

ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ

УДК 519.7

В.А. Утробин

ИНФОРМАЦИОННАЯ ЗНАЧИМОСТЬ КОМПОНЕНТ U -ПРЕОБРАЗОВАНИЯ ПРИ РЕШЕНИИ ЗАДАЧИ ОРИЕНТАЦИИ В ПРОСТРАНСТВЕ ДЛЯ СИСТЕМ ТЕХНИЧЕСКОГО ЗРЕНИЯ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Рассматриваются проблемы построения перцептивного пространства для систем технического зрения с позиций теории активного восприятия по результатам анализа произвольного изображения на системе пространственных градиентов.

Ключевые слова: зрительное восприятие, физико-математические интерпретации, пространственные модели.

Введение

В настоящее время одной из главных проблем при разработке систем зрительного восприятия на этапе понимания результатов наблюдения внешней среды является построение перцептивного (внутреннего для наблюдателя) пространства как модели внешнего. Главная цель построения – это ориентация во внешней среде, понимание своего места в этой среде [1].

В соответствии с теорией активного восприятия (ТАВ) [2, 3], изображение M , представленное своими скалярными компонентами $\mu(x, y)$, подвергается интегральному Q -преобразованию по произвольной области определения G_i (либо по всей области определения G изображения)

$$m(G_i) = \iint_{(x,y) \in G_i} \mu(x, y) dx dy, \quad (1)$$

отображающему изображение в бесконечно гладкое многообразие с потенциальной возможностью последующего выявления внутренней геометрии этого многообразия. Поэтому следующим этапом обработки изображения является пространственное дифференцирование, позволяющее выявлять пятнадцать значений градиентов по направлениям пятнадцати векторных полей Киллинга в потенциальном поле относительно центра области определения A_0 , где находится «тяготеющая» масса (визуальная масса) (1):

$$v^i = \frac{d\varphi(A_0)}{dx^i} = \left| \text{grad}_{x^i} \varphi(A_0) \right|, \quad i = \overline{1, 15}. \quad (2)$$

В результате имеем композицию (U -преобразование)

$$u = v \circ w \equiv d \circ \int_G, \quad (3)$$

а также 16-мерный вектор результата разложения $\mu = \{\mu_i\}$ (рис. 1).

Если систему пространственных производных (2) представить в декартовой системе координат, то каждый из блоков предобработки (рис. 1, фильтры-покрытий, маски, F_i) выявляет дифференциальную структуру по одному из направлений в поле потенциальной функции $\varphi(x, y)$

$$\left(\frac{\partial}{\partial x} \left(\frac{\partial^{j-1}}{\partial x^{k-1} \partial y^r} \right), \frac{\partial}{\partial y} \left(\frac{\partial^{j-1}}{\partial x^k \partial y^{r-1}} \right), \frac{\partial^2}{\partial x \partial y} \left(\frac{\partial^{j-r}}{\partial x^{k-1} \partial y^{r-1}} \right) \right), \quad (4)$$

где $k, r = \overline{1,3}, k + r = 6, j = \overline{1,6}$ (обозначение потенциальной функции опущено).

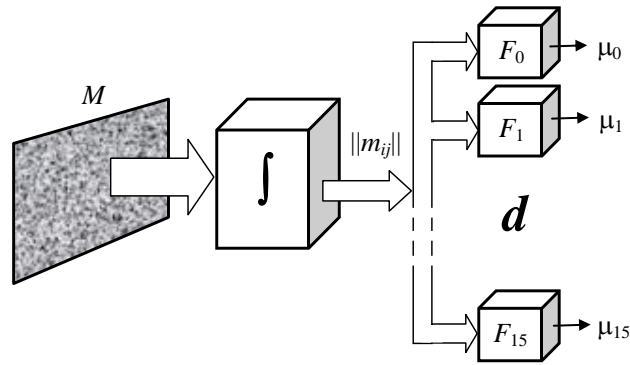


Рис. 1. Система преобработки изображения M

Производные в составе системы (2) – это функции чувствительности, входящие в состав ряда Тейлора при решении задачи активной идентификации (по этой причине соответствующая теория обработки изображений в условиях априорной неопределенности называется теорией активного восприятия).

Ряд Тейлора (относительно центра области определения A_0) имеет следующий вид:

$$\varphi(X) - \varphi(A_0) = d\varphi + \frac{1}{2!} d^2\varphi + \frac{1}{3!} d^3\varphi + \dots, \quad (5)$$

где $d^r\varphi$ – полный дифференциал:

$$d^r\varphi = \sum_{k=0}^r C_r^k \frac{\partial^r}{\partial x^{r-k} \partial y^k} dx^{r-k} dy^k, \quad C_r^k = \frac{r!}{k!(r-k)!}. \quad (6)$$

При разложении произвольного изображения по функциональной схеме рис.1 все компоненты вектора $\mu = \{\mu_i\}$ равнозначны (с точки зрения информационной значимости). При наличии шумов на изображении, либо при анализе относительно сложно организованных сцен компоненты вектора «шумят», т.е. исчезают явные локальные экстремумы в линейчатом спектре разложения. Возникает вопрос, а нельзя ли теоретически обосновать информационную значимость компонент разложения изображения. Результаты теоретических исследований, направленные на решение поставленной проблемы, излагаются в данной работе.

Взвешенное U -преобразование

Множество масок F_i , в соответствии с ТАВ, упорядочены на двумерной решетке $V(x, y)$ (рис. 2, [2, 3]), а по диагональным срезам на системе преобразований (4) принадлежат шести первым полным дифференциалам: $d_1 - d_6$.

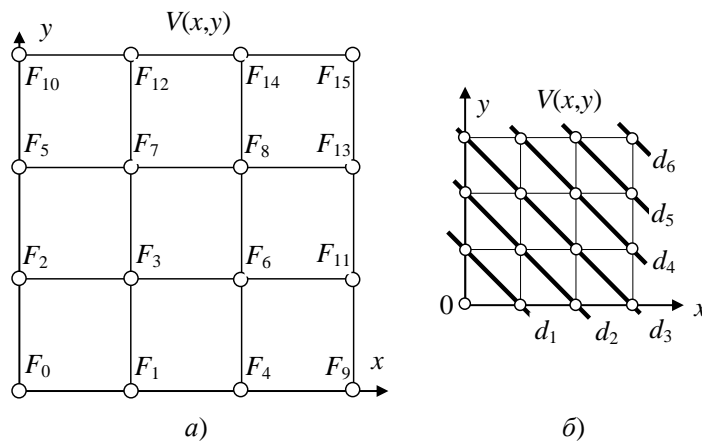


Рис. 2. Двумерная решетка $V(x, y)$ упорядочения масок (а) и срезы d_i на решетке (б)

Три первых полных дифференциала полностью представлены на решетке и реализуются девятью масками – $F_1, F_2, F_4, F_3, F_5, F_9, F_6, F_7, F_{10}$ (нумерация масок по работе [2]):

$$d^1 = \frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy; \quad d^2 = \frac{\partial^2}{\partial x^2} dx^2 + 2 \frac{\partial^2}{\partial x \partial y} dx dy + \frac{\partial^2}{\partial y^2} dy^2;$$

$$d^3 = \frac{\partial^3}{\partial x^3} dx^3 + 3 \frac{\partial^3}{\partial x^2 \partial y} dx^2 dy + 3 \frac{\partial^3}{\partial x \partial y^2} dx dy^2 + \frac{\partial^3}{\partial y^3} dy^3.$$

Три последних полных дифференциала представлены частично и реализуются масками $F_8, F_{11}, F_{12}, F_{13}, F_{14}, F_{15}$:

$$d^4 \approx 6 \frac{\partial^4}{\partial x^2 \partial y^2} dx^2 dy^2 + 4 \frac{\partial^4}{\partial x^3 \partial y} dx^3 dy + 4 \frac{\partial^4}{\partial x \partial y^3} dx dy^3;$$

$$d^5 \approx 10 \frac{\partial^5}{\partial x^3 \partial y^2} dx^3 dy^2 + 10 \frac{\partial^5}{\partial x^2 \partial y^3} dx^2 dy^3; \quad d^6 \approx 20 \frac{\partial^6}{\partial x^3 \partial y^3} dx^3 dy^3.$$

Подставляя выражения полных дифференциалов в (5) с учетом (6), получаем следующий вид ряда Тейлора для потенциальной функции $\varphi(x, y)$ в произвольной точке X области определения относительно центра этой области:

$$\varphi(X) - \varphi(A_0) = \sum_1^3 \mu_i + \frac{1}{2} \sum_4^7 \mu_i + \frac{1}{4} \mu_8 + \frac{1}{6} \sum_9^{12} \mu_i + \frac{1}{12} \sum_{13}^{14} \mu_i + \frac{1}{36} \mu_{15} + 0(\varepsilon). \quad (7)$$

Отсюда следует распределение долевого участия компонент разложения на основе U -преобразования (табл. 1).

Таблица 1

Долевое участие

F_i	1-3	4-7	8	9-12	13-14	15
Кластер	1	2	3	4	5	6
%	49,32	24,66	12,33	8,22	4,11	1,37

При расчете долевого участия компонент вектора разложения принято условие – сумма всех коэффициентов в ряде (7) составляет 100%: $\sum a_i = 2,027(7) \sim 100\%$.

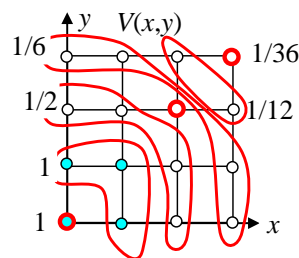


Рис. 3. Кластеры равных долевого участия

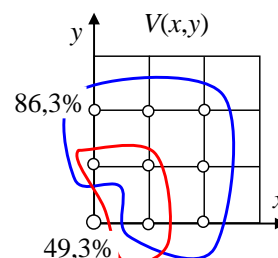


Рис. 4. Максимальный вклад

Таким образом, из анализа ряда Тейлора для U -преобразования на множестве из 15-ти масок (и соответствующих им преобразований из (4)) следует:

1. Существуют подмножества (кластеры), имеющие равный долевого вклад по результатам описания произвольного изображения (табл. 1, рис. 3):

- кластер 1 – $\partial \varphi(x, y) / \partial x, \partial \varphi(x, y) / \partial y; \partial^2 \varphi(x, y) / \partial x \partial y$;
- кластер 2 – $\partial^2 \varphi(x, y) / \partial x^2; \partial^2 \varphi(x, y) / \partial y^2; \partial^3 \varphi(x, y) / \partial x^2 \partial y; \partial^3 \varphi(x, y) / \partial x \partial y^2$;
- кластер 3 – $\partial^4 \varphi(x, y) / \partial x^2 \partial y^2$;
- кластер 4 – $\partial^3 \varphi(x, y) / \partial x^3; \partial^3 \varphi(x, y) / \partial y^3; \partial^4 \varphi(x, y) / \partial x^3 \partial y; \partial^4 \varphi(x, y) / \partial x \partial y^3$;

- кластер 5 – $\partial^5 \varphi(x, y) / \partial x^3 \partial y^2$; $\partial^5 \varphi(x, y) / \partial x^2 \partial y^3$;
- кластер 6 – $\partial^6 \varphi(x, y) / \partial x^3 \partial y^3$.

2. Из табл. 1 видно, что кластер 1 обеспечивает почти 50%-ный вклад, т.е. долевое участие масок данного кластера (а это низкочастотные пространственные компоненты в анализируемом изображении) максимально.

3. Если учесть следующие два кластера, то эти три кластера обеспечивают более 85% информации в общем результате анализа изображения (рис. 4).

4. Оставшиеся три кластера (4, 5, 6), маски которых реализуют относительно высокочастотные преобразования, обеспечивают получение не более 14% информации, относящейся к «малогабаритным» пространственным объектам изображения.

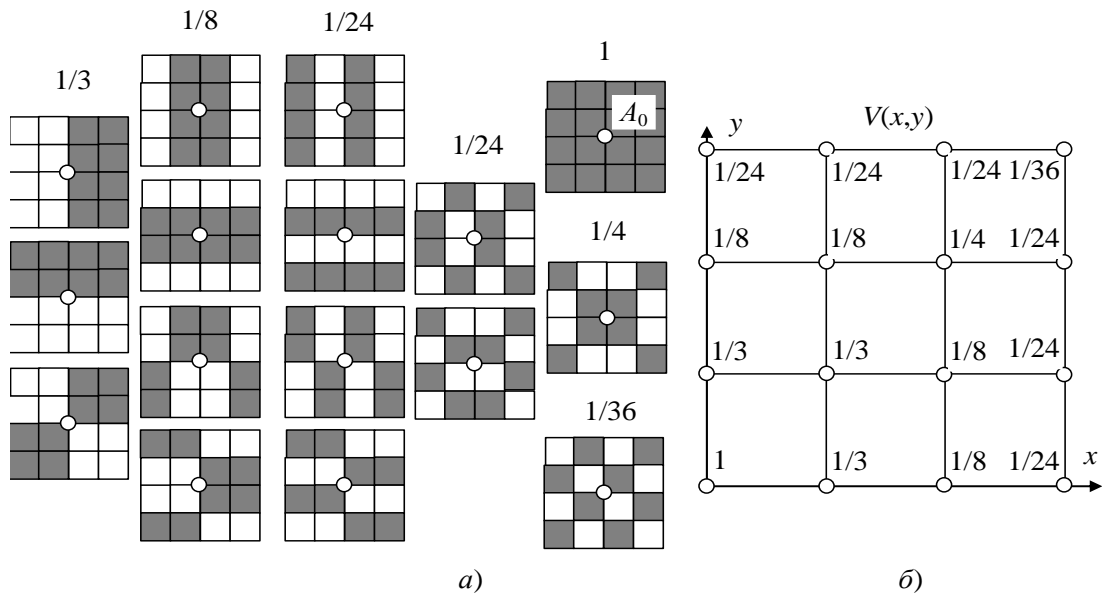


Рис. 5. Поэлементное распределение весовых коэффициентов масок:
 а – внешний вид масок и их веса; б – распределение весов на решетке $V(x,y)$

Каждое преобразование, реализуемое пространственными масками (применяемыми после Q -преобразования), является базисным, т.е. независимым, и одновременно два произвольных преобразования несовместны, т.е. в составе изображения одновременно не может быть образ, соответствующий полному покрытию двух масок, они всегда перекрываются. Поэтому их участие в описании (имеется в виду вероятностная мера) равновероятно и это верно также для масок, входящих в состав кластеров. В результате, учитывая распределения по (7), получаем поэлементное распределение вероятностных долей участия масок, представленное на рис. 5.

На рис. 6 показано изменение вклада масок в зависимости от их принадлежности соответствующим кластерам. На маске F_8 , реализующей преобразование $\partial^4 \varphi(x, y) / \partial x^2 \partial y^2$, наблюдается «всплеск активности».

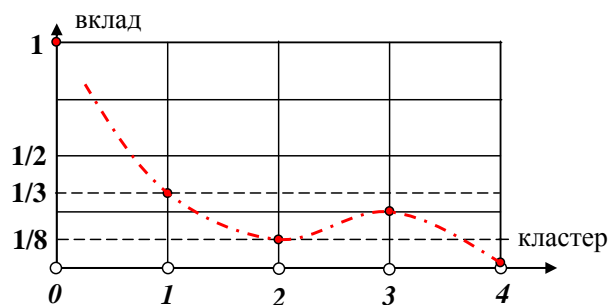


Рис. 6. Изменение вклада масок по их принадлежности

Физико-математические интерпретации

В соответствии с ТАВ, каждой маске F_i соответствует бинарный оператор V_i , аналогично представленный на решетке $V(x, y)$. Для выявления внутренней геометрии многообразия используется касательная в точке A_0 к многообразию, евклидова плоскость, «кусочек» которой с центром A_0 называется планигоном (рис. 5, а).

1. Если маски F_1, F_2 , реализующие преобразования $\partial\varphi(x, y)/\partial x, \partial\varphi(x, y)/\partial y$, выявляют градиентные изменения по направлениям осей x, y декартовой системы координат (и по этой причине максимально значимы при решении задач ориентации в пространстве), то им равнозначна по вкладу маска F_3 , которой соответствует преобразование $\partial^2\varphi(x, y)/\partial x\partial y$, более «специфична».

Каждый фильтр из множества $\{F_i\}$, решая соответствующее дифференциальное уравнение, выявляет интегральную кривую для векторного поля в окрестности точки A_0 . Поэтому F_3 в декартовой системе координат можно поставить в соответствие интегральную кривую вида (рис. 7, а):

$$\mathbf{v} = x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \quad (8)$$

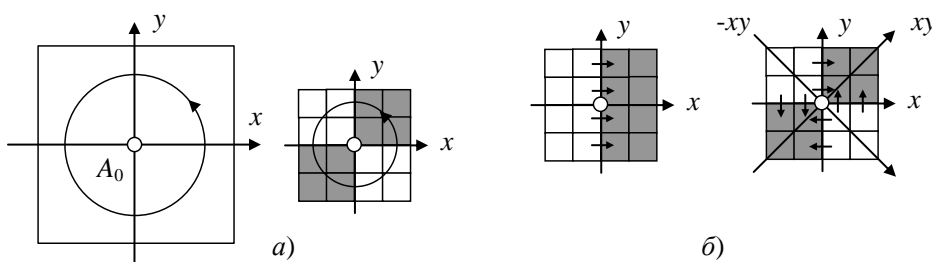


Рис. 7. Интерпретация элементов кластера I:

а – циркуляция для маски F_3 ; б – направления изменения потенциала для F_1 и F_3

Вектору (8) соответствует вектор Киллинга $\mathbf{e}_z = \frac{\partial}{\partial \varphi}$ в сферической системе координат,

где φ – угол поворота вокруг оси z . Следовательно, фильтру F_3 с его преобразованием $\frac{\partial^2}{\partial x\partial y}$ соответствует потенциальное течение циркуляции с особенностью в точке A_0 . В этой точке вдоль оси z существует изолированный вихрь конечной интенсивности с бесконечной угловой скоростью ω .

С другой стороны, если рассматривать направление градиента относительно линии раздела подобластей, например, маски F_1 , то изменение потенциала происходит из подобласти -1 координаты x в $+1$ подобласть. Для F_3 аналогичные изменения наблюдаются для подобластей координат $-xy$ и xy , т.е. подобласти значимости данного фильтра принадлежат диагональной системе координат (рис. 7, б). Следовательно, данная маска (и оператор) отражает преобразование градиента в новой системе координат $-xy, xy$, и ему соответствует преобразование $\frac{\partial}{\partial z_w}$, где направление z_w есть направление xy , точнее, $x = y$.

Таким образом, кластер I, состоящий из элементов F_1, F_2, F_3 , (и им соответствующих операторов V_1, V_2, V_3), – это кластер «градиентных изменений». При этом оператор V_3 (маска F_3) двойственен:

- в декартовой системе координат ему соответствует потенциальное течение циркуляции с особенностью в центре планигона;
- в новой (диагональной) системе координат линейное потенциальное течение обеспечено градиентом. С позиций проективной геометрии – это направление в несобственную точку проективной плоскости (планигон, как картинного пространства), (рис. 8).

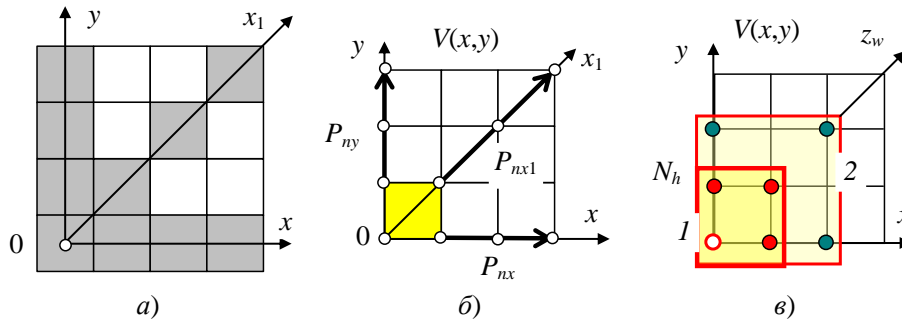


Рис. 8. Координатные направления перцептивного пространства E^3 :
a – на планигоне; *б* – решетке; *в* – двуслойном планигоне (*1, 2* – слои)

Изложенное обосновывает информационную значимость элементов кластера *I*, которые образуют тройку операторов (V_1, V_2, V_3) , являющуюся алгебраической группой (в ТАВ такие группы называются полными). Из вершин графа группы (на рис. 8, б кластер выделен) по направлениям *x, y, x₁* «растут» три орграфа соответствующих полных групп.

2. Рассмотрим тройку (V_4, V_5, V_8) , образующую полную группу. Если первая полная группа (V_1, V_2, V_3) на решетке принадлежит первому срезу (слою) в глубину z_w перцептивного пространства по направлению x_1 (рис. 8, или xu по рис. 7), то вторая – (V_4, V_5, V_8) принадлежит следующему слою по направлению глубины (рис. 8, в).

Если первая группа на системе преобразований – группа градиента, то вторая – группа дивергенции. Действительно.

Уравнение Лапласа для скалярной функции

$$\nabla^2 \varphi = \frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial y^2} + \frac{\partial^2 \varphi}{\partial z_w^2}.$$

Если первым двум компонентам в этом уравнении соответствуют операторы V_4, V_5 , то третьей компоненте – оператор V_8 с преобразованием

$$\frac{\partial^4}{\partial x^2 \partial y^2} \equiv \frac{\partial^2}{\partial z_w^2} \left(\frac{\partial^2}{\partial z_w^2} \right) \equiv \frac{\partial^2}{\partial z^2}$$

в системе координат внешнего наблюдателя (рис. 9).

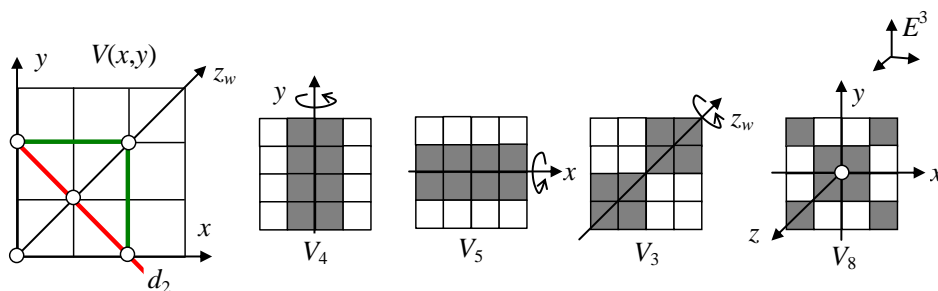


Рис. 9. Интерпретации компонент второго дифференциала и оператора V_8

Поскольку в пространстве определены преобразования градиента и дивергенции, то перцептивное пространство внешнего наблюдателя, представленное на планигоне (рис. 8, а), или представленное через преобразования на решетке $V(x, y)$, – это пространство скалярного потенциала.

Для определения векторного поля полем скалярного потенциала в трехмерном (внешнем) пространстве достаточно существования первого полного дифференциала

$$d\varphi = \frac{\partial\varphi}{\partial x} dx + \frac{\partial\varphi}{\partial y} dy + \frac{\partial\varphi}{\partial z} dz,$$

где компоненты соответствуют преобразованиям, реализуемыми масками F_1, F_2, F_3 . Отсюда (снова) следует информационная значимость элементов кластера I .

3. В состав третьего полного дифференциала входят операторы V_6, V_7, V_9, V_{10} с преобразованиями. Компоненты этого дифференциала входят в состав уравнения Пуассона вектора угловой скорости ω как функции векторного потенциала \mathbf{a} : $\omega = \frac{1}{2} \text{rot rot } \mathbf{a}$ или

$$\begin{cases} \nabla^2 a_x = -2\omega_x; \\ \nabla^2 a_y = -2\omega_y; \\ \nabla^2 a_z = -2\omega_z. \end{cases}$$

С другой стороны, для кругового потенциального течения с изолированным вихрем данные уравнения следует рассматривать независимо друг от друга, т.е. как три варианта ориентированных плоскостей со своими изолированными вихрями (рис. 10). В этом случае допустима следующая система уравнений:

$$\begin{cases} \frac{\partial^3}{\partial x^3} + \frac{\partial^3}{\partial x \partial y^2} + \frac{\partial^3}{\partial x \partial z^2} = -2\omega_x; \\ \frac{\partial^3}{\partial y^3} + \frac{\partial^3}{\partial x^2 \partial y} + \frac{\partial^3}{\partial y \partial z^2} = -2\omega_y; \\ \frac{\partial^3}{\partial z^3} + \frac{\partial^3}{\partial z \partial x^2} + \frac{\partial^3}{\partial z \partial y^2} = -2\omega_z. \end{cases}$$

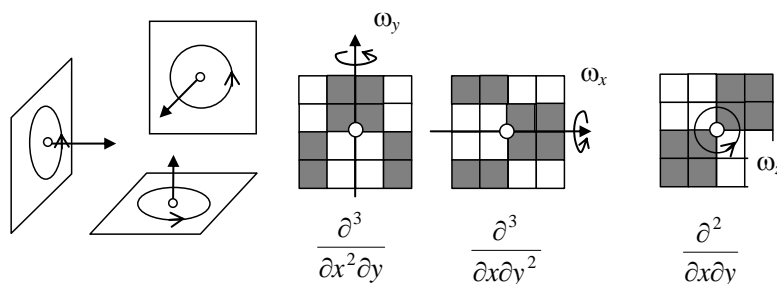


Рис. 10. Три вида изолированных вихрей

Уравнение Пуассона выполняется при условиях, что поле является вихревым, а градиент и дивергенция поля равны нулю, т.е. в системе декартовых координат в разложении отсутствуют компоненты, соответствующие операторам $(V_1, V_2), (V_4, V_5, V_8)$. В этом случае имеем следующее соответствие между компонентами уравнения Пуассона и операторами:

$$\left(\frac{\partial^3}{\partial x^3}; \frac{\partial^3}{\partial x \partial y^2} \right) \Rightarrow (V_9; V_7); \quad \left(\frac{\partial^3}{\partial y^3}; \frac{\partial^3}{\partial x^2 \partial y} \right) \Rightarrow (V_{10}; V_6).$$

Итак, третьему полному дифференциалу с операторами (V_6, V_7, V_9, V_{10}) соответствует уравнение Пуассона. Поэтому это пространство, представленное через соответствующие преобразования на решетке операторов, разделяют пространства скалярного и векторного потенциалом. Первому соответствует пространство внешнего наблюдателя, второму – внутреннего.

Так как $\text{rot} \text{rota} = \text{grad} \text{div} \mathbf{a} - \nabla^2 \mathbf{a}$, то операторы V_9, V_{10} допускают следующие две интерпретации (двойственность):

- компоненты $\text{grad} \text{div} \mathbf{v}$ поля скалярного потенциала φ , где $\mathbf{v} = \text{grad} \varphi$;
- компоненты $\text{rot} \text{rota}$ поля векторного потенциала \mathbf{a} (при дополнительном условии).

В этом случае операторы V_9, V_{10} в составе полных групп $(V_1, V_4, V_9), (V_2, V_5, V_{10})$ внешнего наблюдателя «выполняют» роль в составе поля скалярного потенциала, а в составе полной группы (V_9, V_{10}, V_{15}) – поля векторного потенциала.

Таковы физико-математические интерпретации для задачи понимания масок (операторов), наделенных свойством максимального вклада в анализ произвольного изображения при построении модели «внутреннего мира» – перцептивного пространства внешнего относительно планигона наблюдателя.

Заключение

В работе с позиций анализа ряда Тейлора в полных дифференциалах, используемого при активной идентификации объекта управления, доказываемся возможность информационного обоснования компонент U -преобразования.

Показано следующее:

1. Компоненты U -преобразования, упорядоченные на двумерной решетке реализуемых преобразований, входят в состав шести полных дифференциалов.
2. Компоненты U -преобразования по значимости (весовому коэффициенту) образуют кластеры равных весовых значений и упорядочиваются по значимости.
3. Три первых кластера, упорядоченных по значимости, обеспечивают максимальный вклад в результаты анализа, три последних – высокочастотных менее значимы.
4. Кластеру с максимальным весовым коэффициентом соответствуют три низкочастотные составляющие преобразования, отвечающие за ориентацию в пространстве – построение внутренней системы координат перцептивного пространства.
5. Преобразования, соответствующие второму полному дифференциалу и восьмой компоненте U -преобразования, обеспечивают формирование направления глубины в перцептивном пространстве.
6. Восьмая компонента U -преобразования – единственный элемент своего компакта – по значимости (за исключением элементов первого кластера) максимальна (рис. 6) и отвечает за связь внутренней системы координат с внешней, позволяя тем самым определить место наблюдателя во внешней среде.
7. Преобразования, входящие в состав третьего полного дифференциала, принадлежат уравнению Пуассона и выступают в роли «пространства, разделяющего внешний и внутренний мир» наблюдателя. С позиций физики – разделение сред скалярного и векторного потенциалов.

Библиографический список

1. **Утробин, В.А.** Компьютерная обработка изображений: Информационные модели этапа понимания / В.А. Утробин. – Н. Новгород: НГТУ, 2006. – 247 с.

2. **Утробин, В.А.** Компьютерная обработка изображений. Анализ и синтез / В.А. Утробин. – Н. Новгород: НГТУ, 2013. – 246 с.
3. **Утробин, В.А.** Физические интерпретации элементов алгебры изображения // УФН. 2004. Т.174. №10. С. 1089–1104.

*Дата поступления
в редакцию 15.07.2013*

V.A. Utrobin

SIGNIFICANCE OF INFORMATION OF *U*-TRANSFORM COMPONENTS IN SOLVING THE PROBLEM OF ORIENTATION IN SPACE FOR VISION SYSTEMS

Nyzhny Novgorod state technical university n.a. R.E. Alexeev

Purpose: Building a perceptual (internal observer) space as an external model based on video surveillance.

Method: From the perspective of the theory of active perception, involving the physical-mathematical apparatus investigated the spectral decomposition of the newsworthiness of a priori uncertain picture for description and understanding of the observed images.

The result / area of application: Using of physical and mathematical interpretations of supervised classifications, to prove the existence of change inherent in a system of visual perception of human acceptance of *pâtimernomu* space. The result is the ability to assess the depth of *tencial'naâ* even from the vision of one eye (*monoste-rivospriâtîâ*). Therefore, the system of visual perception uses, with positions on the observed image *nimaniâ*-physical concepts of scalar and vector capacities. Research results would have important applications in machine vision systems.

Conclusions: The studies proved theoretically informational value low-frequency spatial components of the expansion for the task orientation in the environment, assess the position of the observer in this environment, and in-depth assessments of perceptual space.

Key words: visual perception, physical-mathematical interpretation, spatial models.