |g(n_1, n_2)| \leq g_{\max}; \\ 0, & |g(n_1, n_2)| > g_{\max}. \end{cases}$$

Параметры  $g_{\max}$ ,  $a(t_1, t_2)$  и  $\varphi(t_1, t_2)$  имеют тот же смысл, что и выше. Тогда

$$p_{\bar{G}}(\bar{g}) = A \exp \left( \sum_{(n_1, n_2)} |g(n_1, n_2)| \cos(\arg g(n_1, n_2) - \varphi(n_1, n_2)) \right),$$

$$\max_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) = A \exp \left( \sum_{(n_1, n_2)} |g(n_1, n_2)| \right),$$

$$\min_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) = A.$$

Полагая функцию  $\phi$  в виде (5), имеем признак следующего вида:

$$\vartheta_1 = \frac{\sum_{(n_1, n_2)} |g(n_1, n_2)| \cos(\arg g(n_1, n_2) - \varphi(n_1, n_2))}{\sum_{(n_1, n_2)} |g(n_1, n_2)|}.$$

Эта величина характеризует относительную суммарную величину проекции модуля градиента на наиболее вероятное его направление  $\varphi(n_1, n_2)$ .

*Пример 4. Амплитудно-фазовое соответствие*

Зададим плотность вероятностей для отсчётов градиентного поля в виде

$$p_G(n_1, n_2)(g(n_1, n_2)) = \begin{cases} a(n_1, n_2) \cdot \exp \left[ - \left( |g(n_1, n_2)| - \beta(n_1, n_2) \right)^2 \cdot \left| \sin(\arg g(n_1, n_2) - \varphi(n_1, n_2)) \right| \right], & 0 \leq |g(n_1, n_2)| \leq g_{\max}; \\ 0, & |g(n_1, n_2)| > g_{\max}. \end{cases} \quad (6)$$

Значения всех параметров описаны выше. Тогда

$$p_{\bar{G}}(\bar{g}) = A \exp \left[ - \frac{\sum_{(n_1, n_2)} \left( |g(n_1, n_2)| - \beta(n_1, n_2) \right)^2}{\left| \sin(\arg g(n_1, n_2) - \varphi(n_1, n_2)) \right|} \right],$$

$$\begin{aligned} \max_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) &= \max_{\bar{g}} p_{\bar{G}}(\bar{g}) = \max_{\arg \bar{g}} \max_{\bar{g}} p_{\bar{G}}(\bar{g}) = \\ &= \min_{\bar{g}} \max_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) = \min_{\arg \bar{g}} \max_{\bar{g}} p_{\bar{G}}(\bar{g}) = \\ &= \max_{\arg \bar{g}} \min_{\bar{g}} p_{\bar{G}}(\bar{g}) = \max_{\bar{g}} \min_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) = A, \end{aligned}$$

$$\min_{\arg \bar{g}} p_{\bar{G}}(\bar{g}) = A \exp \left[ - \sum_{(n_1, n_2)} \left( |g(n_1, n_2)| - \beta(n_1, n_2) \right)^2 \right],$$

$$\min_{\bar{g}} p_{\bar{G}}(\bar{g}) = A \exp \left[ - \sum_{(n_1, n_2)} \left( \Delta_{(n_1, n_2)}^{\max} \right)^2 \left| \sin(\arg g(n_1, n_2) - \varphi(n_1, n_2)) \right| \right],$$

$$\min_{\arg \bar{g}} \min_{\bar{g}} p_{\bar{G}}(\bar{g}) = A \exp[-\Delta_{\max}]$$

$$\text{Здесь } \Delta_{(n_1, n_2)}^{\max} = \max(g_{\max} - \beta(n_1, n_2), \beta(n_1, n_2)).$$

Также используя функцию  $\phi$  в виде (5), представим весь набор признаков:

$$\vartheta_0 = 1 -$$

$$\frac{\sum_{(n_1, n_2)} \left( |g(n_1, n_2)| - \beta(n_1, n_2) \right)^2 \left| \sin(\arg g(n_1, n_2) - \varphi(n_1, n_2)) \right|}{\sum_{(n_1, n_2)} \left( \Delta_{(n_1, n_2)}^{\max} \right)^2 \left| \sin(\arg g(n_1, n_2) - \varphi(n_1, n_2)) \right|},$$

$$\vartheta_1 = 1 -$$

$$\frac{\sum_{(n_1, n_2)} \left( |g(n_1, n_2)| - \beta(n_1, n_2) \right)^2 \left| \sin(\arg g(n_1, n_2) - \varphi(n_1, n_2)) \right|}{\sum_{(n_1, n_2)} \left( |g(n_1, n_2)| - \beta(n_1, n_2) \right)^2},$$

$$\vartheta_2 = 1 -$$

$$\frac{1}{\Delta_{\max}} \sum_{(n_1, n_2)} \left( |g(n_1, n_2)| - \beta(n_1, n_2) \right)^2 \left| \sin \left( \arg g(n_1, n_2) - \varphi(n_1, n_2) \right) \right|,$$

$$\vartheta_3 = 1 - \frac{1}{\Delta_{\max}} \sum_{(n_1, n_2)} \left( |g(n_1, n_2)| - \beta(n_1, n_2) \right)^2,$$

$$\vartheta_4 = \vartheta_6 = 1,$$

$$\vartheta_5 = 1 -$$

$$\frac{1}{\Delta_{\max}} \sum_{(n_1, n_2)} \left( \Delta_{(n_1, n_2)}^{\max} \right)^2 \left| \sin(\arg g(n_1, n_2) - \varphi(n_1, n_2)) \right|.$$

Приведённые компоненты вектора признаков  $\bar{\vartheta}$  совместно характеризуют согласованность анализируемого градиентного поля с моделью, выбранной в виде указанного распределения с параметрами.

**Метод построения модели градиентного поля с зависимыми отсчётами**

Для построения модели градиентного поля с зависимыми отсчётами используем следующее представление функции плотности:

$$p_{\bar{G}}(\bar{g}) = \sum_{k=0}^{K-1} \prod_{(n_1, n_2)} p_{G(n_1, n_2)}^k(g(n_1, n_2)), \quad (7)$$

где  $p_{G(n_1, n_2)}^k(g(n_1, n_2))$  – плотность вероятностей градиента в конкретном отсчёте  $(n_1, n_2)$  изображения для  $k$ -го «слоя» представления плотности, заданная аналитически с точностью до набора пара-

метров. В качестве такой «элементарной» плотности может выступать любая из плотностей, представленная в приведённых выше примерах 1-4. Более того, при таком представлении можно использовать плотность  $p_{G(n_1, n_2)}^k(g(n_1, n_2))$  в виде, допускающем независимость модуля и фазы градиента:

$$p_{G(n_1, n_2)}^k(g(n_1, n_2)) = p_{|G(n_1, n_2)}^k(|g(n_1, n_2)|) p_{\arg G(n_1, n_2)}^k(\arg g(n_1, n_2)),$$

здесь

- $p_{|G(n_1, n_2)}^k(|g(n_1, n_2)|)$  – плотность распределения вероятностей модуля градиента,
- $p_{\arg G(n_1, n_2)}^k(\arg g(n_1, n_2)) p_{|G(n_1, n_2)}^k(|g(n_1, n_2)|)$  – плотность распределения вероятностей фазы градиента.

Параметрическая оценка плотности в виде (7) может быть получена различными способами. Наиболее удобным способом является использование ОМ-алгоритма (expectation-maximization, EM-algorithm) [32].

*Замечание 2.* Главная проблема использования модели (7) с зависимыми отсчётами заключается в необходимости решения относительно трудоёмкой в вычисленном плане (оптимизационной) задачи расчёта следующих элементов дескриптора:

$$\min_{|\bar{g}|} p_{\bar{G}}(\bar{g}), \max_{|\bar{g}|} p_{\bar{G}}(\bar{g}), \min_{\arg \bar{g}} p_{\bar{G}}(\bar{g}), \max_{\arg \bar{g}} p_{\bar{G}}(\bar{g}),$$

поскольку процесс их вычисления требуется проводить при каждом акте формирования дескриптора. В отличие от указанной четвёрки элементов, величины типа  $\min_{|\bar{g}|} \min_{\arg \bar{g}} p_{\bar{G}}(\bar{g})$  могут

быть вычислены один раз для принятой модели (7), то есть заблаговременно.

**Использование модельно-ориентированного дескриптора поля градиента для распознавания и поиска изображений**

Положительным свойством предложенного дескриптора и его признаков является относительная простота использования их для решения задач анализа, распознавания и поиска изображений, а также обнаружения и распознавания объектов на изображении. В частности, очевидным способом построения решающего правила является использование классической дискриминантной функции байесовского классификатора [28]:

$$d_l(\bar{g}) = P(\Omega_l) p_{\bar{G}(l)}(\bar{g}),$$

где  $P(\Omega_l)$ ,  $p_{\bar{G}(l)}(\bar{g})$  – соответственно, априорная вероятность  $l$ -го класса и соответствующее этому классу значение плотности вероятностей для анали-

зируемого поля градиента  $\bar{g}$  (первый компонент дескриптора). Несмотря на очевидность (и «наивность») такого решения, для некоторых задач (см. второй пример следующего раздела – обнаружение лица на изображении) его оказывается достаточно для получения требуемого качества обработки.

Более сложные решающие правила могут быть сконструированы на основе комбинаций решений отдельных «простых» классификаторов – *экспертов*, каждый из которых использует признаки дескриптора со своей подобласти. Таким образом, потенциальный классифицируемый фрагмент оказывается разбитым на подобласти (возможно, перекрывающиеся) –  $D_i$  ( $i = \overline{0, I-1}$ ), и для каждой подобласти формируется её описание в виде дескриптора и/или его признаков  $\bar{g}^i$  ( $i = \overline{0, I-1}$ ). Выбрав алгоритм классификации для каждого эксперта и метод его настройки, разработчик системы распознавания/анализа/поиска может далее использовать различные существующие стратегии комбинации их решений: голосование большинства или взвешенное голосование, решающий список (см. первый пример следующего раздела и работу [33]), алгоритм вычисления оценок [34], алгебру над алгоритмами [35], алгоритмы бустинга [29-30] и другие.

**Примеры практических задач, решённых с использованием дескриптора поля градиента**

Предложенный модельно-ориентированный дескриптор поля градиента и признаки, рассчитываемые на его основе, были использованы при разработке двух систем анализа изображений. Ниже они представлены как примеры использования предложенного подхода.

*Пример 1.* Целью первой системы было обнаружение легковых транспортных средств на цифровых изображениях оптического дистанционного зондирования Земли (ДЗЗ). Обработываемые изображения были получены путём аэрофотосъёмки, степень их пространственного разрешения соответствовала перспективным средствам космического ДЗЗ, допускающим съёмку с пространственным разрешением порядка 0,2 м (например, американский спутник GeoEye-2 допускает разрешение 0,25–0,3 м). При построении этой системы были использованы различные группы признаков (геометрические, текстурные, структурные). Но большинство признаков (более 80%) задавались как амплитудная величина фазового соответствия (см. пример 3) по специфическим областям объекта обнаружения (подход с комбинацией решений экспертов). В качестве решающего правила (для проверки гипотезы о наличии объекта), следуя предложению предыдущего раздела, выступал решающий список. Система обнаружения тестировалась на 30 изображениях городской застройки (разрешение снимков – 0,1 м), не входящих в состав обучающей выборки. Среднее число объектов, присутствующих на изображениях,

– 35 штук. Пример работы предложенного метода приведён на рисунке ниже.

Полученные в результате эксперимента показатели качества обнаружения:

- вероятность верного обнаружения – 95 %,
- среднее число ложно обнаруженных объектов (в кадре) – 2,8.

Более подробное изложение метода обнаружения, положенного в основу разработанной системы, представлено в работе [33].



Рис. 1. Пример работы системы обнаружения легковых транспортных средств

**Пример 2.** Целью второй системы было обнаружение на цифровом изображении удостоверения личности лица его владельца. Задача усложнялась тем, что не вводилось ограничений на вид документа, на положение искомой фотографии, на тип «растрирования» при его печати; на фотографиях допускалось наличие бликов и засветки как следствий наложения на документ ламинирующего покрытия и неравномерности освещения; присутствовали элементы типографской печати на фотографии человека; допускалось радикальное различие в контрасте переднего (лица) и заднего плана и т.п.

В качестве основного элемента при построении решения был использован единственный признак  $\mathcal{D}_2$ , полученный для модели поля градиента (изображения лица) в виде (7) с функциями плотности  $P_{G(n_1, n_2)}^k(\cdot)$  в виде (6). Вероятность верного обнаружения для разработанной системы (при допущении погрешности в определении местоположения лица и его линейных размеров в 20% и 30% соответственно) составила 98 %. Тестирование проводилось для 500 паспортов, только половина из которых российские.

### Выводы и рекомендации

Предложен модельно-ориентированный дескриптор поля градиента, предназначенный для описания цифровых изображений, а также признаки дескриптора, позволяющие эффективным образом решать задачи анализа, распознавания и поиска изображений. На примерах решения реальных практических задач показана эффективность предложенного дескриптора и подхода к анализу изображений.

Дальнейшими направлениями работ являются:

- сравнение эффективности предложенного модельно-ориентированного дескриптора поля градиента и его признаков с существующими решениями (HOG, SIFT и др.);
- разработка модификации дескриптора, инвариантной к преобразованиям подобия (сдвиг, поворот, масштаб) и/или аффинным преобразованиям;
- разработка моделей поля градиента, допускающих построение численно эффективных алгоритмов вычисления дескрипторов и классификации.

### Благодарности

Работа выполнена при частичной финансовой поддержке:

- грантов РФФИ, проекты 12-07-00021-а, 11-07-12060-офи-м-2011, 11-07-12062-офи-м-2011;
- программы фундаментальных исследований Президиума РАН «Фундаментальные проблемы информатики и информационных технологий», проект 2.12.

### Литература

1. Schmid, C. Evaluation of interest point detectors / C. Schmid, R. Mohr, C. Bauckhage // Int. Journal of Computer Vision. – 2000. – Vol. 37(2). – P. 151-172.
2. Mikolajczyk, K. A performance evaluation of local descriptors / K. Mikolajczyk, C. Schmid // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2005. – Vol. 10(27). – P. 1615-1630.
3. Dalal, N. Histograms of Oriented Gradients for Human Detection / N. Dalal, B. Triggs // Proceedings of IEEE Conference Computer Vision and Pattern Recognition, San Diego, USA. – 2005. – P. 886-893.
4. Forsyth, D.A. Computer Vision: A Modern Approach / D.A. Forsyth, J. Ponce. – New Jersey: Prentice Hall, 2002. – 693 p.
5. Shapiro, L.G. Computer Vision / L.G. Shapiro, G.C. Stockman. – New Jersey: Prentice Hall, 2001. – 608 p.
6. Методы компьютерной обработки изображений. / М.В. Гашников, Н.И. Глумов, Н.Ю.Ильцова, В.В. Мясников [и др.], под общей ред. В.А. Соифера. – 2-е изд., испр. – М.: Физматлит, 2003. – 784 с.
7. Johnson, A. Object recognition by matching oriented points / A. Johnson, M. Hebert // Proceedings of the Conference on Computer Vision and Pattern Recognition, Puerto Rico, USA. – 1997. – P. 684-689.
8. Zabih, R. Non-parametric local transforms for computing visual correspondence / R. Zabih, J. Woodfill // Proceedings of the 3rd European Conference on Computer Vision, Stockholm, Sweden. – 1994. – P. 151-158.
9. Lowe, D. Distinctive image features from scale-invariant keypoints / D. Lowe // International Journal of Computer Vision. – 2004. – Vol. 2(60). – P.91-110.
10. Sarfraz, O. Head Pose Estimation in Face Recognition across Pose Scenarios / O. Sarfraz, S., Hellwich // Proceedings of VISAPP 2008, Int. conference on Computer Vision Theory and Applications, Madeira, Portugal. – 2008. – P. 235-242.
11. Ashbrook, A. Robust recognition of scaled shapes using pairwise geometric histograms / A. Ashbrook, N. Thacker, P. Rockett, C. Brown // Proceedings of the sixth British Machine Vision Conference, Birmingham, UK. – 1995 – P. 503-512.

12. **Belongie, S.** Shape matching and object recognition using shape contexts / S. Belongie, J. Malik, J. Puzicha // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2002. Vol. 24(4). – P. 509-522.
13. **Gabor, D.** Theory of communication / D. Gabor // Journal I.E.E. – 1946. – Vol. 3(93). – P. 429-457.
14. **Mallat, S.** A Wavelet Tour of Signal Processing: The Sparse Way. Third Edition / S. Mallat. – Burlington, MA: Academic Press, 2009.
15. **Deans, S.R.** The Radon Transform and Some of Its Applications / S.R. Deans. – New York: John Wiley & Sons, 1983.
16. **Duda, R.O.** Use of the Hough Transformation to Detect Lines and Curves in Pictures/ R.O. Duda, P.E. Hart // Comm. ACM. – 1972. - Vol. 15. – P. 11–15.
17. **Hu, M.K.** Visual Pattern Recognition by Moment Invariants / M.K. Hu // IRE Trans. Info. Theory. – 1962. – Vol. IT-8. – P. 179–187.
18. **Gool, L.** Affine/photometric invariants for planar intensity patterns / L. Gool, T. Moons, D. Ungureanu // Proceedings of the 4th European Conference on Computer Vision, Cambridge, UK. – 1996. – P.642-651.
19. **Glumov, N.I.** Polynomial bases for image processing in a sliding window / N.I. Glumov, V.V. Myasnikov, V.V. Sergeev // Pattern Recognition and Image Analysis. – 1994. - Vol. 4(4). - P. 408-413. – ISSN: 1054-6618.
20. **Flusser, J.** Moments and Moment Invariants in Pattern Recognition / J. Flusser, T. Suk, B. Zitová. – Wiley & Sons Ltd, 2009. – 312 p.
21. **Koenderink, J.** Representation of local geometry in the visual system / J. Koenderink, A. Doorn // Biological Cybernetics, – 1987. – Vol. 55. – P. 367-375.
22. **Florack, L.** General intensity transformations and second order invariants / L. Florack, B. Romeny, J. Koenderink, M. Viergever // Proceedings of the 7th Scandinavian Conference on Image Analysis, Aalborg, Denmark. – 1991. – P. 338-345.
23. **Freeman, W.** The design and use of steerable filters // W. Freeman, E. Adelson // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1991. – Vol.13(9). – P. 891-906.
24. **Baumberg, A.** Reliable feature matching across widely separated views / A. Baumberg // Proceedings of the Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, USA. – 2000. – P. 774-781.
25. **Schaffalitzky, F.** Multi-view matching for unordered image sets / F. Schaffalitzky, A. Zisserman // Proceedings of the 7th European Conference on Computer Vision, Copenhagen, Denmark. – 2002. – P. 414- 431.
26. **Храмов, А.Г.** Метод поля направлений в анализе и интерпретации диагностических изображений: дис. на соискание ученой степени доктора технических наук по специальности 05.13.17 - Теоретические основы информатики / А.Г. Храмов. – Самара, 2006. – 230 с.
27. **Dalal, N.** Finding People in Images and Videos. PhD Thesis, Institut National Polytechnique de Grenoble / N. Dalal. - Grenoble, 2006. – 135 p.
28. **Fukunaga, K.** Introduction to Statistical Pattern Recognition / K. Fukunaga. – 2nd ed. – New York: Academic Press, Inc, 1991.– 591 p.
29. **Schapire, R.** Boosting: Foundations and Algorithms / R. Schapire, Y. Freund – MIT Press, 2012. – 496 p.
30. **Viola, P.** Rapid Object Detection using a Boosted Cascade of Simple Features / P. Viola, M.J. Jones // Proceedings IEEE Conf. on Computer Vision and Pattern Recognition. – 2001. – P. 511-518.
31. **Cramér, H.** Mathematical Methods of Statistics / H. Cramér. – Princeton: Princeton University Press, 1999. – 575 p.
32. **Dempster, A.P.** Maximum Likelihood from Incomplete Data via the EM Algorithm / A.P. Dempster, N.M. Laird, D.B. Rubin // Journal of the Royal Statistical Society. Series B (Methodological). – 1977. – Vol. 39(1). – P. 1–38.
33. **Мясников, В.В.** Метод обнаружения транспортных средств на цифровых аэрофото- и космических изображениях дистанционного зондирования Земли / В.В. Мясников // Компьютерная оптика. – 2012. – ISSN 0134-2452. (в печати).
34. **Журавлев, Ю.И.** Алгоритмы распознавания, основанные на вычислении оценок / Ю.И. Журавлев, В.В. Никифоров // Кибернетика. — 1971. — №3.
35. **Zhuravlev, J.I.** An algebraic approach to recognition or classifications problems / J.I. Zhuravlev // Pattern Recognition and Image Analysis. — 1998. — Vol. 8(1) — P. 59–100. – ISSN: 1054-6618.

### References

1. **Schmid, C.** Evaluation of interest point detectors / C. Schmid, R. Mohr, C. Bauckhage // Int. Journal of Computer Vision. – 2000. – Vol. 37(2). – P. 151-172.
2. **Mikolajczyk, K.** A performance evaluation of local descriptors / K. Mikolajczyk, C. Schmid // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2005. – Vol. 10(27). – P. 1615-1630.
3. **Dalal, N.** Histograms of Oriented Gradients for Human Detection / N. Dalal, B. Triggs // Proceedings of IEEE Conference Computer Vision and Pattern Recognition, San Diego, USA. – 2005. – P. 886-893.
4. **Forsyth, D.A.** Computer Vision: A Modern Approach / D.A. Forsyth, J. Ponce. – New Jersey: Prentice Hall, 2002. – 693 p.
5. **Shapiro, L.G.** Computer Vision / L.G. Shapiro, G.C. Stockman. – New Jersey: Prentice Hall, 2001. – 608 p.
6. **Methods of computer image processing.** / M.V. Gashnikov, N.I. Glumov, N.U. Ilyasova, V.V. Myasnikov [et al]. - 2-nd edition, reviewed. - Moscow.: "Fizmatlit Publisher", 2003. – 784 p. – (In Russian).
7. **Johnson, A.** Object recognition by matching oriented points / A. Johnson, M. Hebert // Proceedings of the Conference on Computer Vision and Pattern Recognition, Puerto Rico, USA. – 1997. – P. 684-689.
8. **Zabih, R.** Non-parametric local transforms for computing visual correspondence / R. Zabih, J. Woodfill // Proceedings of the 3rd European Conference on Computer Vision, Stockholm, Sweden. – 1994. – P. 151-158.
9. **Lowe, D.** Distinctive image features from scale-invariant keypoints / D. Lowe // International Journal of Computer Vision. – 2004. - Vol. 2(60). – 91-110.
10. **Sarfraz, O.** Head Pose Estimation in Face Recognition across Pose Scenarios / O. Sarfraz, S. Hellwich // Proceedings of VISAPP 2008, Int. conference on Computer Vision Theory and Applications, Madeira, Portugal. – 2008. - P. 235-242.
11. **Ashbrook, A.** Robust recognition of scaled shapes using pairwise geometric histograms / A. Ashbrook, N. Thacker, P. Rockett, C. Brown // Proceedings of the sixth British Machine Vision Conference, Birmingham, UK. – 1995 – P. 503-512.
12. **Belongie, S.** Shape matching and object recognition using shape contexts / S. Belongie, J. Malik, J. Puzicha // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2002. – Vol. 24(4). – P. 509-522.
13. **Gabor, D.** Theory of communication / D. Gabor // Journal I.E.E. – 1946. – Vol. 3(93). – P. 429-457.



14. **Mallat, S.** A Wavelet Tour of Signal Processing: The Sparse Way / S. Mallat. – Burlington, MA: Academic Press, 2009.
15. **Deans, S.R.** The Radon Transform and Some of Its Applications / S.R. Deans. – New York: John Wiley & Sons, 1983.
16. **Duda, R.O.** Use of the Hough Transformation to Detect Lines and Curves in Pictures/ R.O. Duda, P.E. Hart // Comm. ACM. – 1972. – Vol. 15. – P. 11–15.
17. **Hu, M.K.** Visual Pattern Recognition by Moment Invariants / M.K. Hu // IRE Trans. Info. Theory. – 1962. – Vol. IT-8. – P. 179–187.
18. **Gool, L.** Affine/photometric invariants for planar intensity patterns / L. Gool, T. Moons, D. Ungureanu // Proceedings of the 4th European Conference on Computer Vision, Cambridge, UK. – 1996. – P.642-651.
19. **Glumov, N.I.** Polynomial bases for image processing in a sliding window / N.I. Glumov, V.V. Myasnikov, V.V. Sergeyev // Pattern Recognition and Image Analysis. – 1994. - Vol. 4(4). - P. 408-413. – ISSN: 1054-6618.
20. **Flusser, J.** Moments and Moment Invariants in Pattern Recognition / J. Flusser, T. Suk, B. Zitová. – Wiley & Sons Ltd, 2009. – 312 p.
21. **Koenderink, J.** Representation of local geometry in the visual system / J. Koenderink, A. Doorn // Biological Cybernetics. – 1987. – Vol. 55. – P. 367-375.
22. **Florack, L.** General intensity transformations and second order invariants / L. Florack, B. Romeny, J. Koenderink, M. Viergever // Proceedings of the 7th Scandinavian Conference on Image Analysis, Aalborg, Denmark. – 1991. – P. 338-345.
23. **Freeman, W.** The design and use of steerable filters // W. Freeman, E. Adelson / IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1991. – Vol.13(9). – P. 891-906.
24. **Baumberg, A.** Reliable feature matching across widely separated views / A. Baumberg // Proceedings of the Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, USA. – 2000. – P. 774-781.
25. **Schaffalitzky, F.** Multi-view matching for unordered image sets / F. Schaffalitzky, A. Zisserman // Proceedings of the 7th European Conference on Computer Vision, Copenhagen, Denmark. – 2002. – P. 414- 431.
26. **Khramov, A.G.** Method of direction field for analysis and interpretation of diagnostic images. Doctoral thesis, speciality 05.13.17 – Informatics theoretical basis / A.G. Khramov. – Samara, 2006. – 230 p. – (In Russian).
27. **Dalal, N.** Finding People in Images and Videos. PhD Thesis, Institut National Polytechnique de Grenoble / N. Dalal. – Grenoble, 2006. – 135 p.
28. **Fukunaga, K.** Introduction to Statistical Pattern Recognition / K. Fukunaga. – 2nd ed., – New York: Academic Press, Inc, 1991. – 591 p.
29. **Schapire, R.** Boosting: Foundations and Algorithms / R. Schapire, Y. Freund. – MIT Press, 2012. – 496 p.
30. **Viola, P.** Rapid Object Detection using a Boosted Cascade of Simple Features / P. Viola, M.J. Jones // Proceedings IEEE Conf. on Computer Vision and Pattern Recognition, – 2001. – P. 511-518.
31. **Cramér, H.** Mathematical Methods of Statistics / H. Cramér. – Princeton: Princeton University Press, 1999. – 575 p.
32. **Dempster, A.P.** Maximum Likelihood from Incomplete Data via the EM Algorithm / A.P. Dempster, N.M. Laird, D.B. Rubin // Journal of the Royal Statistical Society. Series B (Methodological). – 1977. – Vol. 39(1). – P. 1–38.
33. **Myasnikov, V.V.** Methods of vehicle detection on the aero and space images of Earth remote sensing / V.V. Myasnikov // Computer Optics. – 2012. – ISSN 0134-2452. – (in printing, in Russian).
34. **Zhuravlev, J.I.** Pattern recognition algorithms, based on valuation calculation / J.I. Zhuravlev, V.V. Nikiforov // Cybernetics. — 1971. — №3. – (In Russian)/
35. **Zhuravlev, J.I.** An algebraic approach to recognition or classifications problems / J.I. Zhuravlev // Pattern Recognition and Image Analysis. — 1998. — Vol. 8(1) — P. 59–100. – ISSN: 1054-6618.

## MODEL-BASED GRADIENT FIELD DESCRIPTOR AS A CONVENIENT TOOL FOR IMAGE RECOGNITION AND ANALYSIS

*V.V. Myasnikov*

*Image Processing Systems Institute of the RAS*

### Abstract

In the paper we propose a new descriptor which is used to describe a digital image, specifically, a model-oriented descriptor of gradient field. Derived characteristics of the descriptor, that are considered as features of a digital image, allow to solve effectively the problems of image analysis, recognition and retrieval. The examples of such tasks solutions with the proposed descriptor are given.

Key words: digital images, descriptors, features, analysis, recognition, image retrieval.

*Сведения об авторе Мясников Владислав Валерьевич см. стр. 584 этого номера.*

*Поступила в редакцию 27 сентября 2012 г.*