

УДК 681.518.9; 621.384.3

**С.С. Анциферов**

Московский государственный технический университет МИРЭА, г. Москва  
Россия, 119454, г. Москва, пр. Вернадского, 78

## Метрологические основы формирования информационного пространства образов в интеллектуальных системах обработки

**S.S. Antsyferov**

Moscow State Technical University MIREA, c. Moscow  
Russia, 119454, c. Moscow, Vernadsky ave., 78

## *Metrological Bases for Formation of Image Information Space in Intellectual Systems of Processing*

**С.С. Анциферов**

Московський державний технічний університет МІРЕА, м. Москва  
Росія, 119454, м. Москва, пр. Вернадського, 78

## Метрологічні основи формування інформаційного простору образів в інтелектуальних системах обробки

Проблема обеспечения высокого уровня достоверности обработки информации и распознавания образов пространственно-временных полей является одной из важнейших при создании интеллектуальных систем, т.е. систем, осуществляющих обработку на уровне «понимания» информации. В статье излагаются метрологические основы формирования информационного пространства образов, обеспечивающего более высокий уровень достоверности обработки информации с помощью интеллектуальных систем.

**Ключевые слова:** информационная сложность, идентификационный признак, мера информативности, распознавание образов, шкала, селекция признаков.

The problem of high level of assurance in information processing and images recognition of existential fields is one of the major at creating intellectual systems, i.e. the systems, which are carrying out the processing at level of “understanding” of information. In the article, metrological bases of formation of image information space for higher level of information processing assurance by means of intellectual systems are stated.

**Key words:** information complexity, identification sign, informational content measure, image recognition, scale, selection of signs.

Проблема гарантування високого рівня достовірності обробки інформації та розпізнавання образів просторово-часових полів є однією з найважливіших при створенні інтелектуальних систем, що здійснюють обробку на рівні «розуміння інформації». У статті викладаються метрологічні основи формування інформаційного простору образів, що забезпечує більш високий рівень достовірності інформації за допомогою інтелектуальних систем.

**Ключові слова:** інформаційна складність, ідентифікаційна ознака, міра інформативності, розпізнавання образів, шкала, селекція ознак.

## Введение

В настоящее время специалистами самого различного профиля огромное внимание уделяется проблеме создания интеллектуальных систем, т.е. систем, осуществляющих обработку информации на уровне ее «понимания» с широким использованием

накопленной базы знаний. Актуальность данной проблемы обуславливается как возрастающей информационной сложностью исследуемых явлений и процессов, так и все более нарастающей в целом ряде случаев интенсивностью обрабатываемых информационных потоков. Рост требований к точности (достоверности) результатов обработки только повышает актуальность указанной проблемы. Успешное решение данной проблемы связано не только с определением принципов построения и закономерностей функционирования интеллектуальных систем (ИС), чему уже был посвящен ряд работ [1-9], но и с решением не менее важной задачи – разработкой методологии их метрологического обеспечения. Так, в работах [1-3], [5], [6], [8], [9] были обозначены направления повышения интеллектуального потенциала адаптивных информационно-распознающих систем (АИРС), реализующих методологию структурно-стохастической аппроксимации информации пространственно-временных полей и процесса их обработки, определены принципы адаптации параметров обрабатываемых алгоритмов, в частности путем построения эволюционного процесса поиска оптимальной модели адаптивной обработки в рамках объединения символистского и коннективистского направлений, т.е. объединения логической и нейронной парадигм. В работах [4], [7] были рассмотрены системные принципы метрологического обеспечения ИС, в частности такие метрологические показатели, как скорость адаптации, достоверность и устойчивость результатов в процессе динамических наблюдений. Однако выход АИРС на новый, более высокий, уровень сложности переводит их в разряд развивающихся систем обработки, возможно, с распределенной многокомпонентной и многосвязной структурой, что, в свою очередь, требует при разработке таких систем не только учета закономерностей их построения, функционирования, развития и осуществимости, но и определения адекватных оценок характеристик этих закономерностей (совокупностей системных параметров) путем формирования соответствующих форм представления основных шкал измерения параметров как объективных критериев оценки качества ИС.

Понятие информационной сложности является одним из основных в теории информационного подхода к анализу систем [10]. Смысл этого понятия состоит в том, что любое распределение информации на фоне наложенных на нее логических связей должно обладать определенным содержанием. При анализе тех или иных ситуаций принято говорить о том, что они имеют больший или меньший смысл с точки зрения определенных целей. Тем самым определяется измеримость содержания, смысла ситуации. Концепция информационного поля позволяет найти количественную оценку содержания, смысла на основе прослеживания путей реализации логических связей.

Под информационной сложностью  $C$  пространственно-временного поля (ПВП), поступающего на вход ИС, будем понимать логическое пересечение измеримых интенсивностно-пространственных (геометрических) свойств элементарных структурных компонентов (ЭК) поля –  $S1$  и свойств взаимообусловленности (взаимосвязанности) параметров ЭК между собой, формирующих образы поля, адекватно отображающие состояния исследуемых объектов –  $S2$ :

$$C = S1 \cap S2. \quad (1)$$

Каждому ЭК поставим в соответствие совокупность наблюдаемых или измеримых значений идентификационных признаков, отображающих свойства  $S1$  и будем считать, что для распознавания состояний исследуемого объекта в режиме динамических наблюдений в общем случае в каждый конкретный момент времени используется вектор семантических или количественно измеренных значений признаков

$$B_k = \{\beta_{jk}\} = \beta_{1k}, \beta_{2k}, \dots, \beta_{rk}, \quad j = \overline{1, r}, \quad k = \overline{1, K}. \quad (2)$$

В свою очередь, каждому  $j$ -му признаку поставим в соответствие меру его информативности на  $k$ -й реализации поля:

$$\beta_{jk} \rightarrow \mu_{jk}, \quad j = \overline{1, r}, \quad k = \overline{1, K}. \quad (3)$$

В таком случае многомерный вектор  $M = \{\mu_{jk}\}_{\forall_{jk}}$  образует информационное пространство идентификационных признаков. Мера  $\mu_{jk}$  используется в решающем правиле как весовой коэффициент  $j$ -го признака, что существенно повышает достоверность процедуры распознавания. Вместе с тем не все признаки обладают достаточно высоким значением  $\mu$ , что приводит к распознаванию образов поля с высоким уровнем риска.

**Целью данной работы** является разработка шкал метрологической селекции идентификационных признаков по мерам их информативности, обеспечивающих минимизацию риска распознавания образов ПВП.

## Шкала информационной сложности пространственно-временного поля

В качестве оптимального правила оценки значений мер информативности, учитывающего существование некоторого закона информативной связи между значениями признака в различные моменты времени, будем использовать определение апостериорной плотности вероятности меры в  $k$ -й момент времени при условии, что известны ее оценки во всех предыдущих  $(k-1)$ -х тактах и значения признаков  $\beta_{j1}, \dots, \beta_{jk}$ :

$$\hat{\mu}_{jk} \equiv \tilde{p}(\hat{\mu}_{jk}) = p(\hat{\mu}_{jk} | \hat{\mu}_{j1}, \dots, \hat{\mu}_{jk-1}; \beta_{j1}, \dots, \beta_{jk}). \quad (4)$$

Это соотношение служит основой для построения шкалы мер информативности идентификационных признаков, а следовательно и интенсивностно-пространственных свойств компонентов поля. Однако, практическое использование этого соотношения проблематично из-за необходимости располагать довольно обширной априорной информацией, а также необходимости вычисления многомерных интегралов. Достигнуть упрощения задачи построения шкалы  $\mu_{jk}$  можно, введя некоторые предположения относительно статистики функции  $\mu$ . Будем полагать, что информативная связь между последовательными значениями  $j$ -го признака носит вероятностный характер и аппроксимируется марковским процессом, т.е. закон информативной связи – стохастический марковский процесс. В этом случае, при аппроксимации процессом первого порядка.

$$p(\hat{\mu}_{jk}) = \int_M \tilde{p}(\hat{\mu}_{jk-1}) p(\hat{\mu}_{jk} | \hat{\mu}_{jk-1}) d\hat{\mu}_{jk-1}; \quad (5)$$

$$\hat{\mu}_{jk} \equiv \tilde{p}(\hat{\mu}_{jk}) = \frac{p(\hat{\mu}_{jk}) p(\hat{\mu}_{jk} | \beta_{jk})}{\int_M p(\hat{\mu}_{jk}) p(\hat{\mu}_{jk} | \beta_{jk}) d\hat{\mu}_{jk}}. \quad (6)$$

В соотношениях (5) и (6)  $p(\hat{\mu}_{jk})$  имеет смысл априорной плотности вероятности,  $p(\hat{\mu}_{jk} | \beta_{jk})$  – условной плотности вероятности меры информативности  $j$ -го идентификационного признака на  $k$ -м такте обработки, а  $p(\hat{\mu}_{jk} | \hat{\mu}_{jk-1})$  – плотности вероятности перехода от меры информативности в  $(k-1)$ -й момент времени к мере информативности в  $k$ -й момент. Соотношения (5) и (6) являются математической моделью

шкалы мер информативности каждого из идентификационных признаков, обеспечивающей возможность построения модели стохастического информационного пространства элементарных компонентов, отражающих свойства  $S1$  пространственно-временных полей.

Вид функции  $p(\hat{\mu}_{jk} | \beta_{jk})$ , являющейся одним из основных элементов стохастического информационного пространства, должен соответствовать наихудшему виду функции распределения для каждого из идентификационных признаков, т.е. виду приводящему к образованию зон взаимного перекрытия, размытию границ между образами реализаций полей. Согласно информационной теории идентификации, а также методам стохастического синтеза информационных систем наихудшая плотность распределения значений  $j$ -го признака имеет вид функции Лапласа с максимумом этой функции на границе раздела образов (рис. 1).

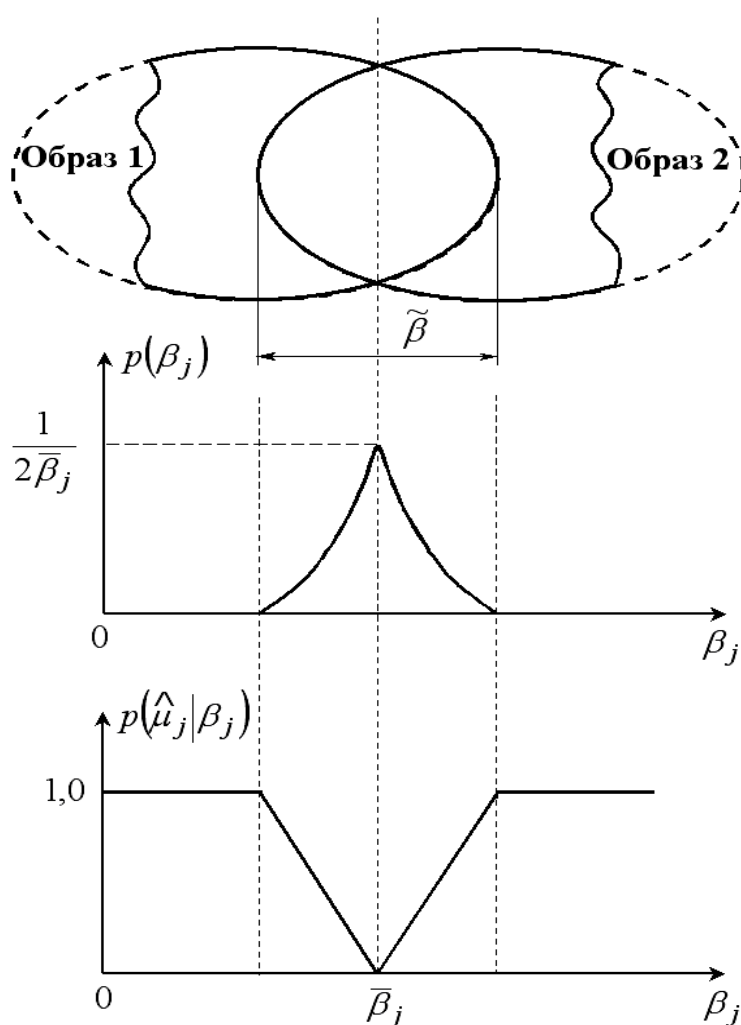


Рисунок 1 – Функции распределения и отображения  $j$ -го идентификационного признака в информационное пространство ИС

Используя аналогичный математический аппарат, можно построить модель и соответствующую ей шкалу, отражающую свойства взаимосвязанности идентификационных признаков ( $S2$ ). Каждая связанная совокупность признаков представляет собой

образ реализации поля, адекватный состоянию исследуемого объекта. При установлении полных вероятностных связей шкала информационной сложности будет выглядеть на  $k$ -м такте обработки, как показано на рис. 2.

Сформированная шкала информационной сложности обеспечивает возможность управления такой важной метрологической характеристикой ИС, как достоверность получаемых результатов. Очевидным подходом к повышению достоверности является информационное «взвешивание» значений идентификационных признаков. Однако, как показывает практика [1-9], простое «взвешивание» не всегда оказывается достаточно эффективным. Поэтому целесообразным представляется выполнение следующего шага, связанного с селекцией признаков по мерам их информативности на шкале информационной сложности.

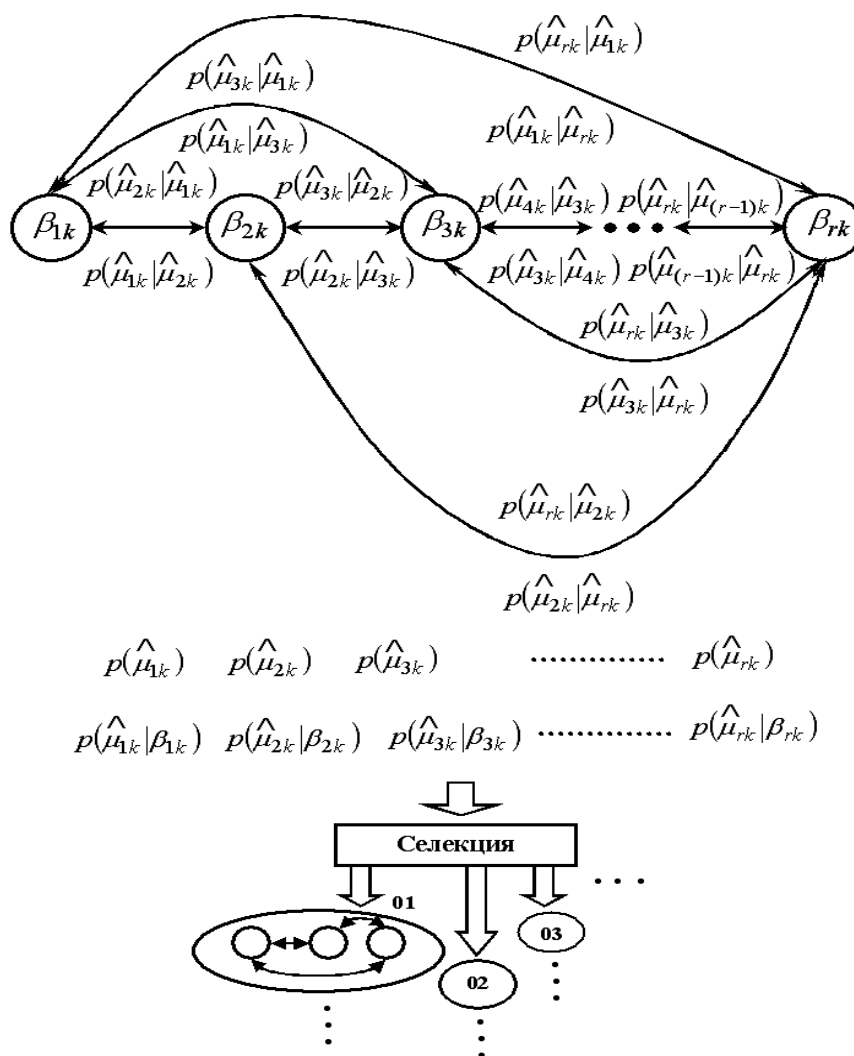


Рисунок 2 – Шкала информационной сложности ПВП

### Механизм селекции признаков по мерам информативности

При разработке механизма селекции будем исходить из того, что мера информативности  $\hat{\mu}_{jk}$  представляет собой дискретную случайную величину, использование

которой либо приводит к увеличению достоверности (обозначим данную ситуацию как  $\mu_1$ ), либо не приводит ( $\mu_2$ ), т.е.  $\mu_1$  – множество значений  $\hat{\mu}_{jk}$ , приводящие к увеличению достоверности, а  $\mu_2$  – множество значений  $\hat{\mu}_{jk}$ , как минимум не приводящее к увеличению достоверности. Если  $p(\hat{\mu}_{jk} \in \mu_1) = P$ , то  $p(\hat{\mu}_{jk} \in \mu_2) = 1 - P$ . Процесс селекции состоит в принятии по каждому признаку одного из двух решений:  $d_1$  – использовать или  $d_2$  – не использовать его в последующей обработке. Каждое решение сопровождается некоторыми потерями, характеризующимися функцией потерь  $G(\mu_i, d_l), i = 1, 2, l = 1, 2$ .

Математическое ожидание функции потерь при фиксированном  $d_l$  представляет собой риск:

$$R(P, d_l) = G(\mu_1, d_l)P + G(\mu_2, d_l)(1 - P). \quad (7)$$

Минимальное по  $l$  значение риска называют байесовским (оптимальным) риском:  $\min R(P, d_l)$ . Решение  $d$ , минимизирующее риск, называется байесовским. В простейшем случае

$$d_B = \begin{cases} d_2, & 0 \leq P \leq 0,5, \\ d_1, & 0,5 \leq P \leq 1. \end{cases} \quad (8)$$

Однако, как показывает практика, такой способ селекции сопровождается большим риском. Так, функция потерь, адекватная решающему правилу (8), может иметь, например, такой вид.

$$G(\mu_i, d_l) = \begin{cases} -102, & \mu_1 \wedge d_1, \\ -100, & \mu_1 \wedge d_2, \\ -98, & \mu_2 \wedge d_1, \\ -100, & \mu_2 \wedge d_2. \end{cases} \quad (9)$$

В этом случае (рис. 3)

$$r_0 = \min_l R(P, d_l) = \min_l \{-98 - 4P, -100\} = \begin{cases} -100, & 0 \leq P \leq 0,5, \\ -98 - 4P, & 0,5 \leq P \leq 1. \end{cases} \quad (10)$$

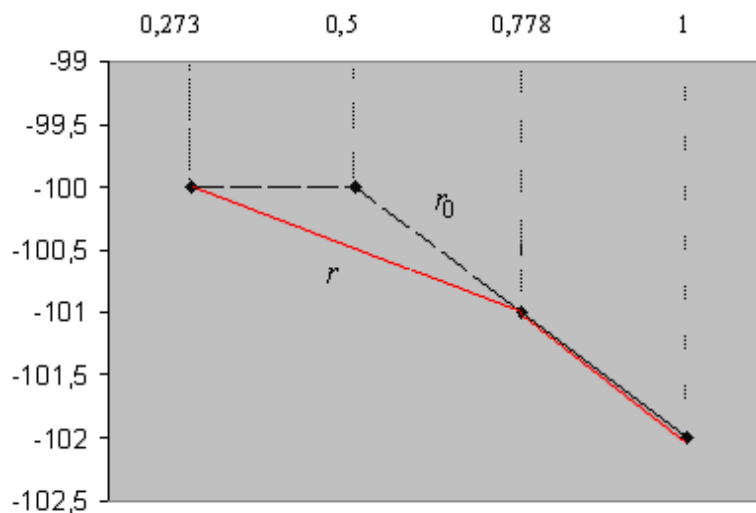


Рисунок 3 – Функция (шкала) риска

Чтобы его снизить, введем в рассмотрение условное распределение  $p(\beta_h|\mu_i)$ , представляющее собой функцию правдоподобия. Тогда апостериорные распределения для мер информативности будут иметь следующий вид:

$$\tilde{p}_1(\hat{\mu}_{jk}) = \frac{p(\beta_h|\mu_1)P}{p(\beta_h)}, \quad (11)$$

$$\tilde{p}_2(\hat{\mu}_{jk}) = \frac{p(\beta_h|\mu_2)(1-P)}{p(\beta_h)}, \quad (12)$$

где  $\beta_h, h=1,2$ ,  $\beta_1$  – множество значений признаков, соответствующих повышению достоверности;  $\beta_2$  – множество значений признаков, не соответствующих повышению достоверности:

$$p(\beta_h) = p(\beta_h|\mu_1)P + p(\beta_h|\mu_2)(1-P). \quad (13)$$

Типичные значения функции  $p(\beta_h|\mu_i)$  согласно экспертным оценкам:

$$p(\beta_h|\mu_i) = \begin{cases} 0,8, & \beta_1 \wedge \mu_1, \\ 0,3, & \beta_1 \wedge \mu_2, \\ 0,2, & \beta_2 \wedge \mu_1, \\ 0,7, & \beta_2 \wedge \mu_2. \end{cases} \quad (14)$$

Тогда

$$p(\beta_h|\mu_i)P = \begin{cases} 0,8, & \beta_1 \wedge \mu_1, \\ 0,3(1-P), & \beta_1 \wedge \mu_2, \\ 0,2P, & \beta_2 \wedge \mu_1, \\ 0,7(1-P), & \beta_2 \wedge \mu_2. \end{cases} \quad (15)$$

$$p(\beta_h) = \begin{cases} 0,3 + 0,5P, & h = 1, \\ 0,7 - 0,5P, & h = 2. \end{cases} \quad (16)$$

$$\tilde{p}_1(\hat{\mu}_{jk}) = \frac{0,8P}{0,3 + 0,5P}, \quad h = 1, \mu_1,$$

$$1 - \tilde{p}_1(\hat{\mu}_{jk}) = \frac{0,3(1-P)}{0,3 + 0,5P}, \quad h = 1, \mu_2, \quad (17)$$

$$\tilde{p}_2(\hat{\mu}_{jk}) = \frac{0,2P}{0,7 - 0,5P}, \quad h = 2, \mu_1,$$

$$1 - \tilde{p}_2(\hat{\mu}_{jk}) = \frac{0,7(1-P)}{0,7 - 0,5P}, \quad h = 2, \mu_2.$$

Условные байесовские риски:

$$r_1 = \min_l R \left[ \tilde{p}_1(\hat{\mu}_{jk},) d_l \right] = \begin{cases} -100, & 0 \leq \tilde{p}_1(\hat{\mu}_{jk}) \leq 0,5, \\ -98 - 4\tilde{p}_1(\hat{\mu}_{jk}), & 0,5 \leq \tilde{p}_1(\hat{\mu}_{jk}) \leq 1. \end{cases} \quad (18)$$

$$r_2 = \min_l R \left[ \tilde{p}_2(\hat{\mu}_{jk},) d_l \right] = \begin{cases} -100, & 0 \leq \tilde{p}_2(\hat{\mu}_{jk}) \leq 0,5, \\ -98 - 4\tilde{p}_2(\hat{\mu}_{jk}), & 0,5 \leq \tilde{p}_2(\hat{\mu}_{jk}) \leq 1. \end{cases} \quad (19)$$

Или

$$r_1 = \begin{cases} -100, & 0 \leq P \leq 0,273, \\ -\frac{29,4 + 52,2P}{0,3 + 0,5P}, & 0,273 \leq P \leq 1. \end{cases} \quad (20)$$

$$r_2 = \begin{cases} -100, & 0 \leq P \leq 0,778, \\ -\frac{68,6 + 48,2P}{0,7 - 0,5P}, & 0,778 \leq P \leq 1. \end{cases} \quad (21)$$

Усредняя условные байесовские риски с вероятностями  $p(\beta_1) = 0,3 + 0,5P$  и  $p(\beta_2) = 0,7 - 0,5P$ , получим безусловный байесовский риск (рис. 3):

$$\begin{aligned} r &= p(\beta_1)r_1 + p(\beta_2)r_2 = (0,3 + 0,5P) \begin{cases} -100, & 0 \leq p \leq 0,273, \\ \frac{29,4 + 52,2P}{0,3 + 0,5P}, & 0,273 \leq p \leq 1 \end{cases} + \\ &+ (0,7 - 0,5P) \begin{cases} -100, & 0 \leq p \leq 0,778, \\ -\frac{68,6 + 48,2P}{0,7 - 0,5P}, & 0,778 \leq p \leq 1 \end{cases} = \\ &= \begin{cases} -100 & 0 \leq P \leq 0,273, \\ -99,4 - 2,2P, & 0,273 \leq P \leq 0,778, \\ -98 - 4P, & 0,778 \leq P \leq 1. \end{cases} \quad (22) \end{aligned}$$

Как видно из рис. 3, к мерам информативности признаков предъявляются вполне «разумные» требования, – порог селекции имеет более низкое значение (0,273) по сравнению с эмпирически задаваемым (0,5), что обеспечивает более высокий уровень достоверности разделения (рис. 2) информационного пространства образов на множество опорных образов пространственно-временных полей.

## Выводы

Как следует из проделанного анализа, сформированные шкалы информационной сложности и риска дают возможность повышать достоверность результатов обработки как за счет информационного «взвешивания» значений признаков, так и их селекции по мерам информативности.

## Литература

1. Анцыферов С.С. Формирование спектра тепловых изображений и распознавание их образов / С.С. Анцыферов // Оптический журнал. – 1999. – Т. 66, № 2. – С. 44-48.
2. Antsyferov S.S. Forming the spectrum of thermal images of objects and recognizing their patterns / S.S. Antsyferov // J.Opt. Technol. – 1999. – Т. 66(12). – P. 1047-1049.
3. Анцыферов С.С. Адаптация информационно-распознающих биомедицинских систем / С.С. Анцыферов, Н.Н. Евтихий // Биомедицинская радиоэлектроника. – 2001. – Т. 3, № 10. – С. 52-57.
4. Анцыферов С.С. Метрология виртуальных систем / С.С. Анцыферов // Измерительная техника. – 2003. – № 5. – С. 17-21.
5. Анцыферов С.С. Адаптивные системы распознавания образов пространственно-временных полей / С.С. Анцыферов, Н.Н. Евтихий // Искусственный интеллект. – 2004. – № 3. – С. 405-416.
6. Анцыферов С.С. Адаптивная обработка информации пространственно-временных изотропных полей / С.С. Анцыферов, Н.Н. Евтихий // Оптический журнал. – 2006. – Т. 3, № 10. – С. 52-57.
7. Анцыферов С.С. Метрология интеллектуальных систем / С.С. Анцыферов // Искусственный интеллект. – 2008. – № 3. – С. 18-27.
8. Анцыферов С.С. Повышение интеллектуального потенциала адаптивных информационно-распознающих систем / С.С. Анцыферов // Искусственный интеллект. – 2010. – № 4. – С. 330-336.
9. Анцыферов С.С. Общие принципы построения и закономерности функционирования интеллектуальных систем / С.С. Анцыферов // Искусственный интеллект. – 2011. – № 3. – С. 6-15.
10. Денисов А.А. Теоретические основы кибернетики. Информационное поле / Денисов А.А. – Л. : ЛПИ, 1975.



## Literatura

1. Antsyferov S.S. Opticheskiy zhurnal. 1999. T. 66. № 2. S. 44-48.
2. Antsyferov S.S. J.Opt. Technol. 1999. T. 66(12). P. 1047-1049.
3. Antsyferov S.S. Biomedicinskaya radioelektronika. 2001. T. 3. № 10. S. 52-57.
4. Antsyferov S.S. Izmeritelinaya tehnika. 2003. № 5. S. 17-21.
5. Antsyferov S.S. Iskusstvennyj intellect. 2004. № 3. S. 405-416.
6. Antsyferov S.S. Opticheskiy zhurnal. 2006. T. 3. № 10. S. 52-57.
7. Antsyferov S.S. Iskusstvennyj intellect. 2008. № 3. S. 18-27.
8. Antsyferov S.S. Iskusstvennyj intellect. 2010. № 4. S. 330-336.
9. Antsyferov S.S. Iskusstvennyj intellect. 2011. № 3. S. 6-15.
10. Denisov A.A. Teoreticheskie osnovy kibernetiki. Informacionnoe pole. L.: LPI, 1975.

**S.S. Antsyferov**

### *Metrological Bases for Formation of Image Information Space in Intellectual Systems of Processing*

Now the huge attention is given by experts of the most various profile to the problem of creation of intellectual systems, i.e. the systems, which are carrying out processing of information at level of its “understanding” with wide use of the saved-up knowledge base. The urgency of this problem is caused as increasing information complexity of the studied phenomena and processes, and more and more accruing in a number of cases by intensity of processed information streams. Growth of requirements to accuracy (reliability) of results of processing only raises an urgency of the specified problem. The successful solution of this problem is connected not only with definition of principles of construction and regularities of functioning of intellectual systems (IS) to that a number of works [1-9] was already devoted, but also with the solution of not less important task, – development of methodology of their metrological providing. So, in the works [1-3], [5], [6], [8], [9] the directions of increase of intellectual potential of the adaptive information distinguishing systems (AIDS) realizing methodology of structural and stochastic approximation of information of existential fields and process of their processing were designated, principles of adaptation of parameters of processing algorithms, in particular, by a way of creation of evolutionary process of search of optimum model of adaptive processing within association of the symbolist and connective directions, i.e. association of logic and neural paradigms are defined. In the works [4], [7], system principles of metrological providing IS, in particular such metrological indicators as speed of adaptation, reliability and stability of results in the course of dynamic supervision were considered. However, the exit of AIDS to new, higher level of complexity, transfers them to the category of developing systems of processing, it is possible with the distributed multicomponent and multicoherent structure that, in turn, demands when developing such systems not only the accounting of regularities of their construction, functioning, development and feasibility, but also definition of adequate estimates of characteristics of these regularities (sets of system parameters) by formation of the corresponding forms of representation of the main scales of measurement of parameters as objective criteria of an assessment of quality of IS.

The concept of information complexity is one of the main in the theory of information approach to the analysis of systems [10]. The sense of this concept consists that any distribution of information against the logic communications imposed on it should possess a certain contents. In the analysis of these or those situations it is accepted to say that they have bigger or smaller sense from the point of view of definite purposes. Measurability of the contents, sense of a situation thereby is defined. The concept of an information field allows to find a quantitative assessment of the contents, sense on the basis of tracing of ways of realization of logic communications.

*Статья поступила в редакцию 05.06.2012.*