ПРОЕКТИВНЫЕ МОРФОЛОГИИ И ИХ ПРИМЕНЕНИЕ В СТРУКТУРНОМ АНАЛИЗЕ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ[1]

© 2007 г. Ю.В. Визильтер, С.Ю.Желтов

Москва, ФГУП ГосНИИАС

Поступила в редакцию 20.10.07 г.

Предложен формализм проективной морфологии, позволяющий описывать и сравнивать структурные модели и структурные объекты на изображениях вне зависимости от специфики прикладной области. Рассмотрены основные задачи структурного анализа цифровых изображений, описаны методы их решения в рамках проективных морфологий.

Введение. Классическим методом распознавания образов является так называемое «признаковое распознавание». В рамках этого подхода предполагается, что любому объекту из предметной области ставится в соответствие некоторый характеризующий его вектор признаков, после чего анализ природы исследуемого объекта полностью осуществляется в пространстве признаков путем явного или неявного сравнения данного вектора признаков с аналогичными векторами признаков других объектов известных классов. Решение этой задачи является на сегодняшний день достаточно глубоко проработанным. Существуют хорошо известные метрические, логические, вероятностные и алгебраические подходы к признаковой классификации и обучению признаковых классификаторов ([1]-[4]). Широкое применение нашли нейросетевые методы распознавания [5]. Среди наиболее перспективных современных методов классификации и обучения можно, например, отметить классификаторы типа Adaboost, основанные на усилении слабых классификаторов [6]. Таким образом, в общетеоретическом плане задача признакового распознавания образов может считаться в значительной степени решенной, хотя, очевидно, прогресс в этой области будет продолжаться, и со временем возникнут еще более эффективные классификаторы по сравнению с ныне существующими.

Однако при попытке приложения методов признакового распознавания к конкретной предметной области, например, к задачам анализа и обработки цифровых изображений сразу возникает целый ряд принципиальных проблем, а именно:

- какой набор признаков следует выбрать для описания объектов каждого конкретного типа?
- как гарантировать соответствие формальной «близости» или «сходства» векторов в пространстве признаков и семантического сходства образов объектов на изображениях?

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

- как обеспечить необходимую устойчивость (вероятность) распознавания объектов в данном пространстве признаков с учетом наличия реальных шумов и искажений, воздействующих на образы объектов?
- как использовать сформированный вектор признаков не только для распознавания, но и для конструирования алгоритма выделения (обнаружения, локализации) объектов заданного типа на изображении в присутствии сложного фона, других объектов, а также помех и искажений?
- как обеспечить вычислительную эффективность конструируемых процедур анализа изображения для заданного типа архитектуры вычислительных средств?

На все эти вопросы общая теория признакового распознавания образов ответить не может, так как проблема получения векторов признаков и их семантической связи с объектами, подлежащими распознаванию, в ней не рассматривается.

В то же время, в каждой конкретной предметной области существует множество подходов к анализу данных, основанных именно на содержательных моделях, учитывающих семантически специфические особенности анализируемых образов. В области обработки изображений это так называемые «яркостно-геометрические» модели, описывающие структуру распределения яркостных и геометрических особенностей объекта на изображении. При всем различии используемого математического аппарата, опора на богатые и содержательные яркостно-геометрические модели изображения безусловно характерна для таких современных методов анализа изображений как:

- *математическая морфология Серра* (ММ), классическое описание которой дано в работах Серра [7,8];
- *морфологический анализ Ю.П. Пытьева* (МП), описанный в работах [9,10];
- *методы голосования*, восходящие к *преобразованию Хафа*, первоначально описанному Хафом [11], и *обобщенному преобразованию Хафа*, предложенному Баллардом [12,13] и затем развитые Дэвисом [14,15].
- *оптимально-вероятностные* (байесовские) методы анализа изображений [16-19].
- *частотные методы фильтрации* на базе двумерного *преобразование Фурье* (включая БПФ, ДКП и др.) и пространственно-частотные методы фильтрации, такие как *кратномасштабный вейвлет-анализ* [20].

Здесь необходимо сразу уточнить используемые термины. Вообще говоря, под *модельным подходом* в компьютерном зрении, как правило, понимается подход, основанный на моделях более высокого уровня, причем *модель* должна включать ряд геометрических примитивов (*элементов*, образующих) и набор связей (*отношений*) между ними. Такие модели адекватно описываются реляционными гиперграфами общего вида и не могут быть сведены к структурам типа массивов или векторов признаков. Однако, как будет показано ниже, *яркостно-геометрические модели*, используемые в перечисленных современных методах анализа изображений непосредственно на уровне пиксельного анализа, имеют более простую структуру, а именно: модель описывает лишь состав входящих в нее элементов и связь между этими образующими и образом в целом, но не рассматривает связи образующих элементов между собой. Такой класс моделей мы предлагаем далее называть *моделями с однородной структурой связей* или просто *однородными структурами*. Такие однородные структуры, в отличие от моделей общего вида, всегда могут быть однозначно охарактеризованы массивами или векторами признаков, которые в семантическом плане можно рассматривать как содержательные *структурные описания* соответствующих объектов. Это позволяет, применительно к конкретной предметной области, решать на их основе не только задачу распознавания, но и такие специфические задачи как фильтрация, сжатие и сегментация данных, а также задачу автоматического обнаружения (выделения и локализации) объектов непосредственно в ходе низкоуровневого анализа растровых данных.

В первом разделе данной статьи будут кратко изложены математические основы предлагаемого морфологического формализма структурного анализа для однородных структур общего вида. Во втором разделе будут рассмотрены основные задачи сравнения и фильтрации цифровых изображений и их решение в рамках рассматриваемого класса проективных морфологий.

1. Проективные морфологии однородных структур

1.1. Проективное пространство образов. Пусть имеется два сорта элементов: *скаляры* и *образы*. На множестве скаляров Ψ определены две операции – *умножение* ($\bullet$) и *объединение* ($\vee$). Умножение определяет на множестве Ψ группу с 1, объединение – полугруппу с 0. Образы принимают значения на множестве Ω, на котором также определена операция *объединения* ($\vee$), задающая на Ω полугруппу с «нулевым образом» $\varnothing$. Кроме этого, на множестве образов Ω определена *норма* μ: $\Omega \rightarrow R$. В дальнейшем мы также будем использовать более привычное обозначение нормы образа $\mu(A)=\|A\|$.

Пусть теперь операция *умножения образа на скаляр* определена таким образом, что выполняются следующие свойства *линейной комбинации образов*:

$$aA=Aa\in\Omega;\ (ab)A=r(bA);\ A\bullet 1=A;$$

$$A\bullet 0=\varnothing;\ a\bullet\varnothing=\varnothing;\ \mu(aA)=|a|\bullet\mu(A); \qquad (1.1)$$

$$A(b\vee c)=Ab\vee Ac;\ r(B\vee C)=aB\vee aC,$$

где $0,1,a,b,c,d \in \Psi$; $A,B,C,\varnothing \in \Omega$.

Введем операцию *проекции образа на образ*, обладающую следующими свойствами:

$$Pr(A,B) \in \Omega;\ Pr(A,B)=Pr(Pr(A,B),B);$$
$$Pr(A,A)=A;\ Pr(\varnothing,A)=\varnothing;\ Pr(A,\varnothing)=\varnothing; \qquad (1.2)$$
$$Pr(a\bullet A,B)=a\bullet Pr(A,B);\ \forall a\neq 0:\ Pr(A,a\bullet B)=Pr(A,B);$$

где $a \in \Psi$; $A,B,\varnothing \in \Omega$.

Как видно, в общем случае оператор проекции асимметричен, в нем различаются *проектируемый образ* (A) и *образ, на который осуществляется проектирование* (B).

Назовем *линейным проектором* оператор проекции следующего вида:

$$Pr(A,B)=r(A,B)\bullet B, \qquad (1.3)$$

где $r(A,B)$ – коэффициент *линейной корреляции*, удовлетворяющий следующим свойствам:

$$r(A,A)=1;\ r(\varnothing,A)=0;$$
$$r(\alpha A,B)=\alpha\bullet r(A,B); \qquad (1.4)$$
$$r(A,\alpha B)=r(A,B),$$

$A,B \in \Omega$, $r(A,B),\alpha \in \Psi$.

То, что оператор (1.3) является оператором проекции в смысле (1.2) легко проверяется непосредственной подстановкой (1.3) и (1.4) в условия (1.2).

Как видно, структурно-корреляционная функция $r(A,B):\{A,B\}\to\Psi$ задает здесь в явном виде систему парных отношений *структурного сходства* или *структурной близости* образов, отражающую в каждом конкретном случае семантические свойства образов предметной области. Это позволяет ввести следующую стандартную меру близости двух образов, инвариантную к их линейному преобразованию (умножению на скаляр) – *нормированный коэффициент линейной корреляции*:

$$K(A,B) = \|Pr(A,B)\|/(\|A\|\bullet\|B\|)= \|r(A,B)\|/\|A\|, \qquad (1.5)$$

$A,B \in \Omega$; со следующими свойствами:

1) $0 \le K(A,B)$;
2) $K(A,A)=1$.
3) $K(A,B)=0 \Leftrightarrow Pr(A,B)=\varnothing$.

Таким образом, мы можем охарактеризовать Ω как замкнутое относительно перечисленных операций с образами и скалярами нормированное пространство, на котором определен оператор проекции. Такое пространство можно назвать *проективным пространством образов*. В частном случае, когда оператор **V** образует абелевы группы и на Ψ, и на Ω, данное однородное пространство становится классическим линейным пространством, свойства которого хорошо известны. При этом, если в таком пространстве определено скалярное произведение, коэффициент линейной корреляции (1.3) может рассматриваться как коэффициент Фурье, определяемый из условия минимума нормы проекции.

1.2. Проективная морфология. Пусть задано некое произвольное множество образов $\mathbf{B}\subseteq\Omega$. Очевидно, что благодаря свойствам линейной комбинации образов (1.1) операции объединения образов и умножения образа на число образуют над множеством **B** замкнутое линейное подпространство $\boldsymbol{B}\subseteq\Omega$, такое, что любой образ, принадлежащий этому подпространству, может быть представлен в форме

$$A=\vee_{k=1..n}(a_k B_k), \tag{1.6}$$

где A – описываемый структурный образ, $\mathbf{B}=\{B_k\}$ – множество входящих в данный образ структурных *примитивов* или *образующих,* $\mathbf{a}=\{a_k\}$ – вектор *весов образующих*, характеризующий степень связи (способ вхождения) образующих из **B** в данный конкретный образ A, n – количество образующих, входящих в данную модель.

Если любые два образа из **B** являются *линейно независимыми*, то есть

$$\forall B_i,B_k\in\mathbf{B},\ k\neq i:\ \forall\alpha\in\Psi:\ B_i\neq\alpha B_k, \tag{1.7}$$

тогда множество **B** можно назвать *базисом структурного описания*, элементы этого множества – *структурными примитивами*, а количество структурных примитивов n – соответственно *размерностью базиса*.

Определим операцию *проекции образа на подпространство*, такую что:

$$\Pr(A,\boldsymbol{B})\in\boldsymbol{B};\ \Pr(A,\boldsymbol{B})=\Pr(\Pr(A,\boldsymbol{B}),\boldsymbol{B});$$

$$\Pr(\varnothing,\boldsymbol{B})=\varnothing;\ \Pr(a\bullet A,\boldsymbol{B})=a\bullet\Pr(A,\boldsymbol{B}), \tag{1.8}$$

где $a\in\Psi$; $A,\varnothing\in\Omega$; $\boldsymbol{B}\subseteq\Omega$.

Пусть теперь существует такой базис примитивов **E** размерности n, что проекция любого образа из Ω на замыкание $\boldsymbol{E}$ удовлетворяет следующему *условию разложимости:*

$$\Pr(A,\boldsymbol{E}) = \vee_{k=1..n}(\Pr(A,E_k)) = \vee_{k=1..n}(r(A,E_k)\bullet E_k). \tag{1.9}$$

То есть *проекция образа на базис есть объединение проекций на его элементы*.

Поскольку любому произвольному образу $A\in\Omega$ может быть поставлен в соответствие структурный вектор признаков $\mathbf{a}=\{r(A,E_k)\}$, каждый элемент которого является коэффициентом структурной корреляции данного образа с одним из структурных примитивов, это позволяет определить *операцию морфологического разложения образа по базису* (*decomposition*) как операцию отображения из пространства образов в пространство векторов признаков размерности n:

$\mathbf{dec_E}(A)=\mathbf{a}(A,\mathbf{E})$: $A\in\Omega\to\Psi^n$. (1.10)

Алгебраическую систему $\{\Psi,\Omega,\bullet,\mathbf{V},\mu,\Pr,\mathbf{E}\}$, для которой справедливо условие (1.9), будем далее называть *проективной морфологией* на Ω. Базис **E** является здесь соответственно *базисом морфологического разложения. Пространство векторов разложений* $\Theta=\Psi^n\subseteq R^n$ по некоторому базису разложения **E** представляет собой проективное векторное пространство, на котором определены операции *объединения векторов, умножения вектора на скаляр, проекции вектора на вектор* и задана *норма (модуль) вектора*:

$\mathbf{a}\vee\mathbf{b}=\{a_1,a_2,..,a_n\}\vee\{b_1,b_2,..,b_n\}=\{a_1\vee b_1,\ a_2\vee b_2,..,a_n\vee b_n\}$

$a\bullet\mathbf{b}=a\bullet\{b_1,b_2,..,b_n\}=\{a\bullet b_1,\ a\bullet b_2,..,a\bullet b_n\}$, (1.11)

$\Pr(\mathbf{a},\mathbf{b})=\alpha(\mathbf{a},\mathbf{b})\bullet\mathbf{b}$,

$\mu(\mathbf{a})\geq 0;\ \mu(\mathbf{0})=0;\ \mu(\mathbf{a}\vee\mathbf{b})\leq\mu(\mathbf{a})+\mu(\mathbf{b}),\ \mu(a\mathbf{b})=|a|\bullet\mu(\mathbf{b})$,

где $a,\alpha\in\Psi$; $\mathbf{a},\mathbf{b}\in\Theta$. В дальнейшем мы также будем использовать более привычное обозначение нормы вектора $\mu(\mathbf{a})=\|\mathbf{a}\|$.

Веденные таким образом морфологические разложения обладают рядом важных и практически значимых свойств. Важнейшим из них является следующее. Пусть в пространстве Ω существует операция проекции Pr(A,B), удовлетворяющая свойствам (1.2)-(1.4) и существует разложение $\mathbf{dec_E}$, описываемое свойствами (1.9)-(1.11), тогда в пространстве разложения Θ может быть однозначно задана операция проекции вектора разложения на вектор разложения

$\Pr(\mathbf{a},\mathbf{b})=r(\mathbf{a},\mathbf{b})\bullet\mathbf{b}$,

удовлетворяющая условиям (1.2)-(1.4), такая, что выполняется следующее *условие соответствия*:

$\forall A,B,C\in\boldsymbol{E}$: $C=\Pr(A,B)=r(A,B)\bullet B$: $\mathbf{a}=\mathbf{dec_E}(A)$, $\mathbf{b}=\mathbf{dec_E}(B)$, $\mathbf{c}=\mathbf{dec_E}(C)$,

$\mathbf{c}=\Pr(\mathbf{a},\mathbf{b})=r(\mathbf{a},\mathbf{b})\bullet\mathbf{b}$; $r(A,B)=r(\mathbf{a},\mathbf{b})$. (1.12)

Докажем это утверждение. Подставляя (1.9) в (1.12), имеем:

$\Pr(A,B)=r(A,B)\bullet B = r(A,B)\bullet(\vee_{k=1..n}(b_kE_k)) = \vee_{k=1..n}(r(A,B)\bullet b_k\bullet E_k)$,

откуда

$\mathbf{dec}(\Pr(A,B)) = r(A,B)\bullet\mathbf{dec}(B)$,

что и требовалось доказать. То есть *линейная корреляция векторов разложений оказывается равна линейной корреляции исходных образов*.

Для оценки близости двух образов A и B в пространстве векторов разложений может быть использован инвариантный относительно умножения образа на скаляр *нормированный коэффициент линейной корреляции разложений*:

$K(\mathbf{a},\mathbf{b}) = \|Pr(\mathbf{a},\mathbf{b})\|/(\|\mathbf{a}\|\bullet\|\mathbf{b}\|)= \|r(\mathbf{a},\mathbf{b})\|/\|\mathbf{a}\|,$ (1.13)

где $\mathbf{a}=\mathbf{dec}(A),\mathbf{b}=\mathbf{dec}(B)\in\Theta$, со следующими свойствами:

1) $0 \le K(\mathbf{a},\mathbf{b})$;
2) $K(\mathbf{a},\mathbf{a})=1$.
3) $K(\mathbf{a},\mathbf{b})=0 \Leftrightarrow Pr(A,B)=\varnothing$.

Таким образом, можно утверждать, что вектора разложений из Θ адекватно описывают структурные отношения образов из Ω, что позволяет обоснованно анализировать (сравнивать) образы и модели данного типа, опираясь на их признаковое описание в виде векторов морфологических разложений.

1.3. Типы проективных морфологий. Поскольку в общем случае оператор, сконструированный путем объединения проекций согласно выражению (1.9), не всегда является проектором в смысле условий (1.8), то для того, чтобы гарантировать их выполнение, необходимо наложить дополнительные условия либо на операцию объединения, либо на вид оператора проекции, либо на способ формирования системы примитивов. Соответственно будут определены различные типы морфологических разложений с различными свойствами. Можно выделить два основных типа таких разложений: монотонные и ортогональные.

<u>*Монотонные разложения*</u>

Пусть множество Ω представляет собой *решетку*, то есть частично упорядоченное множество, в котором для любых двух элементов $A,B\in\Omega$ объединение $A\vee B$ является их *точной верхней границей*, а также можно указать их *точную нижнюю границу* $A\wedge B$. Это позволяет определить для элементов решетки Ω *отношение включения*:

$\forall A,B\in\Omega$: $A\subseteq B \Leftrightarrow A\vee B=B$, $\|A\|\le\|B\|$. (1.14)

Наложим теперь на операцию проектирования два дополнительных требования:

(а) *сохранения включения*:

$\forall A,B,C\in\Omega$, $A\subseteq B \Leftrightarrow Pr(A,C) \subseteq Pr(B,C)$. (1.15)

(б) *монотонности*:

$\forall A,B\in\Omega$, $Pr(B,A) \subseteq Pr(A,A)=A$. (1.16)

Условия (1.14)-(1.16) являются *достаточными* для выполнения условия (1.9), то есть *в случае, когда пространство образов является алгебраической решеткой, любой базис примитивов является базисом морфологического разложения*.

Следует отметить, что в силу условия (1.15) морфологические разложения являются избыточными. Избыточность сообщает им дополнительную устойчивость к возможным шумам и искажениям, в то же время с неоднозначностью (неединственностью) монотонных разложений на практике может быть связан ряд проблем.

<u>*Ортогональные разложения*</u>

Назовем *ортогональными* любые два образа $A,B\in\Omega$, такие что

$A\perp B \Leftrightarrow \{Pr(A,B)=\varnothing;\ Pr(B,A)=\varnothing\}.$ (1.17)

Ортогональным базисом примитивов будет являться такой базис примитивов **E**, в котором:

$\forall E_i,E_k\in\mathbf{E},\ k\neq i:\ E_i\perp E_k.$ (1.18)

В этом случае вектор коэффициентов проекции из выражения (1.10) становится вектором *ортогонального морфологического разложения* образа А по базису **E**.

Наложим теперь на операцию проектирования образа на образ следующее дополнительное требование *сохранения объединения*:

$\forall A,B,C\in\Omega,\ A\vee B \Leftrightarrow Pr(A\vee B,C)= Pr(A,C) \vee Pr(B,C).$ (1.19)

Условия (1.18)-(1.19) являются *достаточными* для выполнения условия (1.9), то есть *любой ортогональный базис примитивов является базисом морфологического разложения*. При этом ортогональное разложение является единственным.

Если некоторое разбиение исходно является монотонным, то в случае выбора ортогонального базиса оно становится *разбиением*, т.е. разложением, являющимся одновременно и ортогональным и монотонным.

В заключение данного раздела можно отметить, что введенный формализм проективной морфологии как алгебры моделей с однородными структурными связями является достаточно общим, так как он позволяет описывать столь различные математические объекты как векторы, функции, множества, логические модели на базе предикатов и ряд других. В силу этого, данный формализм может быть успешно применен для описания и конструирования процедур анализа изображений.

2. Морфологический анализ изображений

2.1. Яркостно-геометрические модели процедур структурного анализа изображений.

Рассмотрим представления значений изображения f в точке (x,y), используемые в ряде популярных методов анализа двумерных данных.

<u>Цифровое изображение (тривиальное пиксельное описание)</u>:

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

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:

$$Pr_F(f,x,y)=\Sigma_n(a_n*\varphi(n,x,y)),$$

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*:

$$A(\mathbf{p})=\bigvee_{q\in Q}(A(\mathbf{q})\bullet\varphi(\mathbf{p},\mathbf{q})), \quad (2.1)$$

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*); '$\bigvee$'$\in${'Σ','MAX','Π','MIN'} – коммутативная и ассоциативная операция *попиксельного* (поэлементного) объединения элементов изображения, которая может иметь смысл либо сложения/умножения элементов, либо взятия максимального/минимального значения.

2.2. Проективная морфология изображений. Пусть на проективном пространстве образов $\{\Psi,\Omega,\bullet,\mathbf{V},\mu,Pr\}$ задана *группа преобразований* $\mathcal{G}(\Omega)$. Алгебраическую систему $\{\Psi,\Omega,\bullet,\mathbf{V},\mu(\Omega),Pr,\mathcal{G}(\Omega)\}$ можно назвать *проективно-геометрическим пространством*, если выполняются следующие условия комбинации образов из Ω с преобразованиями из $\mathcal{G}$:

$$\tau(a\bullet B)=a\bullet\tau(B);$$
$$\tau(A\vee B)=\tau(A)\vee\tau(B); \quad (2.2)$$
$$Pr(\tau(A),\tau(B))= \tau(Pr(A,B)),$$

где $a\in\Psi$; $\tau\in\mathcal{G}(\Omega)$; $A,B\in\Omega$.

При этом $\mathcal{G}(\Omega)$ имеет смысл множества *геометрических преобразований* образов, тогда как умножение образов на скаляры имеет здесь смысл *яркостных преобразований*.

Предположим теперь, что группа преобразований $\mathcal{G}(\Omega)$, является *параметрической*:

$$\mathcal{G}(P)=\{g(\mathbf{p})\in\mathcal{G};\ \mathbf{p}\in P\subseteq R^N\}, \quad (2.3)$$

где N – размерность *пространства параметров исходных образов* P; $g(\mathbf{0})=\varepsilon$ - тождественное преобразование.

Таким образом, множество образов Ω также оказывается *параметризованным*:

$$\Omega(P)=\{A(\mathbf{p})=g(\mathbf{p})(A):\ A\in\Omega,\ g(\mathbf{p})\in\mathcal{G};\ \mathbf{p}\in P\subseteq R^N\}. \quad (2.4)$$

Пусть теперь имеется некоторый n-мерный базис морфологического разложения $\mathbf{E}(P)=\{\varphi_k(\mathbf{p})\in\Omega\}$, на котором также задана группа преобразований

$$\mathcal{R}(Q)=\{\sigma(\mathbf{q})\in\mathcal{R};\ \mathbf{q}\in Q\subseteq R^M\}, \qquad (2.5)$$

где M – размерность пространства параметров преобразования Q, $r(\mathbf{0})=\varepsilon$.

В этом случае (M+1)-мерный вектор $(k\in[1..n],\mathbf{q}\in Q)$ можно назвать *полным вектором параметров морфо-геометрического базиса*

$$\mathbf{E}(P,Q)=\{\varphi(k,\mathbf{p},\mathbf{q})=\sigma(\mathbf{q})(\varphi_k(\mathbf{p}))\in\Omega,\ \mathbf{p}\in P,\ \mathbf{q}\in Q,\ k\in[1..n]\}.$$

С учетом введенных определений выражение (1.9), описывающее разложение образа по базису, примет следующий вид:

$$\Pr(A(\mathbf{p}),\mathbf{E}(\mathbf{p},\mathbf{q}))=\vee_{k=1..n;\ q\in Q}(\Pr(A(\mathbf{p}),\varphi(k,\mathbf{p},\mathbf{q})))=\vee_{k=1..n;\ q\in Q}(A(k,\mathbf{q})\bullet\varphi(k,\mathbf{p},\mathbf{q})), \qquad (2.6)$$

где $A(\mathbf{p})$ – *представление образа* A *в исходном пространстве* P; n – *количество типов примитивов* структурного базиса $\mathbf{E(0)}$; $A(k,\mathbf{q})$ – *представление образа* A *в пространстве параметров разложения* $[1..n]\times Q$, где k – *индикатор типа образующей*.

В дальнейшем для простоты обозначений будем считать, что вектор $\mathbf{q}$ является полным вектором параметров и включает в себя также индикатор типа образующей:

$$\Pr(A(\mathbf{p}),\mathbf{E}(\mathbf{p},\mathbf{q}))=\vee_{q\in Q}(A(\mathbf{q})\bullet\varphi(\mathbf{p},\mathbf{q})). \qquad (2.7)$$

Таким образом, в области яркостно-геометрического анализа изображений в качестве характеристических структурных статистик выступают уже не вектора признаков, а представления характеризуемых образов в пространстве параметров. В дискретном случае они имеют вид M-мерных массивов-*аккумуляторов*.

По аналогии с введенным ранее *морфологическим разложением* (1.9) введем *морфо-геометрическое разложение* как оператор, отображающий образ из $\Omega(P)$ в *параметризованное пространство разложений* $\Theta(Q)$:

$$\mathbf{dec}(A(\mathbf{p}))=A(\mathbf{q}):\ A\in\Omega(P)\to\Theta(Q). \qquad (2.8)$$

Параметризованное пространство разложений $\Theta(Q)$ представляет собой пространство M-мерных массивов, на котором определены поэлементные операции *объединения массивов, умножения массива на скаляр* и задана *норма (модуль) массива:*

$$A(\mathbf{q})\vee B(\mathbf{q}) = a[\mathbf{q}]\vee b[\mathbf{q}],\ \forall\ \mathbf{q}\in Q$$

$$a\bullet B(\mathbf{q})=a\bullet b[\mathbf{q}],\ \forall\ \mathbf{q}\in Q, \qquad (2.9)$$

$$\mu(A(\mathbf{q}))\geq 0;\ \mu(0(\mathbf{q}))=0;\ \mu(A(\mathbf{q})\vee B(\mathbf{q}))\leq \mu(A(\mathbf{q}))+\mu(B(\mathbf{q})),\ \mu(a\bullet B(\mathbf{q}))=|a|\bullet\mu(B(\mathbf{q})),$$

где $a\in\Psi$; $A(\mathbf{q}),B(\mathbf{q})\in\Theta(Q)$, $a[\mathbf{q}]$ – элемент массива $A(\mathbf{q})$.

Определим теперь в пространстве параметров $\Theta(Q)$ операцию *проекции разложения на разложение*, такую что

$$\Pr(A(\mathbf{q}),B(\mathbf{q}))=r(A(\mathbf{p}),B(\mathbf{p}))\bullet B(\mathbf{q}), \qquad (2.10)$$

где $A(\mathbf{p}),B(\mathbf{p})\in\mathbf{E}(P)$; $A(\mathbf{q})=\mathbf{dec}(A(\mathbf{p}))$, $B(\mathbf{q})=\mathbf{dec}(B(\mathbf{p}))\in\Theta(Q)$; $r(A(\mathbf{p}),B(\mathbf{p}))$ - линейная структурная корреляция образов $A(\mathbf{p})$ и $B(\mathbf{p})$. Соответственно *нормированный коэффициент линейной корреляции разложений* будет иметь вид:

$$K(A(\mathbf{q}),B(\mathbf{q})) = \|Pr(A(\mathbf{q}),B(\mathbf{q}))\|/(\|A(\mathbf{q})\|\bullet\|B(\mathbf{q})\|). \qquad (2.11)$$

Ведем понятие *геометрического носителя* образа:

$$S_P(A)=\{\mathbf{p}: A(\mathbf{p})\neq 0\};$$

$$S_Q(A)=\{\mathbf{q}: A(\mathbf{q})\neq 0\}, \qquad (2.12)$$

где $S_P(A)$ – носитель образа A в исходном пространстве параметров; $S_Q(A)$ – носитель образа A в пространстве разложений. С учетом этого, элементам яркостно-геометрической модели изображения можно дать следующую семантическую интерпретацию:

- значения $A(\mathbf{q})$ - коэффициенты разложения (веса образующих $E_k(\mathbf{q})$) описывают яркостную природу образа A;
- множество образующих $\{\varphi(\mathbf{p},\mathbf{q})\}$ и носитель образа $S_Q(A)$ описывают его геометрическую природу.

Построенную таким образом проективную морфологию естественно называть *яркостно-геометрической морфологией* или *проективной морфологией изображений*. Подобные проективные морфологии изображений могут быть, в частности, построены на базе всех методов представления и анализа изображений, рассмотренных в разделе 2.1.

2.3. Структурный анализ изображений с использованием проективных морфологий.

Введенные в предыдущих подразделах обобщенные понятия позволяют с единых позиций описывать и разрабатывать целый ряд базовых аспектов теории обработки и анализа изображений. Рассмотрим кратко каждый из них.

Структурная фильтрация изображений

Назовем *морфологическим преобразованием* изображения, описываемого моделью

$$A(\mathbf{p})=\vee_{q\in Q}(A(\mathbf{q})\bullet\varphi(\mathbf{p},\mathbf{q})), \qquad (2.13)$$

любое преобразование ψ, такое что:

$$\psi(A(\mathbf{p}))=\vee_{q\in Q}(\psi(\mathbf{q})\bullet A(\mathbf{q})\bullet\varphi(\mathbf{p},\mathbf{q})), \qquad (2.14)$$

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

Если оператор ψ обладает свойством алгебраического проектора:

$$\psi(\psi(A(\mathbf{p})))=\psi(A(\mathbf{p})), \qquad (2.15)$$

то такое преобразование можно назвать *морфологическим фильтром*.

Учитывая свойства выражения (2.14), любой морфологический фильтр, весовая функция которого не зависит от анализируемого изображения, можно описать следующим образом:

$$f(\mathrm{A}(\mathbf{p}))=\vee_{q\in Q}(f(\mathbf{q})\bullet \mathrm{A}(\mathbf{q})\bullet \varphi(\mathbf{p},\mathbf{q})), \qquad (2.16)$$

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

То есть любой морфологический фильтр может быть также рассмотрен как оператор проекции исходного образа на область пропускания:

$$f(\mathrm{A}(\mathbf{p}))=\mathrm{Pr}(\mathrm{A}(\mathbf{p}),\mathbf{F(p,q)}), \qquad (2.17)$$

где $\mathbf{F}(\mathbf{p},\mathbf{q})=\{\mathrm{E}(p,q)\in\mathbf{E}(\mathbf{p},\mathbf{q}):\ f(\mathbf{q})\neq 0\}$ – базис примитивов, соответствующий области пропускания $\boldsymbol{F}(\mathbf{q})$; $\mathbf{E(p,q)}$ - исходный базис примитивов рассматриваемого разложения.

Таким образом, часть образующих элементов, присутствующих в полной реконструкции изображения, операция морфологической фильтрации f удаляет из изображения *в зависимости от выбранной области пропускания параметров и вне зависимости от значений коэффициентов разложения*. Например, удаляются все высокие частоты, все малоразмерные структурирующие элементы и т.п.

Отсюда возникает унифицированная двухэтапная схема фильтрации изображений:

Этап1: Деконструкция (разложение, анализ). Проектирование изображения на образующие элементы преобразования.

Этап2: Частичная реконструкция (синтез). Объединение проекций на те элементы, которые находятся в области пропускания фильтра.

Вместе эти два этапа всегда определяют операцию морфологической фильтрации, обладающую свойствами алгебраического проектора.

Практическая применимость морфологической фильтрации изображений основывается на следующих двух предположениях:

- наблюдаемое изображение состоит из двух частей – полезного изображения и помехи или фона;

- полезное изображение состоит из структурных элементов с такими параметрами, которые не содержит (или редко содержит) помеховая или фоновая составляющая.

Рассмотрим типичную ситуацию, когда наблюдаемое изображение $\mathrm{A}(\mathbf{p})$ представляет собой объединение *полезного сигнала* $\mathrm{I}(\mathbf{p})\in\mathbf{I}$ и *помехи* $\xi(\mathbf{p})\in\aleph$:

$$\mathrm{A}(\mathbf{p})= \mathrm{I}(\mathbf{p}) \vee \xi(\mathbf{p}), \qquad (2.18)$$

где $\mathbf{I}$ - класс всех возможных изображений объекта; $\aleph$ - класс всех возможных помех.

Идеальным помеховым фильтром, очевидно, будет являться такой морфологический фильтр, что:

$$\forall\ \mathrm{I}(\mathbf{p})\in\mathbf{I}:\ f(\mathrm{I}(\mathbf{p}))=\mathrm{I}(\mathbf{p});$$

$$\forall\ \xi(\mathbf{p})\in\aleph:\ f(\xi(\mathbf{p}))=0(\mathbf{p}), \tag{2.19}$$

так как это гарантирует полное подавление помех при восстановлении полезного сигнала без искажений:

$$f(A(\mathbf{p}))= f(I(\mathbf{p})) \vee f(\xi(\mathbf{p})) = I(\mathbf{p}). \tag{2.20}$$

Необходимым и достаточным условием возможности построения идеального помехового фильтра (2.19) на базе некоторого морфологического разложения является требование непересечения носителей классов полезного сигнала и помехи в пространстве параметров данного разложения:

$$S_Q(\mathbf{I}) \cap S_Q(\aleph) = \varnothing. \tag{2.21}$$

На практике, разумеется, это условие выполняется редко, в связи с чем, как правило, помеховые фильтры не являются идеальными, то есть не полностью подавляют помехи и при этом искажают исходный сигнал.

<u>*Структурное сжатие изображений*</u>

Морфологическое сжатие (с потерями), так же как и морфологическая фильтрация изображений, основано на структурной деконструкции и последующей частичной реконструкции изображения. Разница заключается в том, что часть образующих элементов, присутствующих в полной реконструкции изображения, операция сжатия удаляет из изображения не только в зависимости от выбранной области пропускания в пространстве параметров, но и *в зависимости от значений коэффициентов разложения данного конкретного изображения*. Элементы разложения, вносящие сравнительно небольшой вклад в изображение – удаляются, остальные элементы - *квантуются*. Это позволяет затем значительно более компактно упаковывать изображение стандартными методами сжатия без потерь, так как уменьшается необходимое число разрядов для описания кантованной функции в пространстве параметров, кроме того, резко уменьшается количество ненулевых элементов и увеличивается вероятность возникновения последовательностей элементов, имеющих одинаковые значения (цепочек повторов).

Рассмотрим формальные свойства оператора морфологического сжатия с потерями как отображения в пространстве однородных морфологических разложений. Назовем *операцией квантования* скаляра $\alpha\in[\alpha_{min},\alpha_{max}]\subseteq\Psi$ на T уровней следующую операцию:

$$\pi_T(\alpha) = \alpha_{min} + [(\alpha-\alpha_{min})*(T-1)/(\alpha_{max}-\alpha_{min})]*((\alpha_{max}-\alpha_{min})/(T-1)), \tag{2.22}$$

где $[\alpha]$ – оператор взятия целой части числа α.

Оператор морфологического сжатия (точнее результат реконструкции и декодирования изображения после сжатия и кодирования), представляющий собой комбинацию морфологического фильтра с оператором квантования, может быть представлен следующим образом:

$$c(\mathrm{A}(\mathbf{p}))=\vee_{q\in Q}(f(\mathbf{q})\bullet\pi_T(\mathrm{A}(\mathbf{q}))\bullet\varphi(\mathbf{p},\mathbf{q})), \quad (2.23)$$

где $f(\mathbf{q})$ – индикаторная функция *области пропускания* выбранного фильтра; T – число уровней квантования, используемых в данном алгоритме сжатия.

Поскольку операция квантования (2.22) является оператором скалярной проекции непрерывного множества $[\alpha_{min},\alpha_{max}]$ на дискретное множество $\{\alpha_{min},\ldots,\alpha_{max}\}$, состоящее из T элементов, отсюда следует, что любой оператор морфологического сжатия (2.23) также является проектором, а значит, и морфологическим фильтром согласно условию (2.15).

Структурные спектры и адаптивная фильтрация на основе спектров

Выше (касаясь фильтрации) мы предположили, что, что полезное изображение состоит из структурных элементов с такими параметрами, которые не содержит (или редко содержит) помеховая или фоновая составляющая. В каких-то случаях соответствующая область пропускания действительно может быть сформирована из априорных соображений (например, если требуется пропускать все «крупные» объекты и удалять все «мелкие»). В других случаях такую область пропускания необходимо формировать адаптивно, опираясь на свойства текущего анализируемого изображения или калибровочной группы изображений из обучающей выборки. Для этого удобно использовать *морфологические спектры*.

В рамках предлагаемого подхода морфологические спектры могут быть рассмотрены как опытные статистики специального вида, описывающие наблюдаемое распределение весов структурных примитивов по одному или нескольким параметрам (координатам) вектора $\mathbf{q}\in Q$. Поскольку в этом случае рассматриваются одни параметры (например, масштаб) и отбрасываются другие (например, положение и ориентация), то на одно и то же значение параметра может приходиться несколько голосующих элементов, входящих в анализируемый образ. В результате возникает *спектр по параметру – статистическое распределение, максимумы которого соответствуют преобладающим в данном образе значениям исследуемого параметра*. Если максимумы в спектре полезного сигнала не имеют аналогов в спектре фонового или шумового сигнала, это позволяет адаптивно сформировать область пропускания морфологического фильтра.

Пусть вектор параметров разложения $\mathbf{q}\in Q$ состоит из двух частей:

$$\mathbf{q}=(\mathbf{q}_1,\mathbf{q}_2), \quad (2.24)$$

где $\mathbf{q}_1$ – подвектор *параметров-характеристик*, $\mathbf{q}_2$ – подвектор *параметров локализации*. Разделение параметров на «характеристики» и «локализаторы» является условным и зависит от того, спектры по каким характеристикам мы собираемся исследовать. Например, если речь идет об образующих, характеризующихся сдвигом, масштабом и поворотом, то в одном случае в качестве исследуемых характеристик могут приниматься параметры сдвига, в другом – масштаба, в третьем – поворота. Оставшиеся параметры при этом, естественно, рассматриваются как параметры локализации.

Интегральным спектром (распределением моментов) *порядка* n образа A($\mathbf{p}$) по характеристикам из $\{\mathbf{q}_1 \in Q\}$ будем называть выражение следующего вида:

$$Sp(A(\mathbf{p}),\mathbf{q}_1) = [\boldsymbol{\Sigma}_{(\mathbf{q}_2 \in Q)} |A(\mathbf{q}_1,\mathbf{q}_2)|^n]^{1/n}. \qquad (2.25)$$

В основном на практике используются *массовые спектры* или *проекции* (n=1), *энергетические спектры* (n=2) и *максимальные спектры* (n=∞). В литературе также встречаются спектры, производные от описанных интегральных. Их естественно назвать *дифференциальными спектрами*. Следуя Марагосу (см. [14]) дифференциальные спектры можно определить следующим образом. Пусть q_1 – скаляр (один параметр из $\mathbf{q}$). Тогда дифференциальный спектр по q_1 определяется на основе интегрального спектра (2.22) как его производная:

$$DSp(A(\mathbf{p}),q_1) = d(Sp(A(\mathbf{p}),q_1))/dq_1, \qquad (2.26)$$

где $d(Sp(A(\mathbf{p}),q_1))$ – приращение интергрального спектра $Sp(A(\mathbf{p}),q_1)$, соответствующее минимальному приращению dq_1 параметра q_1 (определяется дискретом аккумулятора $\Theta(Q)$).

<u>*Структурное сравнение изображений*</u>

Задача сравнения изображений традиционно (в пространстве изображений) рассматривается как задача определения меры сходства (корреляции) между изображениями, что необходимо для решения задач взаимной привязки изображений (*matching*). В предыдущих разделах данная задача была подробно рассмотрена как задача сравнения образов в пространствах однородных структур на основе их линейной корреляции. Рассмотрим теперь более сложную задачу сравнения образа с моделью.

Пусть *модель образа* задана в виде однородного разложения (1.9).

Назовем *индикатором структурной связи* характеристическую функцию вида:

$\chi(x)=\{0$, если $x=0$; 1 – в противном случае$\}$.

Определим *характеристический базис* образа B как

$$\mathbf{E}_\chi(B) = \{\chi(b_k) \bullet E_k,\ E_k \in \mathbf{E}\}, \qquad (2.27)$$

где $\mathbf{E}$ – некоторый исходный базис морфологического разложения. Тогда *морфологическую проекцию образа* A *на модель образа* [B] можно определить как

$$\Pr(A,[B]) = \vee_{k=1..n}(a_k \bullet \chi(b_k) \bullet E_k) = \Pr(A,\boldsymbol{E}_\chi(B)). \qquad (2.28)$$

При этом линейное замыкание характеристического базиса $\boldsymbol{E}_{\chi}(B)\subseteq\Omega$ можно рассматривать как полное множество образов, имеющих такую же структуру, как образ B.

Соответствующая операция проектирования векторов разложения будет иметь вид:

$$\mathrm{Pr}(\mathbf{a},[\mathbf{b}]) = \mathrm{Pr}(\{a_k\},[\{b_k\}]) = \{a_k\bullet\chi(b_k)\}. \quad (2.29)$$

Если выполняется следующее естественное условие

$$\forall k\colon a_k\geq b_k \Rightarrow \mu(\mathbf{a})\geq\mu(\mathbf{b}),$$

тогда для любых двух ненулевых векторов из Θ справедливо неравенство:

$$\|\mathbf{Pr}(\mathbf{a},[\mathbf{b}])\| \leq \|\mathbf{a}\|.$$

Это позволяет определить стандартную меру сходства образа с моделью – *нормированный морфологический коэффициент корреляции*:

$$K_{стр}(A,B)= \|\mathbf{Pr}(\mathbf{a},[\mathbf{b}])\| / \|\mathbf{a}\|, \quad (2.30)$$

где $A,B\in\Omega$; $\mathbf{a}=\mathbf{dec}(A),\mathbf{b}=\mathbf{dec}(B)\in\Theta$, со следующими *стандартными свойствами*:

1) $0 \leq K_{стр}(A,B) \leq 1$;
2) $K_{стр}(A,A) = 1$.
3) $K_{стр}(A,B) = 0 \Leftrightarrow \mathrm{Pr}(A,[B]) = \varnothing$.

Коэффициент корреляции (2.30) позволяет сравнивать однородные структуры с точностью до *класса морфологически эквивалентных структур*:

$$\boldsymbol{B}=\{X\in\Omega\colon K_{стр}(X,B)=1\}. \quad (2.31)$$

Однако это отношение эквивалентности асимметрично, что дает возможность различать среди эквивалентных образов «более простые» и «более сложные» структуры. Если

$$K_{стр}(A,B) = 1,\ K_{стр}(B,A) < 1, \quad (2.32)$$

значит «A сложнее B», и соответственно «B проще A».

Таким образом, задача структурного сравнения изображений успешно решается в пространстве морфологических признаков вне зависимости от семантической природы прикладной области.

Заметим, что для решеток в силу условий (1.14)-(1.16) морфологический коэффициент корреляции со стандартными свойствами типа (2.30) может быть определен непосредственно в пространстве изображений:

$$K_{стр}(A,B)= \|\mathrm{Pr}(A,[B])\| / \|A\|, \quad (2.33)$$

где $A,B\subseteq\Omega$.

Часто также возникает задача сравнения образов, инвариантного относительно геометрических преобразований.

Рассмотрим *отношение эквивалентности* вида

$\forall A(\mathbf{p}),B(\mathbf{p})\in\Omega$: $A(\mathbf{p})\sim B(\mathbf{p}) \Leftrightarrow \exists\mathbf{q}'\in Q$: $A(\mathbf{p})=\sigma(\mathbf{q}')(B(\mathbf{p}))$, $\sigma(\mathbf{q}')\in\mathcal{R}$.

Оно определяет соответствующий *класс эквивалентности*

$$\boldsymbol{B}(\mathbf{p})=\{B(\mathbf{p},\mathbf{q})=\sigma(\mathbf{q})(B(\mathbf{p})),\ \mathbf{q}\in Q,\ \sigma(\mathbf{q})\in\mathcal{R}\}. \quad (2.34)$$

Перейдем в пространство разложений $\Theta(Q)$. Поскольку

$A(\mathbf{p})\sim B(\mathbf{p}) \Leftrightarrow \exists\mathbf{q}'\in Q$: $A(\mathbf{p})=\sigma(\mathbf{q}')(B(\mathbf{p})) \Rightarrow A(\mathbf{q})=B(\mathbf{q}+\mathbf{q}')$,

соответствующий класс *геометрически эквивалентных разложений* можно описать как

$$\boldsymbol{B}(\mathbf{q})=\{B(\mathbf{q}+\mathbf{q}'),\ \mathbf{q}'\in Q\}. \quad (2.35)$$

С учетом этого, *нормированный инвариантный коэффициент морфологической корреляции разложений* может быть определен следующим образом:

$$K_{стр}(A(\mathbf{q}),\boldsymbol{B}(\mathbf{q}))=K_{стр}(A(\mathbf{q}),B(\mathbf{q}')): \forall\mathbf{q}\in Q,\ K_{стр}(A(\mathbf{q}),B(\mathbf{q}))\le K_{стр}(A(\mathbf{q}),B(\mathbf{q}')). \quad (2.36)$$

В наиболее общем виде решение задачи построения инвариантного морфологического описания изображения можно описать следующим образом.

Пусть дано некоторое множество *классов яркостно-геометрической эквивалентности изображений* $\mathbf{[B]}=\{[B_k]\}$, каждый из которых замкнут относительно операций линейной комбинации изображений (объединения и умножения на скаляр), а также относительно геометрических преобразований из $\mathcal{R}$.

Если любые два класса из $\mathbf{[B]}$ являются *линейно независимыми*, то есть

$$\forall[B_k],[B_i]\in\mathbf{[B]},\ k\ne i:\ \forall\alpha\in\Psi, A\in[B_k], B\in[B_i]:\ A\ne\alpha B, \quad (2.37)$$

тогда множество $\mathbf{[B]}$ можно назвать *базисом инвариантного структурного описания*, причем гарантируется инвариантность описания относительно яркостно-геометрических преобразований изображения, определяемых умножением изображения на скаляр и применением к изображению геометрических преобразований из $\mathcal{R}$.

Определим операцию *проекции изображения на подпространство* $\boldsymbol{[B]}$, образованное линейным замыканием множества классов $\mathbf{[B]}$, такую что:

$Pr(A,\boldsymbol{[B]})\in\boldsymbol{[B]}$; $Pr(A,\boldsymbol{[B]})=Pr(Pr(A,\boldsymbol{[B]}),\boldsymbol{[B]})$;

$$Pr(\varnothing,\boldsymbol{[B]})=\varnothing;\ Pr(a\bullet A,\boldsymbol{[B]})=a\bullet Pr(A,\boldsymbol{[B]}), \quad (2.38)$$

где $a\in\Psi$; $A,\varnothing\in\Omega$; $\boldsymbol{[B]}\subseteq\Omega$.

Пусть теперь для некоторого базиса инвариантного описания $\mathbf{[E]}$ размерности n, проекция любого образа из Ω на замыкание $\boldsymbol{[E]}$ удовлетворяет *условию разложимости:*

$$Pr(A,\boldsymbol{[E]}) = \vee_{k=1..n}(Pr(A,[E_k])). \quad (2.39)$$

То есть, как и в случае условия (1.9), *проекция образа на базис есть объединение проекций на его элементы*. Построенную таким образом алгебраическую систему $\{\Psi,\Omega,\bullet,\mathbf{V},\mu,Pr,\mathcal{R},\mathbf{[E]}\}$, для которой справедливо условие (2.39), можно назвать *инвариантной проективной морфологией* на Ω. Базис **[E]** является здесь соответственно *базисом инвариантного морфологического разложения.*

Распознавание объектов на изображениях

С точки зрения базовых методов распознавания образов морфологические разложения, которые ставят в соответствие любому изображению A(**p**) массив параметров (аккумулятор) его разложения A(**q**), есть один из регулярных и обоснованных способов перейти от изображения как двумерной функции яркости к вектору (массиву) в пространстве признаков. При этом каждый отдельный признак (элемент разложения) возникает в результате сравнения анализируемого образа (изображения) с одним из эталонных образов (образующих). Далее в этом пространстве могут решаться все стандартные задачи теории распознавания образов: верификация, идентификация, кластеризация и т.п. Для этого могут применяться любые известные методы решения этих задач: метрические (методы «ближних» и «дальних» соседей), аналитические (разбиение гиперповерхностями), вероятностные (байесовские, максимального правдоподобия и т.п.), нейросетевые и любые другие.

Следует, однако, учитывать, что далеко не всегда традиционные процедуры классификации (или кластеризации) векторов признаков на основе метрических расстояний между ними или соответствующих вероятностных мер оказываются адекватными природе анализируемых структурных образов. Между тем, модификации этих процедур, основанные на использовании описанных выше морфологических мер близости, являются существенно более обоснованными с точки зрения семантики анализируемых объектов.

Нетривиальной (с точки зрения специфики анализа изображений) задачей является здесь задача построения *минимального информативного дескриптора*, предполагающая снижение пространства размерности признаков (выбор минимально необходимого числа характерных образующих), описывающих изображения объекта или класса объектов в задаче верификации (проверки принадлежности предъявляемого образа заданному классу). Эта задача тесно связана с задачей обнаружения (локализации) объектов на изображениях – ее всегда приходится решать на этапе обучения алгоритмов обнаружения.

Заключение. В данной работе описана т.н. «проективная морфология», позволяющая формально описывать и сравнивать структурные модели и структурные объекты вне зависимости от конкретной прикладной области. В качестве базового понятия здесь традиционно рассматривается «составной образ», который строится из первичных «примитивов» путем их объединения при помощи некоторой кумулятивной (ассоциативной и коммутативной) операции с определенными весами-скалярами, численно характеризующими «степень связи» образа с его элементом. «Моделью» (шаблоном) 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 признаков стандартным образом определяются удобные нормированные коэффициенты морфологической 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.).

Применительно к структурному анализу цифровых изображений предложен формализм «проективной морфологии изображений», позволяющий с единых позиций описывать и разрабатывать такие базовые аспекты машинного зрения как структурная фильтрация и сжатие изображений, построение структурных спектров, структурное сравнение изображений и моделей. Показано, что проективные морфологии изображений могут быть, в частности, построены на базе таких методов анализа изображений как корреляционный анализ, математическая морфология Серра, морфологический анализа Ю.П. Пытьева, методы голосования, восходящие к преобразованию Хафа, частотные методы (на базе преобразования Фурье, БПФ, ДКП и др.) и пространственно-частотные методы фильтрации (кратномасштабный вейвлет-анализ).

Основным ограничением предложенного подхода является то, что описанный класс «моделей с однородными структурными связями» позволяет задавать лишь состав входящих в модель (образ) элементов и связь между этими образующими и образом в целом, но не рассматривает связи образующих элементов между собой. Это ограничение является принципиальным, так как модели более общего вида, описываемые произвольными гиперграфами, не могут быть однозначно охарактеризованы однородными массивами или векторами признаков.

Литература

1. **Дуда Р., Харт П.** Распознавание образов и анализ сцен: Пер. с анг. – М.: Мир, 1976, с.511.

2. **Ту Дж., Гонсалес Р.** Принципы распознавания образов. – М.: Мир, 1978, с.411.

3. **Горелик А.Г., Гуревич И.Б.**, Скрипкин В.А. Современное состояние и проблемы распознавания. – М.: Радио и связь, 1985, с.160.

4. **Bender E.A.** Mathematical Methods in Artificial Intelligence. – IEEE Comput. Society Press, Los Alamitos, California, 1996, p.636.

5. **Горбань А.Н.**, Обучение нейронных сетей, М.: СП ПараГраф, 1991.

6. **Freund Y., Schapire R.** A Decision-Theoretic Generalization of On-line Learning and an Application to Boosting. Journal of Computer and System Sciences, 55, 1997, p.119-139.

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

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

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

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

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

12. **Ballard D.H.** Generalizing the Hough transform to detect arbitrary shapes. Pattern Recognition, 1981. № 13(1.1). P. 111-122.

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

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

15. **Davies E.R.** Locating objects from their point features using an optimised Hough-like accumulation technique. Pattern Recogn. 1992. № 13(1.1). P.113-121.

16. **Visilter Y.U., Zheltov S., Stepanov A.** Object Detection and Recognition using Events-based Image Analysis. SPIE Proceedings, International archives of photogrammetry and remote sensing. 1996. № 2823. P.184-195.