

## МЕТРИЧЕСКИЕ СВОЙСТВА ГРАНУЛЯЦИИ ИНФОРМАЦИИ

А. Г. Каграманян<sup>1</sup>, В. П. Машталир<sup>2</sup>, В. В. Шляхов<sup>3</sup>

<sup>1</sup> ХНУ, г. Харьков, Украина

<sup>2</sup> ХНУРЭ, г. Харьков, Украина

<sup>3</sup> ХНУРЭ, г. Харьков, Украина

Для задач грануляции — представления информации с различной степенью ее детализации или обобщения — исследуются метрические свойства конечных разбиений произвольных измеримых множеств. Вводится и обосновывается метрика, для которой указываются равносильные формы представления, предлагаются ее вероятностные интерпретации. Обсуждается гипотеза об общем виде функционала, являющегося метрикой на разбиениях измеримых множеств.

ГРАНУЛЯЦИЯ, МНОЖЕСТВО, РАЗБИЕНИЕ, МЕТРИКА.

### Введение

Granular Computing — грануляция информации, включающая в себя огрубление–детализацию данных, их стратифицированную структуризацию и причинно–следственные отношения в процессе регистрации, трансформации и интерпретации [1], оперирует, как правило, с разбиениями исходного (или индуцированного признакового) пространства [2]. Применение классов эквивалентностей, выступающих в качестве, пожалуй, основного объекта изучения и манипуляций, приводит к необходимости решения целого класса задач. В первую очередь следует указать поиск и анализ внутренних эквивалентностей [3, 4], установление правомочности (согласованности) использования совокупности эквивалентностей [5, 6], в том числе внешних, введение которых диктуется предметной областью, наконец, разумное определение контекстно–зависимых операций на классах эквивалентностей [7, 8]. Достаточно общим вопросом для задач указанного типа является анализ сходства (различия) между ансамблями группировок данных. Однако метрические свойства, вообще говоря, хорошо изучены лишь для анализа в форматах «точка–точка», «точка–множество», «множество–множество».

Наибольшего развития сравнение разбиений достигло при исследовании эффективности кластеризации [9, 10] и сегментации изображений [11–13] в основном в рамках «ground truth» подхода. Иначе говоря, когда результаты автоматической грануляции сравниваются с результатами, полученными человеком. Как правило, самые распространенные коэффициенты сходства (различия) не удовлетворяют неравенству треугольника, что ограничивает диапазон их использования. Вместе с тем известен и ряд метрик на разбиениях. Прежде всего — это метрика, имеющая структурное сходство с метрикой Хаусдорфа, в основу которой положены суммы максимальных мощностей пересечений каждого класса эквивалентностей со всеми элементами разбиений второго множества [14]. Достаточно широ-

кое распространение получила метрика энтропийного типа, названная вариацией информации, которая базируется на понятиях теории информации [15]. При использовании метрики MiCRoM (Minimum Cost Region Matching) поиск расстояния между разбиениями сводится к решению задачи математического программирования, а именно — транспортной задачи [16].

На интуитивном уровне понятно, что метрика на разбиениях должна интегрально учитывать характеристики и сходства, и различия классов эквивалентностей. В этом плане можно выделить простейший инструмент — операции пересечения и симметрической разности, соответственно. Конечномерный случай построения метрики такого типа обоснован в [17], некоторые метрические свойства для разбиений произвольных множеств проанализированы в [18]. Цель работы — обоснование, анализ и интерпретация инструментальных средств «множественного» сравнения, т. е. сравнения фактор–множеств.

### 1. Постановка задачи

Пусть некоторое измеримое множество  $\Omega$  с мерой  $\mu(\cdot)$  характеризует информацию, подлежащую в процессе анализа грануляции. Под мерой понимается количество точек (в конечном случае), длина, площадь, объем, распределение масс, вероятностей и т. п. Будем полагать, что мера конечна, т. е.  $\mu(\Omega) < \infty$  и  $F_\Omega$  — множество всех измеримых подмножеств множества  $\Omega$ . Рассмотрим  $\Pi_\Omega$  — множество конечных (по числу элементов) разбиений множества  $\Omega$ , т. е.  $\alpha \in \Pi_\Omega \Leftrightarrow \alpha = \{A_i\}_{i=1}^n$ :

$$\begin{cases} A_i \in F_\Omega \quad \forall i \in \{1, 2, \dots, n\}; \\ A_i \cap A_{i'} = \emptyset, \quad \forall i \neq i' \in \{1, 2, \dots, n\}; \\ A_1 \cup \dots \cup A_n = \Omega. \end{cases} \quad (1)$$

Пусть даны произвольные разбиения  $\alpha, \beta, \gamma, \dots \in \Pi_\Omega$  и  $\alpha = \{A_i\}_{i=1}^n, \beta = \{B_j\}_{j=1}^m, \gamma = \{C_k\}_{k=1}^l, \dots$ . Задача заключается в установлении, изучении и интерпретации соотношений между  $\alpha, \beta, \gamma, \dots$

## 2. Метрические свойства разбиений

Введем на  $\Pi_\Omega \times \Pi_\Omega$  функционал

$$\rho(\alpha, \beta) = \sum_{i=1}^n \sum_{j=1}^m \mu(A_i \Delta B_j) \mu(A_i \cap B_j), \quad (2)$$

где « $\Delta$ » — симметрическая разность множеств, и покажем, что он представляет собой метрику на произвольных разбиениях измеримых множеств. Предварительно укажем несколько полезных свойств разбиений и найдем эквивалентную форму функционала (2).

**Свойство 1.** Для любого множества  $D \subseteq \Omega$  и произвольного разбиения  $\alpha \in \Pi_\Omega$  справедливо равенство  $\mu(D) = \sum_{i=1}^n \mu(D \cap A_i)$ .

Действительно, разбиение  $\alpha \in \Pi_\Omega$  дробит множество  $D$  на непересекающиеся подмножества  $D \cap A_i$ , и тогда из свойства аддитивности меры следует требуемое равенство.

**Свойство 2.** Пересечение  $\alpha \cap \beta = \{A_i \cap B_j\}_{i=1, n, j=1, m}$  пары  $\alpha, \beta \in \Pi_\Omega$  произвольных разбиений также является разбиением.

Действительно, если выбрать произвольный элемент  $\omega \in \Omega$ , то, поскольку  $\alpha$  и  $\beta$  являются разбиениями, найдутся номера  $i \in \{1, \dots, n\}$  и  $j \in \{1, \dots, m\}$ , для которых  $\omega \in A_i$  и  $\omega \in B_j$ , т. е.  $\omega \in A_i \cap B_j$ . Таким образом,

$$\bigcup_{i=1}^n \bigcup_{j=1}^m \{A_i \cap B_j\} = \Omega. \quad (3)$$

С другой стороны, в силу ассоциативности и коммутативности операции пересечения множеств можно записать:

$$(A_i \cap B_j) \cap (A_{i'} \cap B_{j'}) = (A_i \cap A_{i'}) \cap (B_j \cap B_{j'}).$$

Но поскольку пары  $(i, j)$  и  $(i', j')$  не равны, то либо  $i \neq i'$ , либо  $j \neq j'$ , либо эти неравенства выполняются одновременно, что означает: одно из множеств  $A_i \cap A_{i'}$  или  $B_j \cap B_{j'}$  равно  $\emptyset$ , т. к. они принадлежат разбиениям  $\alpha$  и  $\beta$ , соответственно. Таким образом, в итоге имеем

$$(A_i \cap B_j) \cap (A_{i'} \cap B_{j'}) = \emptyset$$

для любых пар  $(i, j), (i', j') \in \{1, \dots, n\} \times \{1, \dots, m\}$ . В совокупности с равенством (3) получаем выполнение соотношений (1), т. е. справедливость свойства 2.

Рассмотрим пересечения трех произвольных разбиений  $\alpha, \beta, \gamma \in \Pi_\Omega$  и обозначим  $\theta_{ijk} = \mu(A_i \cap B_j \cap C_k)$ ,  $i = \overline{1, n}$ ;  $j = \overline{1, m}$ ;  $k = \overline{1, l}$ .

**Свойство 3.** Для пересечений трех произвольных разбиений  $\alpha, \beta, \gamma \in \Pi_\Omega$  верны соотношения:

$$\mu(C_k) = \sum_{i=1}^n \sum_{j=1}^m \theta_{ijk}, \quad (4)$$

$$\mu(A_i \cap B_j) = \sum_{k=1}^l \theta_{ijk}, \quad (5)$$

$$\mu(A_i \cap C_k) = \sum_{j=1}^m \theta_{ijk}, \quad (6)$$

$$\mu(B_j \cap C_k) = \sum_{i=1}^n \theta_{ijk}. \quad (7)$$

Действительно, если в соответствии со свойством (2) рассмотреть разбиение  $\xi = \alpha \cap \beta$  и применить свойство 1, полагая, что  $D$  — это множество  $C_k$ , то

$$\mu(C_k) = \sum_{i=1}^n \sum_{j=1}^m \mu(C_k \cap (A_i \cap B_j)) = \sum_{i=1}^n \sum_{j=1}^m \theta_{ijk},$$

т. е. получаем равенство (4). Рассуждая совершенно аналогично и используя в качестве  $D$  поочередно множества  $A_i \cap B_j, A_i \cap C_k, B_j \cap C_k$ , а в качестве соответствующего разбиения —  $\gamma, \beta$  и  $\alpha$ , получаем справедливость равенств (5) — (7).

**Лемма.** Функционал  $\rho(\alpha, \beta)$  имеет эквивалентную форму

$$\rho(\alpha, \beta) = \sum_{i=1}^n [\mu(A_i)]^2 + \sum_{j=1}^m [\mu(B_j)]^2 - 2 \sum_{i=1}^n \sum_{j=1}^m [\mu(A_i \cap B_j)]^2.$$

**Доказательство.** Для любых измеримых множеств  $E$  и  $F$  имеет место

$$\mu(E \Delta F) = \mu(E) + \mu(F) - 2\mu(E \cap F). \quad (8)$$

Действительно, учитывая равенства

$$E = (E \setminus F) \cup (E \cap F) \text{ и } F = (F \setminus E) \cup (E \cap F),$$

причем  $E \setminus F$  и  $E \cap F$  не пересекаются, из свойства аддитивности меры получаем:

$$\begin{aligned} \mu(E) &= \mu(E \setminus F) + \mu(E \cap F), \\ \mu(F) &= \mu(F \setminus E) + \mu(E \cap F), \end{aligned}$$

складывая которые с учетом  $E \Delta F = (E \setminus F) \cup (F \setminus E)$ , имеем:

$$\begin{aligned} \mu(E) + \mu(F) &= \mu(E \setminus F) + \mu(F \setminus E) + 2\mu(E \cap F) = \\ &= \mu(E \Delta F) + 2\mu(E \cap F), \end{aligned}$$

что равносильно (8).

Применяя (8) к  $A_i$  и  $B_j$ , представим (2) в виде:

$$\begin{aligned} \rho(\alpha, \beta) &= \sum_{i=1}^n \sum_{j=1}^m \mu(A_i \Delta B_j) \mu(A_i \cap B_j) = \\ &= \sum_{i=1}^n \sum_{j=1}^m \mu(A_i \cap B_j) [\mu(A_i) + \mu(B_j) - 2\mu(A_i \cap B_j)] = \\ &= \sum_{i=1}^n \sum_{j=1}^m \mu(A_i) \mu(A_i \cap B_j) + \\ &+ \sum_{i=1}^n \sum_{j=1}^m \mu(B_j) \mu(A_i \cap B_j) - 2 \sum_{i=1}^n \sum_{j=1}^m [\mu(A_i \cap B_j)]^2 = \\ &= \sum_{i=1}^n \mu(A_i) \sum_{j=1}^m \mu(A_i \cap B_j) + \\ &+ \sum_{j=1}^m \mu(B_j) \sum_{i=1}^n \mu(B_j \cap A_i) - 2 \sum_{i=1}^n \sum_{j=1}^m [\mu(A_i \cap B_j)]^2. \end{aligned}$$

В цепочке этих равенств выделим суммы вида

$$\sum_{j=1}^m \mu(A_i \cap B_j), \sum_{i=1}^n \mu(B_j \cap A_i)$$

и, применяя к ним свойство 1, подразумевая под  $D$   $A_i$  и  $B_j$ , соответственно, получим:

$$\sum_{j=1}^m \mu(A_i \cap B_j) = \mu(A_i), \sum_{i=1}^n \mu(B_j \cap A_i) = \mu(B_j)$$

для любых  $i$  и  $j$ , где  $i = \overline{1, n}$ ,  $j = \overline{1, m}$ . Таким образом,

$$\rho(\alpha, \beta) = \sum_{i=1}^n [\mu(A_i)]^2 + \sum_{j=1}^m [\mu(B_j)]^2 - 2 \sum_{i=1}^n \sum_{j=1}^m [\mu(A_i \cap B_j)]^2,$$

что и требовалось доказать.

**Теорема 1.** Функционал (2), определенный на произвольных измеримых разбиениях  $\alpha, \beta$  измеримого множества  $\Omega$ , является метрикой.

Доказательство сводится к проверке выполнения аксиом рефлексивности, симметричности и неравенства треугольника.

Симметричность непосредственно следует из вида функционала (2). Докажем рефлексивность, а именно:  $\rho(\alpha, \beta) = 0 \Leftrightarrow \alpha = \beta$ . Рассмотрим необходимость. Для этого представим функционал (2) в виде:

$$\rho(\alpha, \alpha) = \sum_{i=1}^n \sum_{j=1}^n \mu(A_i \Delta A_j) \mu(A_i \cap A_j) = \\ = \sum_{i=1}^n \mu(A_i \Delta A_i) \mu(A_i \cap A_i) + 2 \sum_{\substack{i, j=1 \\ i < j}}^n \mu(A_i \Delta A_j) \mu(A_i \cap A_j),$$

т. е.  $\rho(\alpha, \alpha)$  состоит из  $n$  слагаемых с совпадающими номерами, а значит, и с совпадающими элементами разбиения, и из  $n^2 - n$  слагаемых с различными элементами разбиения. Первая группа слагаемых состоит из нулей, т. к.  $\mu(A_i \Delta A_i) = 0$ , для  $i = \overline{1, n}$ , а вторая группа дает нулевые слагаемые, т. к.  $\mu(A_i \cap A_j) = 0$  для  $\forall i, j = \overline{1, n}$  и  $i \neq j$ . Таким образом, в одну сторону рефлексивность доказана.

Покажем достаточность, т. е. пусть для двух, в общем случае, различных разбиений  $\alpha, \beta \in \Pi_\Omega$  имеет место  $\rho(\alpha, \beta) = 0$ . Поскольку все слагаемые функционала (2) положительны, то он равен 0 только тогда, когда каждое из них равно 0. Выберем некоторый элемент разбиения  $A' \in \alpha$ . В исходном функционале он входит в набор «нулевых» слагаемых  $\mu(A' \Delta B_j) \mu(A' \cap B_j) = 0$ , где  $j = \overline{1, m}$ . Если предположить, что  $A'$  не является элементом разбиения  $\beta$ , то для всех  $B_j \in \beta$  выполняется  $\mu(A' \Delta B_j) \neq 0$ . Но тогда для всех номеров  $j$  должно выполняться

$\mu(A' \cap B_j) = 0$ . Однако это возможно только тогда, когда  $A' \neq \emptyset$ , т. к.  $A' \in \Omega$ , а весь набор  $\beta = \{B_j\}_{j=1}^m$  покрывает множество  $\Omega$ . В то же время  $A'$  — элемент покрытия  $\beta$  этого же множества  $\Omega$  и, естественно,  $A' \neq \emptyset$ . Тогда найдутся элементы разбиения  $\beta$ , покрывающие  $A'$  и имеющие с ним непустые пересечения. Но тогда для этих элементов выполняется неравенство  $\mu(A' \cap B_j) \neq 0$ , т. е. получаем противоречие. В итоге можно утверждать, что произвольный элемент  $A' \in \alpha$  является элементом  $\beta$ , или  $\alpha \subset \beta$ . В силу симметрии абсолютно аналогично можно показать, что  $\beta \subset \alpha$ , т. е.  $\alpha = \beta$ . Таким образом, рефлексивность доказана.

Приступим к доказательству неравенства треугольника. Допустим, имеется три произвольных разбиения  $\alpha, \beta, \gamma \in \Pi_\Omega$  множества  $\Omega$ . Необходимо доказать, что

$$\rho(\alpha, \gamma) + \rho(\gamma, \beta) \geq \rho(\alpha, \beta). \quad (9)$$

Используя лемму 1, получим:

$$\rho(\alpha, \gamma) + \rho(\gamma, \beta) = \sum_{i=1}^n [\mu(A_i)]^2 + \sum_{j=1}^m [\mu(B_j)]^2 + \\ + 2 \sum_{k=1}^l [\mu(C_k)]^2 - 2 \sum_{i=1}^n \sum_{k=1}^l [\mu(A_i \cap C_k)]^2 - \\ - 2 \sum_{k=1}^l \sum_{j=1}^m [\mu(C_k \cap B_j)]^2.$$

Отсюда:

$$\rho(\alpha, \gamma) + \rho(\gamma, \beta) - \rho(\alpha, \beta) = \\ = \sum_{k=1}^l [\mu(C_k)]^2 - \sum_{i=1}^n \sum_{k=1}^l [\mu(A_i \cap C_k)]^2 - \\ - \sum_{k=1}^l \sum_{j=1}^m [\mu(B_j \cap C_k)]^2 + \sum_{i=1}^n \sum_{j=1}^m [\mu(A_i \cap B_j)]^2.$$

Тем самым, неравенство (9), используя свойство 3, можно представить в виде:

$$\sum_{k=1}^l \left( \sum_{i=1}^n \sum_{j=1}^m \theta_{ijk} \right)^2 + \sum_{i=1}^n \sum_{j=1}^m \left( \sum_{k=1}^l \theta_{ijk} \right)^2 \geq \\ \geq \sum_{i=1}^n \sum_{k=1}^l \left( \sum_{j=1}^m \theta_{ijk} \right)^2 + \sum_{j=1}^m \sum_{k=1}^l \left( \sum_{i=1}^n \theta_{ijk} \right)^2. \quad (10)$$

Зафиксируем  $l = 1$ , тогда, обозначая  $g_{ij} = \theta_{ij1}$ , для любых  $g_{ij} \geq 0$ ,  $i = \overline{1, n}$ ,  $j = \overline{1, m}$ , имеем:

$$\left( \sum_{i=1}^n \sum_{j=1}^m g_{ij} \right)^2 + \sum_{i=1}^n \sum_{j=1}^m (g_{ij})^2 \geq \sum_{i=1}^n \left( \sum_{j=1}^m g_{ij} \right)^2 + \sum_{j=1}^m \left( \sum_{i=1}^n g_{ij} \right)^2. \quad (11)$$

Заметим, если интерпретировать набор чисел  $g_{ij}$  как  $(n \times m)$  матрицу  $G$ , то в правой части этого неравенства в скобках, соответственно, стоят сумма элементов строк и сумма элементов столбцов. Структура левой части неравенства имеет вид:

$$2 \sum_{i,j} g_{ij}^2 + \sum_{(i,j) \neq (i',j')} g_{ij} g_{i'j'}, \quad (12)$$

а правой:

$$2 \sum_{i,j} g_{ij}^2 + \sum_i \sum_{j \neq j'} g_{ij} g_{ij'} + \sum_j \sum_{i \neq i'} g_{ij} g_{i'j}. \quad (13)$$

Сравнивая (12) и (13), отметим: у них есть общая часть (удвоенная сумма квадратов), в (12) входят все попарные произведения элементов матрицы  $G$  с неравными индексами, а в (13) — все попарные произведения с неравными индексами, но из одной строки или из одного столбца (т. е. отсутствующих попарные произведения, например, диагональных элементов). Таким образом, в (13) число слагаемых меньше, чем в выражении (12), и при этом все слагаемые выражения (13) содержатся в (12), т. е. неравенство (11) доказано (выполняется (10) при  $l = 1$ ). Рассмотрим теперь числа  $\theta_{ijk}$  как набор из  $l$  матриц:

$$G_k = (g_{ij}^k)_{j=1, \dots, n}, \quad g_{ij}^k = \theta_{ijk}, \quad k = \overline{1, l}.$$

Тем самым, неравенство (10) трансформируется к виду:

$$\sum_k \left( \sum_{i,j} g_{ij}^k \right)^2 + \sum_{i,j} \left( \sum_k g_{ij}^k \right)^2 \geq \sum_{i,k} \left( \sum_j g_{ij}^k \right)^2 + \sum_{j,k} \left( \sum_i g_{ij}^k \right)^2$$

или

$$\sum_k \left[ \left( \sum_{i,j} g_{ij}^k \right)^2 + \left( \sum_{i,j} g_{ij}^k \right)^2 \right] + S \geq \sum_k \left[ \left( \sum_i \left( \sum_j g_{ij}^k \right)^2 \right) + \left( \sum_j \left( \sum_i g_{ij}^k \right)^2 \right) \right],$$

где  $S = \sum_{i,j} \left( \sum_k g_{ij}^k \right)^2 - \sum_{i,j,k} \left( g_{ij}^k \right)^2 \geq 0$ ,

поскольку состоит только из двойных произведений элементов матриц  $G_k$ , а все  $g_{ij}^k \geq 0$ . С другой стороны, применяя (11) к каждой из матриц  $G_k$ , суммируя по  $k$  и добавляя в левую часть  $S$ , получаем последнее неравенство, а следовательно, справедливость (10), что завершает доказательство неравенства треугольника и теоремы в целом.

Таким образом, показано, что введенный нами функционал (2) на конечных разбиениях на измеримые подмножества измеримого множества  $\Omega$  является метрикой.

**Теорема 2.** Для произвольной области  $\Omega \subseteq \mathbb{R}^s$  и почти всюду (за исключением множеств меры ноль) положительной функции  $f(x)$ , для которой справедливо  $\int_{\Omega} f(x) dx < \infty$ , и пары произвольных разбиений  $\alpha, \beta \in \Pi_{\Omega}$ , имеет место

$$\begin{aligned} \left( \int_{\Omega} f(x) dx \right)^2 &\geq \sum_{i=1}^n \sum_{j=1}^m \left( \int_{A_i \cap B_j} f(x) dx \right)^2 \\ &+ \sum_{i=1}^n \sum_{i'=i-1}^n \int_{A_i} f(x) dx \int_{A_{i'}} f(x) dx + \\ &+ \sum_{j=1}^m \sum_{j'=j+1}^m \int_{B_j} f(x) dx \int_{B_{j'}} f(x) dx. \end{aligned} \quad (14)$$

**Доказательство.** Поскольку функции  $f(x)$  индуцируют меру на измеримых подмножествах  $A \subseteq \Omega$  в виде интеграла  $\mu(A) = \int_A f(x) dx$ , тогда для двух конечных разбиений  $\alpha, \beta$  будем иметь:

$$d(\alpha, \beta) = \sum_{i=1}^n \sum_{j=1}^m \int_{A_i \Delta B_j} f(x) dx \int_{A_i \cap B_j} f(x) dx.$$

В соответствии с леммой, принимая во внимание неотрицательность функции  $d(\alpha, \beta)$ , получаем интегральное неравенство

$$\sum_{i=1}^n \left[ \int_{A_i} f(x) dx \right]^2 + \sum_{j=1}^m \left[ \int_{B_j} f(x) dx \right]^2 \geq 2 \sum_{i=1}^n \sum_{j=1}^m \left[ \int_{A_i \cap B_j} f(x) dx \right]^2,$$

дополняя в котором левую часть до сумм полных квадратов, находим

$$\begin{aligned} \left[ \sum_{i=1}^n \int_{A_i} f(x) dx \right]^2 + \left[ \sum_{j=1}^m \int_{B_j} f(x) dx \right]^2 &\geq 2 \sum_{i=1}^n \sum_{j=1}^m \int_{A_i \cap B_j} f(x) dx \int_{A_i \cap B_j} f(x) dx + \\ &+ 2 \sum_{i=1}^n \sum_{i'=i+1}^n \int_{A_i} f(x) dx \int_{A_{i'}} f(x) dx + 2 \sum_{j=1}^m \sum_{j'=j+1}^m \int_{B_j} f(x) dx \int_{B_{j'}} f(x) dx. \end{aligned}$$

Учитывая, что  $\alpha, \beta$  — конечные (по числу классов эквивалентностей) разбиения множества  $\Omega$ , имеем:

$$\sum_{i=1}^n \int_{A_i} f(x) dx = \int_{\Omega} f(x) dx, \quad \sum_{j=1}^m \int_{B_j} f(x) dx = \int_{\Omega} f(x) dx,$$

откуда сразу же следует неравенство (14), что и требовалось доказать.

На множестве разбиений  $\Pi_{\Omega}$  введем функционал

$$S(\alpha) = \sum_{i=1}^n [\mu(A_i)]^2 \quad (15)$$

и обозначим  $\alpha\beta = \alpha \cap \beta$ , тогда справедлива

**Теорема 3.** Для любых трех разбиений  $\alpha, \beta, \gamma \in \Pi_{\Omega}$  множества  $\Omega$  справедливо неравенство

$$S(\gamma) \geq S(\alpha\gamma) + S(\beta\gamma) - S(\alpha\beta). \quad (16)$$

**Доказательство.** Во-первых, метрика на разбиениях  $\rho(\alpha, \beta)$  с учетом леммы 1 может быть выражена через функционал (15):

$$\rho(\alpha, \beta) = S(\alpha) + S(\beta) - 2S(\alpha\beta).$$

Во-вторых, из неравенства треугольника для метрики (2) и любых трех разбиений  $\alpha, \beta, \gamma \in \Pi_\Omega$  получаем:

$$S(\alpha) + S(\gamma) - 2S(\alpha\gamma) + S(\beta) + S(\gamma) - 2S(\beta\gamma) \geq S(\alpha) + S(\beta) - 2S(\alpha\beta),$$

откуда следует требуемое неравенство (16). Таким образом, теорема 3 доказана.

Из теорем 2 и 3 вытекает

*Следствие 1.* Для произвольной области  $\Omega \subseteq \mathbb{R}^s$  и функции  $f(x)$ , удовлетворяющей условиям теоремы 2 и произвольных разбиений  $\alpha, \beta, \gamma \in \Pi_\Omega$ , имеет место

$$\left(\int_{\Omega} f(x) dx\right)^2 \geq \sum_{i=1}^n \sum_{k=1}^l \left(\int_{A_i \cap C_k} f(x) dx\right)^2 + \sum_{j=1}^m \sum_{k=1}^l \left(\int_{B_j \cap C_k} f(x) dx\right)^2 + 2 \sum_{k=1}^l \sum_{i=1}^n \int_{C_k} f(x) dx \int_{C_k} f(x) dx - \sum_{i=1}^n \sum_{j=1}^m \left(\int_{A_i \cap B_j} f(x) dx\right)^2.$$

**Доказательство.** Для произвольного конечного разбиения  $\alpha \in \Pi_\Omega$  функционал (12) примет вид:

$$S(\alpha) = \sum_{i=1}^n \left[ \int_{A_i} f(x) dx \right]^2.$$

Неравенство (16) в этой ситуации станет интегральным и будет иметь форму:

$$\sum_{k=1}^l \left(\int_{C_k} f(x) dx\right)^2 \geq \sum_{i=1}^n \sum_{k=1}^l \left(\int_{A_i \cap C_k} f(x) dx\right)^2 + \sum_{j=1}^m \sum_{k=1}^l \left(\int_{B_j \cap C_k} f(x) dx\right)^2 - \sum_{i=1}^n \sum_{j=1}^m \left(\int_{A_i \cap B_j} f(x) dx\right)^2.$$

Если к левой и правой части этого неравенства прибавить слагаемое

$$2 \sum_{k=1}^l \sum_{k'=k+1}^l \int_{C_k} f(x) dx \int_{C_{k'}} f(x) dx,$$

то его правая часть сразу примет требуемый вид, а левая часть с учетом того, что  $\gamma = \{C_k\}_{k=1}^l$  — разбиение области  $\Omega$ , будет представлена в форме:

$$\sum_{k=1}^l \left(\int_{C_k} f(x) dx\right)^2 + 2 \sum_{k=1}^l \sum_{k'=k+1}^l \int_{C_k} f(x) dx \int_{C_{k'}} f(x) dx = \left(\sum_{k=1}^l \int_{C_k} f(x) dx\right)^2 = \left(\int_{\Omega} f(x) dx\right)^2,$$

что и требовалось доказать.

### 3. Вероятностная интерпретация метрики

Из свойства 1 следует, что для любого  $D \subseteq \Omega$  при  $\mu(A_i) \neq 0$

$$\frac{\mu(D)}{\mu(\Omega)} = \sum_{i=1}^n \frac{\mu(D \cap A_i)}{\mu(\Omega)} = \sum_{i=1}^n \frac{\mu(D \cap A_i)}{\mu(A_i)} \frac{\mu(A_i)}{\mu(\Omega)}, \quad (17)$$

тогда, прибегая к терминологии теории вероятностей, можно говорить о наборе  $\{A_i\}_{i=1}^n$  как о наборе гипотез, и соотношение (17) представляет собой формулу полной вероятности

$$p(D) = \sum_{i=1}^n p(A_i) p(D/A_i),$$

где  $p(D) = \mu(D)/\mu(\Omega)$ ,  $p(A_i) = \mu(A_i)/\mu(\Omega)$  — вероятности события  $D$  и гипотез  $A_i$ , а

$$p(D/A_i) = \frac{\mu(D \cap A_i)}{\mu(\Omega)} \frac{\mu(\Omega)}{\mu(A_i)} = \frac{\mu(D \cap A_i)/\mu(\Omega)}{\mu(A_i)/\mu(\Omega)} = \frac{p(D \cap A_i)}{p(A_i)}$$

является условной вероятностью события  $D$  при условии выполнения гипотезы  $A_i$ . Таким образом, можно заключить: свойство 1 может быть получено и как следствие формулы полной вероятности. Но тогда возможна и более тесная связь метрических свойств грануляции информации с иными понятиями теории вероятностей. Остановимся на этом подробнее.

Нетрудно заметить, что если  $\Omega \subset \mathbb{R}^s$  — некоторое измеримое множество и имеет место схема геометрической вероятности, тогда вероятность любого события  $D$  равна

$$p(D) = \frac{\mu(D)}{\mu(\Omega)},$$

где  $\mu$  — мера, соответствующая  $\mathbb{R}^s$  ( $s=1$  — длина,  $s=2$  — площадь и т. д.). С другой стороны, если  $\rho(\alpha, \beta)$  нормировать, а именно разделить на  $\mu^2(\Omega)$ , то получим

$$\bar{\rho}(\alpha, \beta) = \frac{\rho(\alpha, \beta)}{\mu^2(\Omega)} = \sum_{i=1}^n \sum_{j=1}^m \frac{\mu(A_i \Delta B_j)}{\mu(\Omega)} \frac{\mu(A_i \cap B_j)}{\mu(\Omega)} = \sum_{i=1}^n \sum_{j=1}^m p(A_i \Delta B_j) p(A_i \cap B_j), \quad (18)$$

где  $p(A_i \Delta B_j)$  и  $p(A_i \cap B_j)$  — вероятности событий  $A_i \Delta B_j$  и  $A_i \cap B_j$ , соответственно. Для метрики (18), использовался функционал (2). В случае, если мы будем пользоваться равенством из леммы, то

$$\bar{\rho}(\alpha, \beta) = \sum_{i=1}^n [p(A_i)]^2 + \sum_{j=1}^m [p(B_j)]^2 - 2 \sum_{i=1}^n \sum_{j=1}^m [p(A_i \cap B_j)]^2 \quad (19)$$

Если рассматривать множество  $\Omega$  как пространства элементарных исходов, являющееся элементом вероятностного пространства  $\langle \Omega, \mathcal{F}_\Omega, p \rangle$ , где  $\Omega$  — в общем случае произвольное множество,

$\mathcal{F}_\Omega$  —  $\delta$ -алгебра его подмножеств, включающая в себя пустое множество  $\emptyset$  и  $\Omega$ , а  $p$  — вероятностная мера, относительно которой все элементы  $\mathcal{F}_\Omega$  измеримы, т. е. для любого  $D \in \mathcal{F}_\Omega$  существует число  $0 \leq p(D) \leq 1$ , удовлетворяющее аксиоматике теории вероятностей А. Н. Колмогорова, то на исходном вероятностном пространстве можно рассматривать различные полные группы событий или наборы гипотез. Более точно, учитывая, что для произвольного набора гипотез  $\{S_i\}_{i=1}^n$  имеет место:  $S_i \in \mathcal{F}_\Omega$ ,  $i = \overline{1, n}$ ,  $\bigcup_{i=1}^n S_i = \Omega$ ,  $S_i \cap S_{i'} = \emptyset \quad \forall i, i' : i \neq i' \in \{1, \dots, n\}$ , то метрику (2) и ее равносильную форму из леммы для двух наборов гипотез  $K_1 = \{S_i\}_{i=1}^n$  и  $K_2 = \{Q_j\}_{j=1}^m$  можно представить в виде:

$$\begin{aligned} \bar{\rho}(K_1, K_2) &= \sum_{i=1}^n \sum_{j=1}^m p(S_i \Delta Q_j) p(S_i \cap Q_j) = \\ &= \sum_{i=1}^n [p(S_i)]^2 + \sum_{j=1}^m [p(Q_j)]^2 - 2 \sum_{i=1}^n \sum_{j=1}^m [p(S_i \cap Q_j)]^2. \end{aligned}$$

С другой стороны, можно рассматривать две дискретные случайные величины  $\xi_1$  и  $\xi_2$ , заданные конечным числом значений, т. е.

$\xi_1$	$x_1$	...	$x_i$	...	$x_n$
$p$	$u_1$	...	$u_i$	...	$u_n$
$\xi_2$	$y_1$	...	$y_j$	...	$y_m$
$p$	$v_1$	...	$v_j$	...	$v_m$

Здесь  $u_i = p(\xi_1 = x_i)$ ,  $i = \overline{1, n}$ ;  $v_j = p(\xi_2 = y_j)$ ,  $j = \overline{1, m}$ , а  $\{x_i\}_{i=1}^n, \{y_j\}_{j=1}^m$  — наборы различных действительных (комплексных) чисел. Используя равенство (19), можно ввести метрику на множестве этих случайных величин:

$$\begin{aligned} \bar{\rho}(\xi_1, \xi_2) &= \sum_{i=1}^n [p(\xi_1 = x_i)]^2 + \sum_{j=1}^m [p(\xi_2 = y_j)]^2 - \\ &- 2 \sum_{i=1}^n \sum_{j=1}^m [p(\xi_1 = x_i, \xi_2 = y_j)]^2. \end{aligned}$$

Если  $\xi_1$  и  $\xi_2$  независимы, то

$$p(\xi_1 = x_i, \xi_2 = y_j) = p(\xi_1 = x_i) p(\xi_2 = y_j), \quad (3.34)$$

тогда, подставляя (3.32) в равенство (3.31), получим:

$$\begin{aligned} \bar{\rho}(\xi_1, \xi_2) &= \sum_{i=1}^n [p(\xi_1 = x_i)]^2 + \sum_{j=1}^m [p(\xi_2 = y_j)]^2 - \\ &- 2 \sum_{i=1}^n [p(\xi_1 = x_i)]^2 \sum_{j=1}^m [p(\xi_2 = y_j)]^2 = \\ &= \sum_{i=1}^n [p(\xi_1 = x_i)]^2 \left\{ 1 - \sum_{j=1}^m [p(\xi_2 = y_j)]^2 \right\} + \\ &+ \sum_{j=1}^m [p(\xi_2 = y_j)]^2 \left\{ 1 - \sum_{i=1}^n [p(\xi_1 = x_i)]^2 \right\}. \end{aligned}$$

Если предположить, что распределения случайных величин  $\xi_1$  и  $\xi_2$  равномерны, т. е. все их значения равновероятны:  $p(\xi_1 = x_i) = 1/n$ ,  $p(\xi_2 = y_j) = 1/m$  при любых  $i = \overline{1, n}$  и  $j = \overline{1, m}$ , то

$$\bar{\rho}(\xi_1, \xi_2) = \frac{n}{n^2} \left(1 - \frac{m}{m^2}\right) + \frac{m}{m^2} \left(1 - \frac{n}{n^2}\right) = \frac{1}{n} + \frac{1}{m} - \frac{2}{nm}.$$

Таким образом, фактически доказана

**Теорема 4.** Расстояние между двумя независимыми конечными равномерными распределениями в смысле равенства (2) равно

$$\bar{\rho}(\xi_1, \xi_2) = \frac{1}{n} + \frac{1}{m} - \frac{2}{nm}. \quad (20)$$

Отметим, что если говорить о независимых случайных величинах, то нельзя говорить об их равенстве. В связи с этим свойство рефлексивности в данном случае отсутствует, что видно и из равенства (20).

Рассмотрим вероятностные интерпретации теорем предыдущего раздела. Из теоремы 2 вытекают три любопытных следствия.

**Следствие 2 (для гипотез).** Если в некотором вероятностном пространстве имеется два набора гипотез  $K_1 = \{S_i\}_{i=1}^n$  и  $K_2 = \{Q_j\}_{j=1}^m$ , то

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^m [p(S_i \cap Q_j)]^2 + \sum_{i=1}^n \sum_{i'=i+1}^n p(S_i) p(S_{i'}) + \\ + \sum_{j=1}^m \sum_{j'=j+1}^m p(Q_j) p(Q_{j'}) \leq 1. \end{aligned}$$

Доказательство с учетом неотрицательности функционала  $\bar{\rho}(K_1, K_2)$  совершенно аналогично доказательству теоремы 2.

**Следствие 3.** Если в некотором вероятностном пространстве произвольное событие  $D$  разбито на два набора попарно непересекающихся подсобытий  $\{E_i\}_{i=1}^n$  и  $\{\Phi_j\}_{j=1}^m$ , образующих относительно  $D$  полную группу, т. е.  $D = \sum_{i=1}^n E_i = \sum_{j=1}^m \Phi_j$ , то

$$\begin{aligned} \sum_{i=1}^n \sum_{j=1}^m [p(E_i \cap \Phi_j)]^2 + \sum_{i=1}^n \sum_{i'=i+1}^n p(E_i) p(E_{i'}) + \\ + \sum_{j=1}^m \sum_{j'=j+1}^m p(\Phi_j) p(\Phi_{j'}) \leq [p(D)]^2. \end{aligned}$$

**Следствие 4 (для дискретных случайных величин).**

Пусть на одном вероятностном пространстве заданы две дискретные случайные величины  $\xi_1$  и  $\xi_2$  с конечным числом значений, т. е.

$$u_i = p(\xi_1 = x_i), i = \overline{1, n}; v_j = p(\xi_2 = y_j), j = \overline{1, m},$$

а  $\{x_i\}_{i=1}^n, \{y_j\}_{j=1}^m$  — наборы различных действительных (комплексных) чисел, то

$$\sum_{i=1}^n \sum_{j=1}^m [p(\xi_1 = x_i, \xi_2 = y_j)]^2 + \sum_{i=1}^n \sum_{i'=i+1}^n u_i u_{i'} + \sum_{j=1}^m \sum_{j'=j+1}^m v_j v_{j'} \leq 1.$$

Если эти случайные величины  $\xi_1$  и  $\xi_2$  еще и независимы, т. е.

$$p(\xi_1 = x_i, \xi_2 = y_j) = p(\xi_1 = x_i) p(\xi_2 = y_j) = u_i v_j$$

для любых  $i = \overline{1, n}$  и  $j = \overline{1, m}$ , то

$$\sum_{i=1}^n u_i^2 \sum_{j=1}^m v_j^2 + \sum_{i=1}^n \sum_{l'=i+1}^n u_l u_{l'} + \sum_{j=1}^m \sum_{j'=j+1}^m v_j v_{j'} \leq 1.$$

Следствие 4 для своего обоснования использует рассуждения, полностью аналогичные рассуждениям, доказывающим теорему 3, неотрицательность функционала  $\bar{r}(\xi_1, \xi_2)$  и равенства

$$\sum_{i=1}^n u_i = \sum_{j=1}^m v_j = 1.$$

Рассмотрим теперь следствия из теоремы 3.

*Следствие 5 (для гипотез).* Для любых трех наборов гипотез,  $K_1 = \{S_i\}_{i=1}^n$ ,  $K_2 = \{Q_j\}_{j=1}^m$ ,  $K_3 = \{R_k\}_{k=1}^l$ , заданных на одном пространстве элементарных исходов  $\Omega$ , вероятности, связанные с этими гипотезами, удовлетворяют неравенству

$$\sum_{k=1}^l [p(R_k)]^2 \geq \sum_{i=1}^n \sum_{k=1}^l [p(S_i \cap R_k)]^2 + \sum_{j=1}^m \sum_{k=1}^l [p(Q_j \cap R_k)]^2 - \sum_{i=1}^n \sum_{j=1}^m [p(S_i \cap Q_j)]^2.$$

*Следствие 6 (для дискретных случайных величин).*

Пусть на одном пространстве элементарных исходов заданы три дискретные случайные величины с конечным числом значений, т. е.  $u_i = p(\xi_1 = x_i)$ ,  $i = \overline{1, n}$ ;  $v_j = p(\xi_2 = y_j)$ ,  $j = \overline{1, m}$ ;  $t_k = p(\xi_3 = z_k)$ ,  $k = \overline{1, l}$ , где  $\{x_i\}_{i=1}^n$ ,  $\{y_j\}_{j=1}^m$ ,  $\{z_k\}_{k=1}^l$  — наборы действительных (комплексных) чисел, тогда:

$$\begin{aligned} \sum_{k=1}^l t_k^2 &\geq \sum_{i=1}^n \sum_{k=1}^l [p(\xi_1 = x_i, \xi_3 = z_k)]^2 + \\ &+ \sum_{j=1}^m \sum_{k=1}^l [p(\xi_2 = y_j, \xi_3 = z_k)]^2 - \\ &- \sum_{i=1}^n \sum_{j=1}^m [p(\xi_1 = x_i, \xi_2 = y_j)]^2. \end{aligned}$$

Если эти случайные величины независимы в совокупности, т. е.

$$\begin{cases} p(\xi_1 = x_i, \xi_2 = y_j) = u_i v_j, & i = \overline{1, n}, j = \overline{1, m}; \\ p(\xi_1 = x_i, \xi_3 = z_k) = u_i t_k, & i = \overline{1, n}, k = \overline{1, l}; \\ p(\xi_2 = y_j, \xi_3 = z_k) = v_j t_k, & j = \overline{1, m}, k = \overline{1, l}, \end{cases}$$

то вероятности их распределений будут удовлетворять неравенству

$$\sum_{k=1}^l t_k^2 \geq \sum_{i=1}^n \sum_{k=1}^l u_i^2 t_k^2 + \sum_{j=1}^m \sum_{k=1}^l v_j^2 t_k^2 - \sum_{i=1}^n \sum_{j=1}^m u_i^2 v_j^2.$$

Приведенные вероятностные интерпретации могут найти широкое применение в мульталгебраических системах [6, 8] и методах грануляции информации [1–5].

### Результаты и перспективы исследований

Необходимо подчеркнуть, что проведенное изучение метрических свойств разбиений направлено не столько на получение инструментария для оценки эффективности алгоритмов продуцирования классов эквивалентностей (в задачах кластеризации данных, сегментации изображений и т. п.), сколько на создание объективных предпосылок для синтеза методов интеллектуальной интерпретации фактор-множеств. В качестве типичного примера можно указать поиск компромисса между чрезмерной и недостаточной сегментацией изображений. В этом плане (точнее говоря, при стратифицированной грануляции информации) возникает необходимость исследования метрических свойств частично упорядоченных фактор-множеств — вложенных разбиений. Определенную перспективу составляет и поиск на разбиениях так называемых подобных метрик первого типа ( $\rho'(a, \beta) = \text{ср}''(f(a), f(\beta))$ ,  $\text{с} = \text{const}$ ) и второго типа ( $\rho'(f(a), f(\beta)) = \rho''(f(\gamma), f(\delta)) \Leftrightarrow \rho''(a, \beta) = \rho''(\gamma, \delta)$ ,  $a, \beta, \gamma, \delta \in \Pi_\Omega$ ). Поясним подробнее на примере.

Поскольку под энтропией понимается функционал вида  $H(q) = -\sum_{i=1}^n q_i \ln q_i$ , где  $\bar{q} = (q_1, \dots, q_n)$  — вообще говоря, некоторое дискретное распределение:  $\sum_{i=1}^n q_i = 1$ ,  $0 \leq q_i \leq 1$ ,  $i = \overline{1, n}$ , то, если даны два любых распределения  $\bar{u} = (u_1, \dots, u_n)$  и  $\bar{v} = (v_1, \dots, v_m)$ , или, что равносильно, в нашем случае (см. раздел 4)  $u_i = p(A_i)$ ,  $i = \overline{1, n}$ ,  $v_j = p(B_j)$ ,  $j = \overline{1, m}$ , тогда метрика энтропийного типа примет вид [15]:

$$\begin{aligned} \bar{h}(\bar{u}, \bar{v}) = h(\alpha, \beta) &= -\sum_{i=1}^n p(A_i) \ln p(A_i) - \sum_{j=1}^m p(B_j) \ln p(B_j) + \\ &+ 2 \sum_{i=1}^n \sum_{j=1}^m p(A_i \cap B_j) \ln p(A_i \cap B_j) = \\ &= H(\alpha) + H(\beta) - 2H(\alpha \cap \beta). \end{aligned}$$

Сравнивая полученное равенство с (19), можно отметить явное сходство между ними. Напрашивается вывод об общей формуле:

$$\rho(\alpha, \beta) = \sum_{i=1}^n f(p(A_i)) + \sum_{j=1}^m f(p(B_j)) - 2 \sum_{i=1}^n \sum_{j=1}^m f(p(A_i \cap B_j)).$$

В нашем случае используется  $f(x) = x^2$ . В энтропийной метрике —  $f(x) = -x \ln x$ . Возможно, допустимо использовать другие функции  $f(x)$  для получения новых метрик на множестве конечных разбиений измеримого множества  $\Omega$ . Ответы на вопросы, каковы эти функции и каким свойствам они должны удовлетворять, остаются открытыми.

**Список литературы:** 1. *Bargiela A., Pedrycz W.* Granular computing: an introduction. Boston, Kluwer Academic Publishers, The Kluwer International Series in Engineering and Computer Science. Vol. 717. 2002. 478 P. 2. *Yao Y.Y.* A partition model of granular computing // Transactions on Rough Sets I / J.F. Peters, et al. (Eds.). — Lecture Notes in Computer Science. — Berlin Heidelberg: Springer-Verlag. — Vol 3100. — 2004 — P. 232-253. 3. *Yao J.T., Yao Y.Y.* Induction of classification rules by granular computing // Rough Sets and Current Trends in Computing / J.J. Alpigini, et al. (Eds.). — Lecture Notes in Artificial Intelligence. — Berlin Heidelberg: Springer-Verlag. — Vol. 2475. — 2002. — P. 331-338. 4. *Yao Y.Y.* Perspectives of granular computing // Proceedings of IEEE International Conference on Granular Computing. — Vol. 1. — 2005. — P. 85-90. 5. *Doherty P., Lukaszewicz W., Szalas A.* Information granules for intelligent knowledge structures // Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing / G. Wang, et al. (Eds.). — Lecture Notes in Artificial Intelligence. — Berlin Heidelberg: Springer-Verlag. — Vol. 2639. — 2003. — P. 405-412. 6. *Машталир В.П., Шляхов В.В.* Индуцированная согласованность отношений в задачах грануляции информации // Бионика интеллекта. — №1 (64). — 2006. — С. 19-26. 7. *Lin T.Y.* Granular computing (Structures, representations, and applications) // Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing / G. Wang, et al. (Eds.). — Lecture Notes in Artificial Intelligence. — Berlin Heidelberg: Springer-Verlag. — Vol. 2639. — 2003. — P. 16-24. 8. *Машталир В.П., Шляхов В.В.* Свойства мультиталгебраических систем в задачах компаративного распознавания // Кибернетика и системный анализ. — 2003. — №6. — С. 12-32. 9. *Jain A.K., Murty M.N., Flynn P.J.* Data clustering: a review // ACM Computing Surveys. — Vol. 31, No. 3. — 1999. — P. 264-323. 10. *Jiang X., Marti C., Irniger C., Bunke H.* Image segmentation evaluation by techniques of comparing clusterings // Image

Analysis and Processing / F. Roliand, S. Vitulano (Eds.). — Lecture Notes in Computer Science. — Berlin Heidelberg: Springer-Verlag. — Vol. 3617. — 2005. — P. 344-351. 11. *Jiang X.* Performance evaluation of image segmentation algorithms. // Handbook of Pattern Recognition and Computer Vision / C.H. Chen, P.S.P. Wang (Eds.). — Singapore: World Scientific. — 2005. — P. 525-542. 12. *Jiang X., Marti C., Irniger C., Bunke H.* Distance measures for image segmentation evaluation // EURASIP Journal on Applied Signal Processing. — Vol. 2006. — Article ID 35909. — 10 P. 13. *Jiang X., Marti C., Irniger C., Bunke H.* Image segmentation evaluation by techniques of comparing clusterings. // Image Analysis and Processing / F. Roliand, S. Vitulano (Eds.). — Lecture Notes in Computer Science. — Berlin Heidelberg: Springer-Verlag. — Vol. 3617. — 2005. P. 344-351. 14. *Dongen S.* Performance criteria for graph clustering and Markov cluster experiments // Technical Report: INS-R0012. — Amsterdam: CWI (Centre for Mathematics and Computer Science) — 2000. 15. *Meila M.* Comparing clusterings by the variation of information // Learning Theory and Kernel Machines. — Lecture Notes in Computer Science. — Berlin Heidelberg: Springer-Verlag. — Vol. 2777. — 2003. — P. 173-187. 16. *Stehling R.O., Nascimento M.A., Falcao A.X.* MiCRoM: A metric to compare segmented images // VISUAL 2002 / S.-K. Chang, Z. Chen, S.-Y. Lee. (Eds.). — Lecture Notes in Computer Science. — Berlin Heidelberg: Springer-Verlag. — Vol. 2314. — 2002. P. 12-23. 17. *Mashtalir V., Mikhnova E., Shlyakhov V., Yegorova E.* A Novel metric on partitions for image segmentation. // Proceedings of IEEE International Conference on Advanced Video and Signal Based Surveillance. *avss*. — 2006. — P. 18. 18. *Kinoshenko D., Mashtalir V., Shlyakhov V.* A partition metric for clustering features analysis // International Journal «Information Theories and Applications» — Vol. 14, No 3. — 2007. — P. 230-236.

*Поступила в редколлегию 18.01.07*