ИСПОЛЬЗОВАНИЕ ПРОЕКТИВНЫХ МОРФОЛОГИЙ В ЗАДАЧАХ ОБНАРУЖЕНИЯ И ИДЕНТИФИКАЦИИ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ[1]

Ю.В. Визильтер, С.Ю.Желтов

Москва, ФГУП ГосНИИАС

Поступила в редакцию 20.10.07 г.

Описан подход к разработке алгоритмов обнаружения и идентификации объектов на изображениях, опирающийся на формализм проективной морфологии. Рассмотрены методы конструирования алгоритмов проективного морфологического анализа изображения. Рассмотрена задача автоматического построения модели объекта по заданной обучающей выборке. Для вероятностного описания алгоритмов обнаружения сложных структурных объектов предложена модификация метода обобщенного анализа свидетельств, основанная на моделях и процедурах проективной морфологии.

Введение. В статье [1] была предложена проективная морфология, представляющая собой алгебру структур с однородной связью, позволяющая формально описывать и сравнивать структурные модели и структурные объекты вне зависимости от конкретной прикладной области. Базовым понятием здесь является «составной образ», который строится из первичных «примитивов» или других образов при помощи их объединения некоторой кумулятивной (ассоциативной и коммутативной) операцией с некоторыми весами-скалярами, численно характеризующими «степень связи» образа с его элементом. «Моделью» образа считается его структурное описание со свободными (переменными) связями (весами). Важным элементом данного формализма является наличие оператора проекции образа на образ (а также образа на модель другого образа), который формально описывает семантически нагруженные понятия «сходства» или «близости» образов. Специфическим свойством проективной морфологии является то, что вес каждой образующей в модели образа определяется проекцией этого образа на данную образующую. Показано, что введенные таким образом однородные структуры всегда могут быть однозначно охарактеризованы массивами или векторами признаков, которые в семантическом плане можно рассматривать как содержательные структурные описания соответствующих объектов. При этом в пространстве морфологических разложений стандартным образом определяются удобные нормированные коэффициенты морофологической корреляции, позволяющие оценивать степень сходства образов с точностью до некоторых классов эквивалентности, которые, в свою очередь, соответствуют определенным семантическим таксонам предметной области («фигуры одной формы», «модели одного класса» и т.п.).

На основе проективной морфологии изображений предложен обобщенный формализм морфологического сравнения изображений, определены меры структурного сходства между ними, описано решение задач взаимной структурной привязки

[1] *Работа выполнена при поддержке РФФИ, проект 05-08-18234-а.*

изображений (matching). Показано, что предложенный формализм позволяет рассматривать в рамках единого подхода не только корреляцию функций яркости и процедуры морфологического анализа Ю.П. Пытьева [2,3], но также структурные модели изображения, опирающиеся на операторы математической морфологии Серра [4,5], методы голосования, восходящие к преобразованию Хафа [6-9], частотные методы фильтрации (на базе преобразования Фурье, БПФ, ДКП и др.), пространственно-частотные методы фильтрации (кратномасштабный вейвлет-анализ) [10].

В статье [1] была также описана структурная фильтрация цифровых изображений с использованием проективных морфологий.

В данной статье с точки зрения проективной морфологии будут рассмотрены основные проблемы, связанные с конструированием алгоритмов обнаружения и идентификации объектов на изображениях.

1. Модели с однородными связями и морфологические фильтры. Рассмотрим представления значений изображения f в точке (x,y), используемые в ряде популярных методов анализа двумерных данных.

Цифровое изображение (тривиальное пиксельное описание):

$$Im(f,x,y)= \Sigma_{ij}(a_{ij}*\varphi(i,j,x,y)) = MAX_{xy} (a_{ij}*\varphi(i,j,x,y)),$$

где $\varphi(i,j,x,y)$ – индикаторная функция отдельного пикселя вида

$$\varphi(i,j,x,y)= \{1: x=i,y=j;\ 0: x\neq i,y\neq j\};$$

(x,y) – положение пикселя; a_{ij} – значение цифрового изображения в точке (i,j). Данное формальное описание (разбиение) изображения соответствует корреляционному сравнению изображений как функций яркости.

Дискретное косинусное преобразование Фурье:

$$FUR(f,x,y))=\Sigma_{ij}(a_{ij}*\varphi(\omega_i,\omega_j,x,y)),$$

где ω_i,ω_j.- пространственные частоты, $\varphi(\omega_i,\omega_j,x,y)= \cos(\omega_i,ix)*\cos(\omega_j,iy)$; a_{ij} – коэффициент разложения Фурье.

Кратномасштабное вейвлет-преобразование:

$$W(f,x,y))=\Sigma_{ij}(a_{ijRn}*\varphi(i,j,R,n,x,y)),$$

где $\varphi(0,0,1,n)$ – эталонный вейвлет n-го типа, (i,j) – положение вейвлета, R – коэффициент масштаба; a_{ijRn} – коэффициент вейвлет-преобразования.

Морфологическое открытие:

$$O_M(f,x,y))=MAX_{ij}(a_{ij}*\varphi(i,j,x,y)),$$

где $\varphi(0,0)=\{0,1\}$ – эталонный *бинарный* структурирующий элемент, (i,j) – положение структурирующего элемента; $\{a_{ij}\}$– результат *эрозии* изображения $f(x,y)$.

Кратномасштабное морфологическое открытие:

$$O_{MM}(f,x,y))=MAX_{ij}(a_{ijRn}*\varphi(i,j,R,n,x,y)),$$

где $\varphi(0,0,1,n)=\{0,1\}$ – эталонный *бинарный* структурирующий элемент n-го типа, (i,j) – положение структурирующего элемента, R – коэффициент масштаба; a_{ijRn} – результат кратномасштабной эрозии изображения $f(x,y)$.

Морфология на базе преобразования Хафа (HT):

$$O_{HT}(f,x,y))= MAX_{ij}(a_{ij}*\varphi(\rho_i,\theta_j,x,y)),$$

где (ρ_i,θ_j) – параметры нормальной параметризации прямой, $\varphi(\rho_i,\theta_j)$ – прямая линия с соответствующими параметрами; $\{a_{ij}\}$ – содержимое бинаризованного аккумулятора преобразования Хафа.

Морфология на базе преобразования Хафа в скользящем окне (Local HT):

$$O_{LHT}(f,x,y))=MAX_{ij\theta}(a_{ij\theta}*\varphi(i,j,\theta,x,y)),$$

где $\varphi(0,0,0)$ – эталонный *структурирующий элемент в виде прямолинейного отрезка фиксированного размера*, (i,j) – положение центра структурирующего элемента, θ – угол поворота отрезка; $\{a_{ij\theta}\}$ – содержимое бинаризованного аккумулятора преобразования Хафа в скользящем окне.

Морфологическое сравнение изображений Ю.П.Пытьева:

$$Pr_{\mathbf{F}}(f,x,y)=\Sigma_n(a_n*\varphi(n,x,y)),$$

где $\varphi(n,x,y)$ – индикаторная функция n-й области разбиения кадра $F_n\subseteq\mathbf{F}=\{F_i\}$ на участки однородной яркости вида:

$$\varphi(n,x,y)= \{1: (x,y)\in F_n;\ 0: (x,y)\notin F_n\};$$

a_n – средняя яркость изображения $f(x,y)$по области F_n.

Таким образом, во всех случаях просматривается следующая единая схема *структурного представления изображения*:

$$A(\mathbf{p})=\vee_{q\in Q}(A(\mathbf{q})\bullet\varphi(\mathbf{p},\mathbf{q})), \qquad (1.1)$$

где $\mathbf{p}=(x,y)$ – вектор пиксельных координат в исходном пространстве изображения; $A(\mathbf{p})=Im(x,y)$ – анализируемое изображение, заданное как двумерная скалярная функция интенсивности сигнала (яркости); $\varphi(\mathbf{p},\mathbf{q})=\varphi(x,y,\mathbf{q})$ – набор образующих (примитивов) данного структурного разложения, также заданных как параметризованные двумерные функции яркости; $\mathbf{q}$ – вектор параметров элемента разложения; $A(\mathbf{q})$ – образ изображения в пространстве параметров (*массив-аккумулятор*); ‘$\vee$’$\in${‘Σ’,‘MAX’,‘П’,‘MIN’} – коммутативная и ассоциативная операция *попиксельного* (поэлементного) объединения элементов изображения, которая может иметь смысл либо сложения/умножения элементов, либо взятия максимального/минимального значения.

Модели типа (1.1) можно назвать *моделями с однородными связями* или просто *однородными моделями* так как все образующие элементы входят в состав модели одним и тем же способом, объединяются при помощи некоторой кумулятивной (коммутативной и ассоциативной) операции, а модель описывает лишь структурный состав образа, но не отношения между его элементами.

Прежде, чем перейти непосредственно к задаче локализации объектов, введем еще несколько определений в соответствии с работой [1].

Назовем *морфологическим преобразованием* модели (1.1) любое преобразование ψ, такое что:

$$\psi(\mathrm{A}(\mathbf{p}))=\bigvee_{q\in Q}(\psi(\mathbf{q})\bullet \mathrm{A}(\mathbf{q})\bullet\varphi(\mathbf{p},\mathbf{q})), \quad (1.2)$$

где $\psi(\mathbf{q})\in\Psi$ – *весовая функция* данного преобразования в пространстве параметров.

Если оператор ψ обладает свойством алгебраического проектора:

$$\psi(\psi(\mathrm{A}(\mathbf{p})))=\psi(\mathrm{A}(\mathbf{p})), \quad (1.3)$$

то такое преобразование можно назвать *морфологическим фильтром*.

Учитывая свойства выражений (1.1)-(1.3), любой морфологический фильтр, весовая функция которого не зависит от анализируемого изображения, можно описать следующим образом:

$$f(\mathrm{A}(\mathbf{p}))=\bigvee_{q\in Q}(f(\mathbf{q})\bullet \mathrm{A}(\mathbf{q})\bullet\varphi(\mathbf{p},\mathbf{q})), \quad (1.4)$$

где $f(\mathbf{q})=\{0,1\}$ – индикаторная функция *области пропускания* данного фильтра.

С точки зрения проективной морфологии ([1]) любой морфологический фильтр может быть также рассмотрен как *оператор проекции исходного образа на область пропускания*, рассматриваемую как подпространство, образованное линейным замыканием множества структурных примитивов $\mathbf{F}(\mathbf{p},\mathbf{q})=\{\varphi(\mathbf{p},\mathbf{q})\colon f(\mathbf{q})\neq 0\}$.

Таким образом, возникает унифицированная двухэтапная схема фильтрации изображений:

- *Деконструкция* (разложение, анализ) – проектирование изображения на образующие элементы преобразования;
- *Частичная реконструкция* (синтез) – объединение проекций на те элементы, которые находятся в области пропускания фильтра.

Совместно эти два этапа всегда определяют операцию морфологической фильтрации, обладающую свойствами алгебраического проектора (1.3).

2. Конструирование алгоритмов обнаружения объектов на изображениях. Пусть имеется *морфо-геометрическая модель объекта* вида:

$$\mathrm{M}(\mathbf{p},\mathbf{u})=\bigvee_{q\in Q}(\mathrm{M}(\mathbf{u},\mathbf{q})\bullet\varphi(\mathbf{p},\mathbf{q})) , \qquad (2.1)$$

где $\mathbf{u}\in Q$ – вектор параметров локализации объекта $\mathrm{M}(\mathbf{p},\mathbf{u})$; $\mathbf{q}\in Q$ – вектор параметров локализации структурных примитивов данного разложения $\varphi(\mathbf{p},\mathbf{q})$; $\mathrm{M}(\mathbf{u},\mathbf{q})=\{0,1\}$ –*модель локализации объекта*, которая описывает все допустимые соответствия между значениями параметров локализации образа в целом и параметрами локализации составляющих его геометрических примитивов.

Рассмотрим задачу обнаружения всех объектов, описываемых моделью (2.1), на наблюдаемом изображении, описываемом моделью (1.1).

$$\mathrm{A}(\mathbf{p})=\bigvee_{q\in Q}(\mathrm{A}(\mathbf{q})\bullet\varphi(\mathbf{p},\mathbf{q})). \qquad (2.2)$$

Следуя принятой в работе [1] методологии, определим проекцию изображения (2.2) на модель (80):

$$\Pr(\mathrm{A}(\mathbf{p}),\mathrm{M}(\mathbf{p},\mathbf{u})) = \bigvee_{q\in Q}(\mathrm{M}(\mathbf{u},\mathbf{q})\bullet\mathrm{A}(\mathbf{q})\bullet\varphi(\mathbf{p},\mathbf{q})) \qquad (2.3)$$

и соответствующий коэффициент структурной корреляции

$$\mathrm{K}(\mathrm{A}(\mathbf{p}),\mathrm{M}(\mathbf{p},\mathbf{u})) = |\Pr(\mathrm{A}(\mathbf{p}),\mathrm{M}(\mathbf{p},\mathbf{u}))| / |\mathrm{A}(\mathbf{q})|. \qquad (2.4)$$

Как видно, коэффициент корреляции оказывается функцией параметра **u**, то есть представляет собой *корреляционное поле*

$$\mathrm{K}(\mathbf{u}) = \sum_{\mathbf{q}\in Q}|\mathrm{M}(\mathbf{u},\mathbf{q})\bullet\mathrm{A}(\mathbf{q})| / \sum_{q\in Q}|\mathrm{A}(\mathbf{q})|, \qquad (2.5)$$

локальные максимумы которого соответствуют параметрам наиболее достоверной локализации объекта, описываемого моделью (2.1).

Поскольку геометрическая модель $\mathrm{M}(\mathbf{u},\mathbf{q})$ симметрична по отношению к векторам параметров **u** и **q**, возникают две равновозможные стратегии анализа изображения с целью обнаружения всех объектов, описываемых моделью (2.1).

<u>Анализ изображения «сверху вниз»</u> осуществляется путем принудительного вычисления значений корреляционного поля $\mathrm{K}(\mathbf{u})$ для всех возможных значений параметров вектора **u**. Данный способ сплошной проверки всех возможных вариантов локализации объекта можно также назвать процедурой *согласованной морфологической фильтрации*, поскольку операция проецирования (2.3) описывает структурный фильтр, удовлетворяющий условиям (1.2)-(1.3).

<u>Анализ изображения «снизу вверх»</u> осуществляется путем обнаружения значимых особенностей изображения и их голосования в пользу возможных значений параметров вектора **u**, определяемых выражением $\mathrm{M}(\mathbf{u},\mathbf{q})=1$.

Пусть *морфологическим событием* (event) $e(\mathbf{q})$ считается факт обнаружения в процессе анализа изображения $\mathrm{A}(\mathbf{p})$ такого значения **q**, что $\Pr(\mathrm{A}(\mathrm{p}),\varphi(\mathbf{p},\mathbf{q}))>\varnothing$, т.е. $\mathrm{A}(\mathbf{q})>0$.

Назовем *пространственной функцией отклика* события $e(\mathbf{q})$ выражение

$$\Phi(\mathbf{p},\mathbf{q}) = \bigvee_{\mathbf{u}\in Q} M(\mathbf{u},\mathbf{q})\bullet\varphi(\mathbf{p},\mathbf{q}), \tag{2.6}$$

и *корреляционной функцией отклика* события $e(\mathbf{q})$ выражение

$$K(\mathbf{u},\mathbf{q}) = |M(\mathbf{u},\mathbf{q})| / \sum_{q\in Q}|A(\mathbf{q})|. \tag{2.7}$$

С учетом этих обозначений выражения (82) и (2.2) соответственно принимают вид:

$$Pr(A(\mathbf{p}),M(\mathbf{p},\mathbf{u})) = \bigvee_{q\in Q} \Phi(\mathbf{p},\mathbf{q})\bullet A(\mathbf{q}), \tag{2.8}$$

$$K(\mathbf{u}) = \sum_{\mathbf{q}\in Q} K(\mathbf{u},\mathbf{q})\bullet|A(\mathbf{q})|. \tag{2.9}$$

Выражения (2.8)-(2.9) позволяют организовать процесс вычисления проекции (2.3) и корреляционного поля (2.5) как процесс объединения и суммирования функций отклика обнаруженных структурных событий с соответствующими весами. При этом вычисления производятся не для всех структурных примитивов, а только для тех, в которых «зарегистрированы» соответствующие «события». Такой способ управляемого данными анализа изображений будем называть процедурой *голосования морфологических событий*.

Во многих случаях количество морфологических событий, определенных выше как факт обнаружения непустых проекций на структурные элементы базиса разложения, оказывается слишком велико, и тогда целесообразным оказывается определить некоторый порог t, выделяющий среди всех событий множество *значимых событий*:

$$\boldsymbol{e}_t(q) = \{e_t(\mathbf{q})=\{B(\mathbf{q})>t\}\}. \tag{2.10}$$

Пусть π_t представляет собой операцию порогового отсечения с порогом $t\in\Psi$:

$$\pi_t(a) = \{a, \text{ если } |a|>t;\ 0 - \text{в противном случае}\}. \tag{2.11}$$

Тогда можно определить следующие модифицированные функции:

$$Pr_t(A(\mathbf{p}),M(\mathbf{p},\mathbf{u})) = \bigvee_{q\in Q} \Phi(\mathbf{p},\mathbf{q})\bullet\pi_t(A(\mathbf{q})), \tag{2.12}$$

$$K_t(\mathbf{u}) = \sum_{\mathbf{q}\in Q} K(\mathbf{u},\mathbf{q})\bullet|\pi_t(A(\mathbf{q}))|. \tag{2.13}$$

В вычислительном плане работа только со значимыми событиями дает существенное преимущество при построении алгоритмов голосования событий. В плане достоверности обнаружения объекта (привязки фрагмента) связь между полным решением (2.7) и выборочным решением (2.13) определяется следующим утверждением.

Утверждение 1. Для любого фиксированного порога K_{min} *правила принятия решения* об обнаружении объекта (позиционировании фрагмента изображения) по полной корреляционной статистике $K(\mathbf{u})>K_{min}$ условие обнаружения по выборочной статистике $K_t(\mathbf{u})>K_{min}$ является *достаточным*.

Доказательство. Доказательство основано на свойствах оператора $\pi_t(a)$:

$$\forall\mathbf{q}: |A(\mathbf{q})| \ge |\pi_t(A(\mathbf{q}))| \Rightarrow \sum_{\mathbf{q}\in Q} K(\mathbf{u},\mathbf{q})\bullet|A(\mathbf{q})| \ge \sum_{\mathbf{q}\in Q} K(\mathbf{u},\mathbf{q})\bullet|\pi_t(A(\mathbf{q}))| \Rightarrow$$

$$\Rightarrow K(\mathbf{u}) \ge K_t(\mathbf{u}),$$

то есть

$K_t(\mathbf{u})>K_{min} \Rightarrow K(\mathbf{u})>K_{min}$,

ч.т.д.♦

Как было ранее отмечено в работе [11], в отличие от «переборных» корреляционных методов обнаружения, методы, основанные на голосовании элементов изображения в пользу гипотез о присутствии и положении на изображении искомого объекта, обладают следующими возможностями *увеличения вычислительной эффективности* процедур анализа изображения:

- Независимое (в смысле очередности подачи голосов) голосование образующих элементов в пользу параметров локализации объекта;
- Возможность декомпозиции или редукции модели объекта (группировка или уменьшение количества анализируемых образующих в модели);
- Возможность декомпозиции или редукции вектора параметров (группировка или уменьшение количества анализируемых параметров модели).

Рассмотрим эти методы применительно к введенным общим определениям.

3. Аккумулирование свидетельств. Возможность *независимого аккумулирования голосов* элементов при вычислении выражений типа (2.6) или (2.10) обеспечивается тем, что используемые для построения морфологических проекций процедуры $\mathbf{V}=\{\Sigma,\Pi,MAX,MIN\}$ являются операциями *кумулятивного накопления* в том смысле, что

$A\vee B=B\vee A$;

$(A\vee B)\vee C=A\vee(B\vee C)$,

то есть при последовательном поэлементном применении этих операций к некоторому произвольному множеству элементов, результат итогового накопленного значения зависит не от порядка опроса элементов множества, а от всей совокупности содержащихся в множестве элементов информации**.**

Таким образом, в процессе заполнения аккумулятора процедуры обнаружения каждый структурирующий элемент может делать свой вклад в результат независимо от всех остальных. Это позволяет реализовать для всех структурных разложений единообразный процесс обнаружения объектов «снизу вверх» методом голосования обнаруженных морфологических особенностей изображения, описанный выше.

В случае реализации в параллельных (многопроцессорных) вычислительных средах независимое голосование образующих элементов позволяет добиться *максимально возможной степени распараллеливания вычислений.*

В случае оптимизации алгоритма под последовательную (однопроцессорную) вычислительную архитектуру та же самая независимость голосования позволяет ставить интересные задачи построения *оптимальных в вычислительном смысле процедур обнаружения при фиксированном уровне надежности детектирования*. При этом участвующие в голосовании события ранжируются по критерию, пропорциональному значимости и обратно пропорциональному вычислительным затратам на детектирование данного события. Это определяет порядок анализа изображения, при котором гипотезы о наличии и положении объекта предварительно генерируются, либо отбрасываются с минимальными затратами вычислительных ресурсов, а затем, если необходимо, поэтапно уточняются по той же схеме.

Однако методы голосования вычислительно эффективны лишь в том случае, когда удается использовать пространство параметров невысокой размерности. Т.н. "проклятие размерности" для подобных алгоритмов заключается в том, что с ростом размерности пространства параметров резко возрастают затраты памяти, и одновременно падает быстродействие. Поэтому, для практического использования кумулятивных методик, вопрос выбора пространства параметров является одним из основных. Имеет смысл рассмотреть два основных подхода к снижению размерности пространства параметров. Первый из них основан на *параметрической редукции*, второй - на *структурном загрублении* модели объекта.

4. Декомпозиция и редукция вектора параметров. Пусть вектор параметров локализации объекта $\mathbf{u}\in Q$ в морфо-геометрической модели (2.1) состоит из двух частей:

$$\mathbf{u}=(\mathbf{u}_1,\mathbf{u}_2),\ M(\mathbf{u},\mathbf{q})=M(\mathbf{u}_1,\mathbf{u}_2,\mathbf{q})\in\{0,1\}. \qquad (4.1)$$

Такое разбиение вектора (и пространства) параметров назовем *декомпозицией*.

Попробуем теперь исключить из рассмотрения подвектор $\mathbf{u}_2$. Для этого введем *параметрически редуцированную модель локализации*

$$M(\mathbf{u}_1,\mathbf{q}) = \mathrm{MAX}_{u_2\in Q}\ M(\mathbf{u}_1,\mathbf{u}_2,\mathbf{q}). \qquad (4.2)$$

Модель (4.2) будем называть параметрически редуцированной, так как она опирается на редуцированный (усеченный) вектор параметров локализации. Смысл ее заключается в том, что если в исходной модели $M(\mathbf{u}_1,\mathbf{u}_2,\mathbf{q})$ для данного значения $\mathbf{u}_1$ существует хотя бы одно значение $\mathbf{u}_2$, такое что $M(\mathbf{u}_1,\mathbf{u}_2,\mathbf{q})=1$, то $M(\mathbf{u}_1,\mathbf{q})=1$. В противном случае $M(\mathbf{u}_1,\mathbf{q})=0$.

С учетом (4.2) параметрически редуцированная модель объекта примет следующий вид:

$$M(\mathbf{p},\mathbf{u}_1)=\bigvee_{q\in Q}(M(\mathbf{u}_1,\mathbf{q})\bullet\varphi(\mathbf{p},\mathbf{q})), \qquad (4.3)$$

а выражение для корреляционного поля (2.3) соответственно

$$K(\mathbf{u}_1) = \sum_{\mathbf{q}\in Q}|M(\mathbf{u}_1,\mathbf{q})\bullet A(\mathbf{q})| \,/\, \sum_{q\in Q}|A(\mathbf{q})|. \qquad (4.4)$$

Связь между результатами обнаружения объекта по полной модели (2.4) и параметрически редуцированной модели (4.3) определяется следующим утверждением.

Утверждение 2. Для любого фиксированного порога K_{min} правила принятия решения об обнаружении объекта (позиционировании фрагмента изображения) по полной корреляционной статистике $K(\mathbf{u})>K_{min}$ условие обнаружения по выборочной статистике $K(\mathbf{u}_1) >K_{min}$ является *необходимым*.

Доказательство. Заметим прежде всего, что из (4.2) следует

$$\forall \mathbf{u}_1,\mathbf{u}_2\text{: } M(\mathbf{u}_1,\mathbf{q}) \ge M(\mathbf{u}_1,\mathbf{u}_2,\mathbf{q}),$$

кроме того,

$$M(\mathbf{u},\mathbf{q}),M(\mathbf{u}_1,\mathbf{q})\in\{0,1\} \Rightarrow M(\mathbf{u},\mathbf{q})=|M(\mathbf{u},\mathbf{q})|;\ M(\mathbf{u}_1,\mathbf{q})= |M(\mathbf{u}_1,\mathbf{q})|.$$

Отсюда

$$K(\mathbf{u}_1) - K(\mathbf{u}) = \sum_{\mathbf{q}\in Q} ((M(\mathbf{u}_1,\mathbf{q})-M(\mathbf{u}_1,\mathbf{u}_2,\mathbf{q}))\bullet|A(\mathbf{q})|) \,/\, \sum_{q\in Q}|A(\mathbf{q})| \ge 0 \Rightarrow K(\mathbf{u}_1) \ge K(\mathbf{u}),$$

то есть

$$K(\mathbf{u})>K_{min} \Rightarrow K(\mathbf{u}_1)>K_{min},$$

ч.т.д.♦

Таким образом, обнаружение объектов по параметрически редуцированной модели гарантирует, что ни один из возможных кандидатов не будет пропущен, однако не гарантирует, что все обнаруженные кандидаты соответствуют исходной полной модели объекта. Это значит, что мы можем сначала оценить вектор параметров $\mathbf{u}_1$, а затем, используя эту оценку - вектор оставшихся параметров $\mathbf{u}_2$. С точки зрения экономии вычислений выгода от такого разделения очевидна, так как вместо работы в пространстве параметров размерности $\dim(\mathbf{u})$ мы сначала работаем в пространстве размерности $\dim(\mathbf{u}_1)$, а затем - в пространстве размерности $\dim(\mathbf{u}_2)$. Такая процедура обнаружения основана *декомпозиции* вектора параметров. Если же достаточно определить только подвектор параметров размерности $\mathbf{u}_1$, не уточняя оставшиеся значения параметров, то формируется процедура обнаружения, основанная на *редукции* вектора параметров.

Именно декомпозиция и редукция вектора параметров регистрации и позволяют строить эффективные в вычислительном смысле инвариантные алгоритмы голосования морфологических событий.

5. Структурное загрубление модели объекта. Пусть имеется морфо-геометрическая модель объекта $M(\mathbf{p},\mathbf{u})$ вида (2.1). Перейдем от этой модели к модели:

$$M'(\mathbf{p},\mathbf{u})=\bigvee_{q\in Q'}(M'(\mathbf{u},\mathbf{q})\bullet\varphi(\mathbf{p},\mathbf{q})),$$

$M'(\mathbf{u},\mathbf{q})\in\{0,1\},\ \forall\mathbf{u}\in Q:\ M'(\mathbf{u},\mathbf{q}) \geq M(\mathbf{u},\mathbf{q})$. (5.1)

Такая операция перехода от данной конкретной модели объекта к более абстрактной (или, что то же самое, от «более сложной» к «более простой») называется *структурным загрублением* модели объекта.

Покажем теперь, что обнаружение объекта по загрубленной модели M' есть необходимое условие для его обнаружения по полной модели M.

Рассмотрим корреляционное поле

$K'(\mathbf{u}) = \sum_{\mathbf{q}\in Q}|M'(\mathbf{u},\mathbf{q})\bullet A(\mathbf{q})| / \sum_{q\in Q}|A(\mathbf{q})|$. (5.2)

Утверждение 3. Для любого фиксированного порога K_{min} правила принятия решения об обнаружении объекта (позиционировании фрагмента изображения) по полной корреляционной статистике $K(\mathbf{u})>K_{min}$ условие обнаружения по загрубленной модели $K'(\mathbf{u})>K_{min}$ является *необходимым*.

Доказательство. Поскольку из (5.1) следует

$\forall\mathbf{u}$: $M'(\mathbf{u},\mathbf{q}) \geq M(\mathbf{u},\mathbf{q})$,

кроме того,

$M(\mathbf{u},\mathbf{q}),M'(\mathbf{u},\mathbf{q})\in\{0,1\} \Rightarrow M(\mathbf{u},\mathbf{q})=|M(\mathbf{u},\mathbf{q})|;\ M'(\mathbf{u},\mathbf{q})= |M'(\mathbf{u},\mathbf{q})|$.

Отсюда

$K'(\mathbf{u}) - K(\mathbf{u}) = \sum_{\mathbf{q}\in Q} ((M'(\mathbf{u},\mathbf{q})-M(\mathbf{u},\mathbf{q}))\bullet|A(\mathbf{q})|) / \sum_{q\in Q}|A(\mathbf{q})| \geq 0 \Rightarrow K'(\mathbf{u}) \geq K(\mathbf{u})$,

то есть

$K'(\mathbf{u})>K_{min} \Rightarrow K(\mathbf{u})>K_{min}$,

ч.т.д.♦

Таким образом, как и в случае параметрической редукции, работая с загрубленной моделью, мы не теряем ни одного из возможных кандидатов.

В то же время, как при использовании загрубленной модели объекта, так и в случае параметрической редукции, платой за достижение вычислительной эффективности становится потеря однозначности интерпретации результатов анализа аккумулятора. Поэтому их *постпроверка* путем повторного анализа изображения с использованием полной модели объекта становится обязательной.

6. Автоматическое конструирование алгоритмов обнаружения объектов. В предыдущих параграфах предполагалось, что модель вида (2.1)-(2.2) исходно задана, т.е. сконструирована разработчиком, исходя из каких-то его априорных знаний о предметной области. Однако в большинстве современных задач анализа изображений модель объекта требуется получить (*автоматически сконструировать*) в процессе *обучения*, опираясь изображения объектов различных классов, имеющиеся в обучающей выборке. В рамках описываемого подхода *модельное описание объекта* (*дескриптор формы изображения* объекта) представляет собой набор образующих элементов данного типа морфологического разложения, которые являются характерными (значимыми) для разложения именно данного объекта (класса объектов). Выделение таких значимых черт есть некая процедура статистического анализа множества морфологических разложений изображений из обучающей выборки. Задача *построения дескриптора* состоит в том, чтобы для данной выборки определить такой минимально возможный набор элементов разложения, чтобы он обеспечивал вероятность обнаружения объекта не ниже заданной.

В случае если выборка невелика, решение находится простым перебором всех сочетаний структурных элементов, выделенных на большинстве эталонных изображений из выборки как значимые. Для одного эталонного изображения такое «обучение» сводится к простой процедуре типа построения Look-Up-Table по эталону в методе GHT.

В случае больших объемов выборки необходимо использовать такие методы формирования «сборных моделей» как генетические алгоритмы [12] или такие алгоритмы отбора слабых классификаторов как Adaboost [13].

Рассмотрим вкратце идею применения метода искусственного генетического отбора для формирования алгоритмов обнаружения объектов на изображениях.

Пусть дан *набор эталонных изображений* $\mathbf{A}=\{A_i(\mathbf{p})\in\Omega\}$ вместе с *обучающей информацией* об истинных параметрах локализации объектов на эталонных изображениях $\mathbf{U}(\mathbf{A})=\{A_i(\mathbf{u})\in\{0,1\},\ \mathbf{u}\in Q\}$, где Q – пространство параметров локализации; $A_i(\mathbf{u})$ – списки параметров локализации объектов на изображениях $A_i(\mathbf{p})\in\mathbf{A}$.

Требуется сформировать модель объекта $M(\mathbf{p},\mathbf{u})$ вида:

$$M(\mathbf{p},\mathbf{u})=\vee_{k=1..n}\,(M_k(\mathbf{u},\mathbf{q}_k)\bullet\varphi_k(t_k,\mathbf{p},\mathbf{q}_k)), \qquad (6.1)$$

где n-количество значимых для модели яркостно-геометрических примитивов в наборе $\{\varphi_k(t_k,\mathbf{p},\mathbf{q}_k)\}\subseteq\Omega$; t_k, - тип k-го примитива; $\mathbf{q}_k\in Q$ – геометрические параметры k-го примитива; $\mathbf{u}\in Q$ – вектор параметров локализации объекта $M(\mathbf{p},\mathbf{u})$; $M(\mathbf{u},\mathbf{q})=\cup\{M_k(\mathbf{u},\mathbf{q}_k)\}=\{0,1\}$ – модель локализации объекта, которая описывает все допустимые соответствия между значениями параметров локализации образа в целом и параметрами локализации составляющих его геометрических примитивов различного типа.

Пусть данной модели соответствует корреляционное поле $K(\mathbf{u},\mathbf{q})$ (2.7) и пороговое правило принятия решения об обнаружении объекта $D(A,M,\mathbf{u})=\{0,1\}$ вида:

$$D(A(\mathbf{p}),M(\mathbf{p},\mathbf{u})) = \{1 \text{ если } K(\mathbf{u})>K_{min};\ 0 \text{ – в противном случае}\} \qquad (6.2)$$

Определим функционал качества обнаружения $\mathcal{F}(M,\mathbf{A})$, штрафующий несовпадение *множества результатов* $\mathbf{D}(M,\mathbf{A})=\{D(A_i(\mathbf{p}),M(\mathbf{p},\mathbf{u}))\in\{0,1\},\mathbf{u}\in Q\}$ и обучающей информации $\mathbf{U}(\mathbf{A})$ на обучающей выборке A. При решении содержательных задач функционал качества должен быть составлен так, чтобы учитывать ошибки первого и второго рода («необнаружения» и «ложные обнаружения»), а также различным образом штрафовать аномальные и нормальные ошибки (то есть различать «необнаружение объекта» и «неточную локализацию»).

Таким образом, в формальном плане задача формирования модели объекта по обучающей выборке состоит в минимизации функционала качества обнаружения:

$$\mathcal{F}(M,\mathbf{A})=\mathrm{func}(\mathbf{D}(M,\mathbf{A}),\mathbf{U}(\mathbf{A}))\to\min(M). \qquad (6.3)$$

При этом задача оптимизации (6.3) оказывается чрезвычайно многокритериальной, так как переменными параметрами здесь являются практически все составляющие модели (6.1): количество и набор типов яркостно-геометрических примитивов; размерность и состав пространства параметров локализации объектов и примитивов; выбор вариантов расположения (локализации) примитивов относительно образа в целом.

Кроме этого, на практике необходимо учитывать также критерии *вычислительной реализуемости*. Прежде всего, это ограничения на необходимые для осуществления обнаружения:

- *время вычислений*, характеризуемое функционалом

$$\mathcal{T}(M)= \sum_{k=1..n} (T(\varphi_k(t_k,\mathbf{q}_k))) + T_A(K(\mathbf{u})), \qquad (6.4)$$

где $T(\varphi_k(t_k,\mathbf{q}_k),\mathbf{A})$ – время анализа изображений из A, определяемое регистрацией и обработкой событий, связанных с соответсвующими структурирующими элементами; $T_A(K(\mathbf{u}))$ – время анализа корреляционного поля и принятия решения.

- *объем памяти*, характеризуемый функционалом

$$\mathcal{V}(M)= \mathcal{V}(A(P)) + \mathcal{V}(\Theta(Q)) + \mathcal{V}(K(U)), \qquad (6.5)$$

где $\mathcal{V}(A(P))$ – определяется размерностью исходного изображения и количеством необходимых для вычисления буферных массивов; $\mathcal{V}(\Theta(Q))$ – определяется размерностью и дискретом пространства параметров разложения (размерами массивов-аккумуляторов); $\mathcal{V}(K(U))$ – определяется размерностью и дискретом формируемого корреляционного поля.

Как правило, ставятся задачи минимизации времени вычислений (6.4) при ограничениях на используемый объем памяти (7.2), либо ограничения обоих типов непосредственно учитываются при оптимизации функционала (6.3), то есть решается условная задача оптимизации

$$\mathcal{F}(M,\mathbf{A})\to\min(M):\ \mathcal{T}(M)\le\mathcal{T}_{\max};\ \mathcal{V}(M)\le\mathcal{V}_{\max}. \qquad (6.6)$$

Как видно, задача получается крайне многокритериальной и противоречивой. В самом деле – в то время как повышение вероятности обнаружения требует построения мощных моделей с большим количеством элементов, сложными связями и множеством параметров, требования вычислительной реализуемости, напротив, требуют снижать размерность пространства параметров, уменьшать количество типов и число образующих в составе модели. Такие многокритериальные задачи сложной структурной оптимизации плохо поддаются решению традиционными аналитическими или вычислительными параметрическими методами.

В работе [12] предлагается следующая схема применения генетического алгоритма для задач данного типа.

1. Каждому *гену* соответствует один из возможных структурных примитивов, характеризуемый набором $\{M_k(\mathbf{u},\mathbf{q}_k),t_k,\mathbf{q}_k\}$.

2. *Хромосома* - последовательность генов ограниченной длины. Каждая хромосома соответствует одной из возможных моделей $M(\mathbf{p},\mathbf{u})$, определяемых на основе генов выражением (6.1).

3. *Функция качества* для хромосомы вычисляется в духе критериев (6.3)-(6.6) с учетом аппаратно-программной архитектуры предполагаемого вычислителя.

4. *Операция скрещивания* позволяет конструировать новые модели объектов (и соответственно процедуры обнаружения) на базе уже построенных. Новая процедура формируется путем перегруппировки составных частей (групп элементарных составляющих) существующих решений.

5. *Операция мутации* позволяет изменить параметры локализации $\{M_k(\mathbf{u},\mathbf{q}_k),\mathbf{q}_k\}$ для выбранного элемента модели.

6. Генетический отбор осуществляется путем итеративного «размножения», тестирования и отбора в каждом поколении хромосом моделей с наименьшим значением функции качества. При этом на каждом этапе итеративного процесса случайным образом осуществляются мутации параметров и скрещивание моделей.

После определенного числа итераций процесс имитации генетического отбора, как правило, сходится. Результатом применения описанного метода является, с одной стороны, последовательность элементарных процедур обнаружения структурных примитивов, которая формирует искомую процедуру обнаружения объекта в целом, с другой стороны – соответствующая однородная структурная модель искомого объекта. При этом сформированная морфологическая процедура обнаружения удовлетворяет заданным требованиям по вероятности и точности обнаружения, а также условиям, налагаемым на объем памяти и количество вычислений.

Необходимо отметить, что поскольку теория генетических алгоритмов не гарантирует нахождения оптимального решения задач оптимизации, то полученный результат понимается лишь как некоторая квази-оптимальная, то есть близкая к оптимальной, но не обязательно оптимальная процедура обнаружения искомого объекта. Аналогичным образом, и сформированная таким методом модель объекта может рассматриваться не как «истинная» или хотя бы «оптимальная» в некотором классе рассматриваемых моделей, но лишь как «адекватная выборке», то есть позволяющая осуществить качественное обнаружение объекта на имеющихся в обучающей и тестовых выборках изображениях.

7. Вероятностный аспект алгоритмов морфологического обнаружения объектов. Вне зависимости от того, какую вычислительную структуру алгоритмов обнаружения мы используем (согласованная фильтрация «сверху-вниз» или голосование «снизу-вверх»), при структурном анализе изображений всегда встает вопрос о вероятности обнаружения или необнаружения объекта, описываемого моделью (2.1).

В работах [11,14] для разработки широкого круга алгоритмов вероятностного анализа цифровых изображений, обеспечивающих обнаружение сложных объектов, заданных модельными описаниями, был предложен *обобщённый метод анализа свидетельств*. В рамках данного метода каждая характерная особенность изображения рассматривается как возможное *свидетельство* в пользу гипотезы (ряда гипотез) о наличии и характеристиках искомого объекта в соответствии с заданной яркостно-геометрической моделью, причем голосование свидетельствующих событий в пользу гипотез осуществляется непосредственно в процессе низкоуровневой обработки, и каждое выявленное событие (характерная черта) соответствующего типа инициирует обработку лишь тех гипотез, на апостериорную вероятность которых данное событие может повлиять. Рассмотрим теперь, как схема событийного анализа изображений может быть распространена на случай морфологических свидетельств. При этом рассматривавшимся ранее абстрактным “событиям“ и “гипотезам“ необходимо придать смысл “морфологических событий“ и “морфологических гипотез“, а связям между ними – смысл яркостно-геометрических связей, определяемых вхождением одних структурных элементов в другие.

В этом случае, как и ранее, *морфологическим событием* (event) $e(\mathbf{q})$ может считаться факт обнаружения в процессе анализа изображения $A(\mathbf{p})$ такого значения параметров локализации структурного примитива $\mathbf{q}$, что $\Pr(A(\mathbf{p}),\varphi(\mathbf{p},\mathbf{q}))>\varnothing$, т.е. $A(\mathbf{q})>0$.

Морфологической гипотезой $H(\mathbf{u})$ является гипотеза о том, что на изображении $A(\mathbf{p})$ присутствует объект, описываемый моделью $M(\mathbf{u},\mathbf{q})$ (2.1), где $\mathbf{u}\in Q$ – вектор параметров локализации объекта $M(\mathbf{p},\mathbf{u})$; $\mathbf{q}\in Q$ – вектор параметров локализации структурных примитивов данного разложения $\varphi(\mathbf{p},\mathbf{q})$.

И если для детерминированного случая мы рассматривали только *логическую модель объекта* $M(\mathbf{u},\mathbf{q})=\{0,1\}$, которая описывает все описывающую все логически допустимые геометрические соответствия между значениями параметров локализации образа в целом и параметрами локализации составляющих его геометрических примитивов, то в условиях реальных изображений, в присутствии шумов и искажений к ней также необходимо добавить *вероятностную модель объекта* $\boldsymbol{H}(\mathbf{u},\mathbf{q})$, задающую вероятностную меру связи каждого данного морфологического события с каждой данной морфологической гипотезой в виде следующего двухкомпонентного вектора:

$$\boldsymbol{H}(\mathbf{u},\mathbf{q})=\{P(e(\mathbf{q})/H(\mathbf{u})),P(e(\mathbf{q})/H^c(\mathbf{u}))\}, \quad (7.1)$$

где $P(e/H)$ – условная вероятность возникновения события e при условии истинности гипотезы H, $P(e/H^c)$ - условная вероятность возникновения события e при условии истинности гипотезы «не H».

Пусть теперь наблюдается изображение $A(\mathbf{p})$, и необходимо определить апостериорную вероятность некоторой гипотезы $H(\mathbf{u})$ относительно видимой сцены. Тогда формула Байеса принимает вид

$$P(H(\mathbf{u})/A(\mathbf{p}))=[P(H(\mathbf{u}))\times P(A(\mathbf{p})/H(\mathbf{u}))]/$$
$$[P(H(\mathbf{u}))\times P(A(\mathbf{p})/H(\mathbf{u}))+P(H^C(\mathbf{u}))\times P(A(\mathbf{p})/H^C(\mathbf{u}))], \quad (7.2)$$

где $P(H(\mathbf{u}))$, $P(H^C(\mathbf{u}))$ – априорные вероятности соответствующих гипотез.

Изображение $A(\mathbf{p})$ здесь также рассматривается как событие или, точнее, должно рассматриваться связанное с данным изображением *совокупное морфологическое событие*

$$E(A(\mathbf{p}))=E(\mathbf{q})=\{e(\mathbf{q})=(A(\mathbf{q})>t),\ \mathbf{q}\in Q\}, \quad (7.3)$$

где t – порог чувствительности алгоритма обнаружения структурных примитивов. То есть будем считать, что в процессе структурного анализа изображения Im выявлен ряд событий $e(\mathbf{q})$, совокупность которых и составляет $E(\mathbf{q})$. $E(\mathbf{q})$ можно также рассматривать как бинарную функцию со значениями $\{0,1\}$, где $E(\mathbf{q})=1$ означает, что событие $e(\mathbf{q})$ произошло, а $E(\mathbf{q})=0$ означает, что событие $e(\mathbf{q})$ не зарегистрировано.

Заметим, что, вообще говоря, переход от изображения, представленного в виде дискретного двумерного числового поля $A(\mathbf{p})$ к представлению в виде множества событий $E(\mathbf{q})$ не является очевидным, ни даже обоснованным. Однако такой переход будет являться обоснованным в том случае, если $E(\mathrm{A}(\mathbf{p}))$ является *достаточной статистикой* для $A(\mathbf{p})$ относительно гипотезы $H(\mathbf{u})$, то есть:

$$\mathrm{P}(\mathrm{A}(\mathbf{p})/E(\mathbf{q}),H(\mathbf{u})) = \mathrm{P}(\mathrm{A}(\mathbf{p})/E(\mathbf{q})).$$

Если при этом все события $e(\mathbf{q})$ являются независимыми в совокупности (что справедливо для монотонных разбиений, а также некоторых других ортогональных разложений), то соответствующие модели (7.1) условные вероятности события $E(\mathbf{q})$ имеют вид:

$$\mathrm{P}(E(\mathbf{q})/H(\mathbf{u}))= \Pi_{E(\mathbf{q})=1}\{\mathrm{P}(e(\mathbf{q})/H(\mathbf{u}))\}\times\Pi_{E(\mathbf{q})=0}\{1-\mathrm{P}(e(\mathbf{q})/H(\mathbf{u}))\};$$

$$\mathrm{P}(E(\mathbf{q})/H^{\mathrm{C}}(\mathbf{u}))= \Pi_{E(\mathbf{q})=1}\{\mathrm{P}(e(\mathbf{q})/H^{\mathrm{C}}(\mathbf{u}))\}\times\Pi_{E(\mathbf{q})=0}\{1-\mathrm{P}(e(\mathbf{q})/H^{\mathrm{C}}(\mathbf{u}))\}, \qquad (7.4)$$

где $\Pi_{\chi(\mathbf{q})}$ – обозначает произведение с итерацией по элементам множества $\chi(\mathbf{q})\in\mathrm{Q}$.

Подставляя (7.4) в (7.2), имеем:

$$\mathrm{P}(H(\mathbf{u})/\mathrm{A}(\mathbf{p}))=[\mathrm{P}(H(\mathbf{u}))\times\mathrm{P}(E(\mathbf{q})/H(\mathbf{u}))]/$$
$$[\mathrm{P}(H(\mathbf{u}))\times\mathrm{P}(E(\mathbf{q})/H(\mathbf{u}))+\mathrm{P}(H^{\mathrm{C}}(\mathbf{u}))\times\mathrm{P}(E(\mathbf{q})/H^{\mathrm{C}}(\mathbf{u}))]. \qquad (7.5)$$

Выражение (7.5) дает возможность определить важное понятие *"влияющего события"* или *"свидетельства"*. Любое событие $e(\mathbf{q}')$, такое что

$$\mathrm{P}(e(\mathbf{q}')/H(\mathbf{u})) \neq \mathrm{P}(e(\mathbf{q}')/H^{\mathrm{C}}(\mathbf{u})) \qquad (7.6)$$

является *влияющим событием* для гипотезы $H(\mathbf{u})$. Это позволяет считать, что произведение в формуле (7.4) берется не по всем возможным морфологическим событиям из Q, а только по совокупности влияющих событий для каждой исследуемой гипотезы $H(\mathbf{u})$. Согласно семантике задачи, это, очевидно, соответствует *носителю яркостно-геометрической модели* $\pi_t(\mathrm{M}(\mathbf{u},\mathbf{q}))$, рассматриваемой с учетом порога чувствительности алгоритма обнаружения структурных примитивов t.

С точки зрения конструирования конкретных алгоритмов анализа изображений принципиальным моментом здесь является то, что выражение (7.4) имеет форму однородного произведения, что позволяет, как и в детерминированном случае, организовать процесс вычисления вероятностных оценок достоверности гипотез о присутствии объектов, описываемых вероятностно-геометрическими моделями как процесс независимого голосования морфологических событий, обладающий аналогичными возможностями увеличения вычислительной эффективности процедур анализа изображения.

При этом для различных типов морфологических разложений необходимо использовать различные процедуры анализа (объединения) свидетельств:

– для *разбиений и ряда ортогональных разложений* может непосредственно использоваться *байесовское* правило объединения независимых свидетельств (7.5);

– для *иерархических разбиений и разложений* (см. ниже) могут применяться методики объединения свидетельств на базе Марковских реляционных гиперграфов, описанные в работе [11];

– для *наиболее общего случая морфологического анализа* методика объединения свидетельств в пользу сложных пересекающихся гипотез, описанная в работе [14], позволяет формировать лишь две различные независимые оценки достоверности гипотезы {Bel,Pls}. В случае монотонных разбиений эта методика дает оценку в виде доверительного интервала [Bel,Pls].

Более того, поскольку получение достоверных оценок вероятностных связей между моделями объектов и их элементами на изображениях является достаточно сложной и не всегда разрешимой задачей, на практике часто оказывается удобно при составлении модели объекта использовать не истинно вероятностные, а всего лишь нормированные на интервале [0,1] *нечеткие* меры принадлежности элемента данной модели. В качестве таких мер, в частности, удобно использовать описанные выше морфологические коэффициенты корреляции. В этом случае вместо *бинарной модели объекта* $M(\mathbf{u},\mathbf{q})=\{0,1\}$ при построении алгоритма локализации рассматривается соответствующая нечеткая модель объекта $M(\mathbf{u},\mathbf{q})=[0,1]$. Объединение свидетельств, характеризуемых такими псевдо-вероятностными мерами, описывают известные инструменты нечеткой логики. Таким образом, детерминированный и вероятностный подходы объединяются в некоторую общую схему, в которой речь идет об объединении морфологических свидетельств, обладающих определенной мерой *достоверности*.

8. Иерархический структурный анализ. *Иерархический морфологический анализ* предполагает, что к результату первичного морфологического анализа исходного изображения (т.е. к сформированному разложению изображения в пространстве морфологических гипотез) вновь применяется морфологический анализ (разложение первого разложения во второе разложение, соответствующее пространству «сборных» гипотез, состоящих из групп гипотез первого уровня). Такую процедуру можно назвать *морфологическим разложением второго порядка*. Далее по индукции могут быть определены морфологические разложения любого N-го порядка.

Дадим формальные определения. Назовем моделью образа нулевого порядка выражение вида

$$M(\mathbf{q}^0)=\vee_{q^1\in Q^1}(M(\mathbf{q}^1)\bullet\varphi^1(\mathbf{q}^0,\mathbf{q}^1)), \qquad (8.1)$$

где $\mathbf{p}=\mathbf{q}^0$. Далее, по индукции модели более высоких порядков определяются выражением

$$M(\mathbf{q}^1)=\vee_{q^2\in Q^2}(M(\mathbf{q}^2)\bullet\varphi^2(\mathbf{q}^1,\mathbf{q}^2)),$$

$M(\mathbf{q}^{n-1})=\bigvee_{q^n\in Q^n}(M(\mathbf{q}^n)\bullet\varphi^n(\mathbf{q}^{n-1},\mathbf{q}^n)).$ (8.2)

Определим проекцию изображения $A(\mathbf{p})=A(\mathbf{q}^0)$ на модель первого порядка $M(\mathbf{q}^0)$ (8.1) стандартным образом:

$Pr(A(\mathbf{q}^0),M(\mathbf{q}^0))= \bigvee_{q^1\in Q^1}(A(\mathbf{q}^1)\bullet\varphi^1(\mathbf{q}^0,\mathbf{q}^1)),$

$A(\mathbf{q}^1)=Pr(A(\mathbf{q}^0),\varphi^1(\mathbf{q}^0,\mathbf{q}^1)).$ (8.3)

Полученный образ $A(\mathbf{q}^1)$, в свою очередь, спроецируем на модель второго порядка $M(\mathbf{q}^1)$ (8.2):

$Pr(A(\mathbf{q}^1),M(\mathbf{q}^1))= \bigvee_{q^2\in Q^2}(A(\mathbf{q}^2)\bullet\varphi^2(\mathbf{q}^1,\mathbf{q}^2)),$

$A(\mathbf{q}^2)=Pr(A(\mathbf{q}^1),\varphi^2(\mathbf{q}^1,\mathbf{q}^2)).$

Комбинация этих проекций позволяет определить *проекцию второго порядка*:

$Pr^2(A(\mathbf{q}^0),M(\mathbf{q}^1))= \bigvee_{q^1\in Q^1}(Pr(A(\mathbf{q}^1),M(\mathbf{q}^1))\bullet\varphi^1(\mathbf{q}^0,\mathbf{q}^1)).$ (8.4)

То, что выражение (8.4) также является оператором проекции, проверяется непосредственной подстановкой. Аналогичным образом могут быть определены и морфологические проекции более высоких порядков.

Яркостно-геометрический смысл иерархических морфологических моделей заключается в том, что это морфологические модели, состоящие из элементов, которые в свою очередь состоят из еще более простых элементов. Например, почтовый индекс состоит из отдельных символов, каждый из которых состоит из набора прямолинейных отрезков. В рамках иерархического анализа, чтобы сформировать гипотезу о значении индекса, следует сначала выделить каждый символ в отдельности, хотя ничто не мешает построить для заданного изображения N-значного индекса такую процедуру, в которой каждая черта нижнего уровня будет сразу и непосредственно голосовать в пользу гипотезы верхнего уровня.

Как мы уже отмечали выше, снижение числа гипотез при голосовании имеет большое значение с точки зрения снижения размерности пространства параметров, но при этом необходимо также избежать роста числа сочетаний голосующих элементов. Для этого в стратегиях голосования часто вводятся *дополнительные условия (фильтры) на группы голосующих черт*. Например, в пользу гипотезы об окружности голосуют не любые две точки, а две точки, находящиеся на некотором ограниченном расстоянии друг от друга; в пользу гипотезы о прямоугольнике голосуют только пары параллельных линий и т.п. Легко заметить, что на самом деле процедура выделения небольших групп черт, параметры которых удовлетворяют некоторым заданным отношениям – сами по себе есть процедуры структурного анализа второго уровня. А поскольку их результаты далее голосуют в пользу более сложных моделей – это уже элемент морфологического анализа третьего уровня. Таким образом, мы имеем здесь движение в сторону построения иерархической многоуровневой системы морфологического анализа. Замечательно то, что движение в направлении усложнения системы морфологического анализа является прямым следствием стремления к достижению вычислительной эффективности.

Помимо того, что иерархический анализ изображений экономит вычислительные ресурсы, он имеет еще два важных свойства. Во-первых, иерархический анализ позволяет унифицировать первый этап обнаружения в том случае, если необходимо обнаруживать на изображении не один, а несколько различных типов объектов, модели которых при этом могут быть построены из образующих одного типа (из единого структурного «алфавита»). Во-вторых, хотя морфологический анализ структур с однородными связями не рассматривает отношения между элементами модели, а только ее состав и отношение частей к целому, отношения между элементами модели в некоторых случаях все же могут быть косвенно описаны при помощи иерархических структурных моделей. Это можно продемонстрировать на примере задачи обнаружения объекта, состоящего из двух параллельных линий. Иерархическая процедура поиска такого объекта может быть реализована путем морфологической обработки (открытия) маской из двух точек на строке аккумулятора пространства Хафа классической параметризации. При этом взаимное расположение точек маски будет соответствовать расстоянию между прямыми. В то же время, для произвольного объекта из двух непараллельных прямых такая процедура построена быть не может, так как в пространстве Хафа характеристическая фигура из двух точек перестанет быть постоянной по форме при сдвиге и вращении объекта. Этот пример, как мы видим, демонстрирует существенные ограничения исследуемого морфологического подхода даже для случая иерархического морфологического анализа по сравнению с общим модельным походом, способным описывать произвольные классы моделей (в терминах не только элементов, но и отношений) и разрабатывать для них соответствующие процедуры обнаружения и идентификации на изображениях.

Заключение. В данной статье описан общий подход к конструированию алгоритмов обнаружения и идентификации на изображениях объектов, описываемых структурными моделями с однородными связями, опирающийся на введенный в работе [1] формализм проективной морфологии. Рассмотрены методы анализа изображения «сверху-вниз» и «снизу-вверх». Показано, что в отличие от «переборных» корреляционных методов обнаружения, морфологические методы, основанные на голосовании структурных элементов изображения в пользу гипотез о присутствии и положении на изображении искомого объекта, обладают следующими возможностями увеличения вычислительной эффективности:

- Независимое (в смысле очередности подачи голосов) голосование образующих элементов в пользу параметров локализации объекта;
- Возможность декомпозиции или редукции модели объекта (группировка или уменьшение количества анализируемых образующих в модели);
- Возможность декомпозиции или редукции вектора параметров (группировка или уменьшение количества анализируемых параметров модели).

Доказано, что обнаружение объекта на основе загрубленной модели либо редуцированного вектора параметров локализации является необходимым условием обнаружения объекта по полной модели. При этом платой за достижение вычислительной эффективности становится потеря однозначности интерпретации результатов анализа аккумулятора морфологического преобразования, поэтому их постпроверка путем повторного анализа изображения с использованием полной модели объекта становится обязательной.

Рассмотрена задача автоматического построения морфологической модели объекта по заданной обучающей выборке. Описан подход к решению данной задачи, основанный на использовании генетических алгоритмов отбора.

Рассмотрен вероятностный аспект алгоритмов морфологического обнаружения объектов. Дана морфологическая интерпретация предложенного ранее обобщённого метода анализа свидетельств. Непосредственно показано, каким образом проективная морфология позволяет придать рассматривавшимся ранее абстрактным “событиям“ и “гипотезам“ смысл “морфологических событий“ и “морфологических гипотез“.

Описан иерархический морфологический анализ. Иерархические структурные модели могут быть представлены как модели, состоящие из элементов, которые в свою очередь состоят из еще более простых элементов. При этом снижение числа гипотез при голосовании имеет большое значение с точки зрения снижения размерности пространства параметров морфологического разложения. Коме того, иерархический анализ позволяет унифицировать первый этап обнаружения объектов в том случае, когда необходимо обнаруживать на изображении не один, а несколько различных типов объектов, модели которых при этом могут быть построены из образующих одного типа.

Литература

1. **Визильтер Ю.В., Желтов С.Ю**. Проективные морфологии и их применение в структурном анализе цифровых изображений. Теория и системы управления, 2007. *(в печати).*

2. **Serra J.** Image Analysis and Mathematical Morphology, Academic Press, London, 1982.

3. **Serra J.** Introduction to mathematical morphology. Computer Vision, Graphics and Image Processing, 1986. V. 35, № 3.

4. **Пытьев Ю.П.** Морфологический анализ изображений. Доклады АН СССР, 1983. Т. 269. № 5. С. 1061-1064.

5. **Пытьев Ю.П.** Задачи морфологического анализа изображений // Математические методы исследования природных ресурсов Земли из Космоса. М.: Наука, 1984. С.41-83.

6. **Hough P.V.C.** Methods and Means for Recognizing Complex Patterns. US, Patent 3069654, 1962.

7. **Ballard D.H.** Generalizing the Hough transform to detect arbitrary shapes. Pattern Recognition, 1981. № 13(1.2). P. 111-122.

8. **Ballard D.H., Brown C.M.** Computer Vision. Prentice-Hall, Englewood Cliffs, New Jersey, 1982.

9. **Davies E.R.** Machine Vision: Theory, Algorithms, Practicalities. San Diego: Academic Pres, 1997. P.750.

10. **Гонсалес Р., Вудс Р.**, Цифровая обработка изображений. – М.: Техносфера, 2005, с.1072.

11. **Визильтер Ю.В.** Применение метода анализа морфологических свидетельств в задачах машинного зрения. Вестник компьютерных и информационных технологий, N9, 2007.

12. **Buryak D.Y., Vizilter. Y.V.** Application of genetic algorithms for automated construction of image analysis procedures. Pattern Recognition and Image Analysis, Vol. 13, No. 1, 2003, pp. 77-79.

13. **Бекетова И.В., Каратеев С.Л., Визильтер Ю.В., Бондаренко А.В., Желтов С.Ю.** Автоматическое обнаружение лиц на цифровых изображениях на основе метода адаптивной классификации AdaBoost. Вестник компьютерных и информационных технологий, N5, 2007.

14. **Визильтер Ю.В.** Объединение свидетельств при проверке сложных гипотез в задачах распознавания образов. Мехатроника, автоматизация, управление. 2006. № 3. С.26-32.