

УДК 681.51:007.5

В. А. Гороховатский

Харьковский национальный университет радиоэлектроники
проспект Ленина, 14, 61166 Харьков, Украина

Фильтрация ложных соответствий дескрипторов ключевых точек на основе анализа геометрических данных

Рассмотрены вопросы усовершенствования структурных подходов при анализе данных в системах компьютерного зрения. Исследованы методы фильтрации на основе пространственной информации как путь устранения ложных признаков. Проведена формализация задачи, предложены подходы для кластеризации в пространстве геометрических преобразований, обсуждаются результаты компьютерного моделирования.

Ключевые слова: компьютерное зрение, сопоставление визуальных объектов, ключевые точки, дескрипторы, соответствия, фильтрация, геометрические преобразования, кластеризация.

Анализ проблемы и постановка задачи

Ряд методов распознавания изображений строится на анализе соответствий *ключевых точек* (КТ) визуального объекта, которые описаны посредством *дескрипторов* [1–4]. Дескриптор представляет собой пару $\lambda = (v, c)$, где $v \in R^n$ — вектор действительных чисел размерности n ; $c = (x, y)$ — пространственные координаты КТ. Такое структурное описание позволяет решать задачу распознавания в условиях частичного представления анализируемых объектов и устранять ложные элементы описания.

Пусть U — некоторый универсум дескрипторов $\lambda \in U$. Обозначим Λ^1, Λ^2 — конечные множества дескрипторов, $\Lambda^1, \Lambda^2 \subset U$, $\Lambda^1 = \{\lambda_i^1\}_{i=1}^{\mu_1}$, $\Lambda^2 = \{\lambda_k^2\}_{k=1}^{\mu_2}$, μ_1, μ_2 — мощности множеств. Пусть $\Omega: \Lambda^1 \rightarrow \Lambda^2$ некоторое отображение из множества Λ^1 во множество Λ^2 . В результате применения Ω получаем конечное множество $\theta = \{\theta_j\}_{j=1}^m$, $\theta \subset \Theta$, парных *соответствий* элементов $\lambda_i^1 \rightarrow \lambda_k^2$, $\lambda_i^1 \in \Lambda^1$, $\lambda_k^2 \in \Lambda^2$ (m — количество, Θ — универсум соответствий). Множество θ представимо в виде элементов бинарной матрицы

© В. А. Гороховатский

$$\theta_j = \theta(i, k) = \begin{cases} 1, & \lambda_i^1 \rightarrow \lambda_k^2, \\ 0, & \text{else.} \end{cases} \quad (1)$$

Число сформированных соответствий m удовлетворяет неравенству $0 \leq m \leq \mu_{12}$, где $\mu_{12} = \mu_1 \mu_2$. Построение соответствия θ_j означает получение некоторого отношения для пары элементов из двух рассматриваемых множеств. Основной способ формирования соответствий — анализ близости двух дескрипторов из U , который определяется как некоторая мера сходства $d(\lambda_i^1, \lambda_k^2)$ или метрика $\rho(\lambda_i^1, \lambda_k^2)$. На основе подсчета соответствий можно получить оценку сходства между множествами Λ^1, Λ^2 в виде:

$$d(\Lambda^1, \Lambda^2) = [\mu_*]^{-1} \sum_{i=1}^{\mu_1} \sum_{k=1}^{\mu_2} \theta(i, k) = [\mu_*]^{-1} \sum_{j=1}^m \theta_j, \quad (2)$$

где μ_* — нормировочный коэффициент. В случае $\mu_* = \mu_{12}$ выражение (2) можно считать метрикой [5]. В задачах распознавания визуальных объектов, когда одно из множеств, например Λ^1 , считают эталоном, а для каждого $\lambda_i^1 \in \Lambda^1$ формируется только одно соответствие, нормировка в выражении (2) с коэффициентом $\mu_* = \mu(1)$ приводит к мере сходства, отражающей долю элементов из Λ^1 , нашедших соответствие в Λ^2 . В общем случае структурная мера сходства множеств (2) оценивает долю мощности общего подмножества этих множеств.

Переход в пространство КТ существенно сокращает объем информации о визуальных объектах, часто обеспечивая при этом достаточно высокий уровень надежности при распознавании [6, 9], так как посредством КТ удается сохранить важную для распознавания информацию. Одним из путей достижения хорошего качества распознавания в современных методах SIFT [2] есть анализ альтернативных соответствий для каждого из дескрипторов. Однако даже посредством дополнительной обработки часто не удается полностью устранить ложные элементы, т.к. дескрипторы компонент бывают близкими для разных объектов и фона. Основной причиной есть то, что мера сходства (2) не отражает важной информации о геометрической структуре состоящих из компонент объектов и может включать ложные соответствия. Особенно усложняется ситуация в условиях действия сложного фона. На рис. 1 приведено оверлейное представление соответствий для одного из объектов базы soil-20 [10], когда на фоне представлен совершенно другой объект. Точки эталона для удобства показаны темными кружочками, а точки фона — светлыми. Здесь наблюдается наличие значительного числа ложных соответствий, что может привести к принятию неправильного решения.

Цель работы — применение подходов, связанных с кластеризацией на множестве значений преобразований ключевых точек в целях устранения ложных признаков при сопоставлении изображений.

Задачи исследования — построение отображения из множества ключевых точек во множество значений их преобразований на основе соответствий дескрип-

торов объекта и эталона; анализ и выбор вариантов кластеризации на множестве преобразований; изучение свойств, особенностей применения, сравнение с другими подходами, оценка эффективности предлагаемых способов обработки.

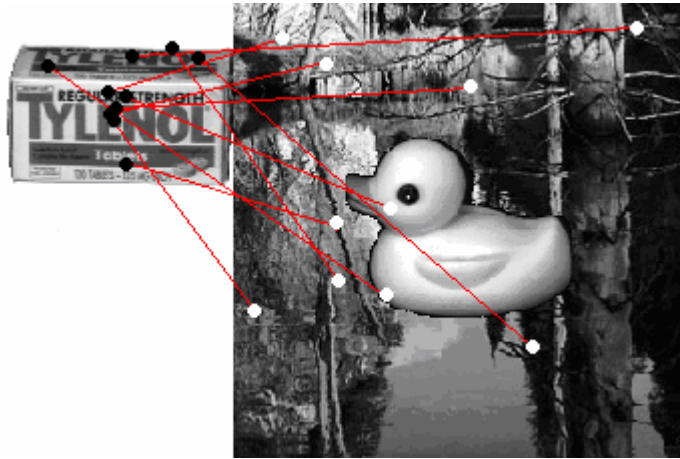


Рис. 1. Ложные соответствия при сопоставлении объектов

Обработка соответствий с использованием геометрической информации

Основным подходом к устранению ложных соответствий есть использование пространственной (геометрической) информации, содержащейся в координатах $s = (x, y)$ дескрипторов, на основе которых получено множество соответствий. Таким путем можно осуществить фильтрацию соответствий, установленных на основе анализа только значений дескрипторов. Использование геометрической информации лежит в основе построения ряда известных подходов [1–4]. Так, метод геометрического хеширования состоит в получении для множества КТ набора инвариантных признаков на основе координат триад точек и поиске близких значений инвариантов на множестве КТ сопоставляемого изображения [1]. Таким путем осуществляется переход из пространства координат КТ в пространство двухкомпонентных вещественных векторов, инвариантных к аффинным преобразованиям. Другой путь может быть связан с формированием пространственных *кластеров* на множестве соответствий, формирующих группы точек с одинаковыми значениями геометрических преобразований. В основе этого лежит одинаковость значений преобразований для всех точек объекта. Кластеризация в отличие от инвариантных признаков позволяет дополнительно осуществить оценку параметров преобразований, необходимую в задачах нормализации изображений [9]. Таким образом, оценка сходства, параметров и принятие решения о соответствии объектов могут быть осуществлены на основе комплексного использования негеометрических и геометрических данных, что в целом повышает надежность распознавания.

Рассмотрим структуру отдельной компоненты вектора соответствий $\theta = \{\theta_j\}_{j=1}^m$ в виде $\theta_j = (v_{ij}, c_{ij}, v_{kj}, c_{kj})$. В общем виде фильтрацию соответствий

можно определить как отображение $F : \Theta \rightarrow \Theta$, которое для конкретного множества соответствий приобретает вид $F : \theta \rightarrow \theta_F$, где $\theta_F \subseteq \theta$ — подмножество θ , элементы которого обладают заданными свойствами, причем $\theta, \theta_F \subset \Theta$. Каждый элемент θ_j в рамках классической теории нормализации [9] дает возможность оценить величины геометрических преобразований по координатам c_{ij}, c_{kj} . Например, на базе одного значения θ_j есть возможность определить величины смещений либо параметр однопараметрической группы (поворот или масштаб). Соответствие пары точек (θ_a, θ_b) при условии $c_{ia} \neq c_{ib}, c_{ka} \neq c_{kb}$ является основой для вычисления величин трех преобразований: поворота, масштаба и смещений. Шесть параметров аффинной группы можно определить по трем соответствиям $(\theta_a, \theta_b, \theta_l)$. Важным моментом при построении геометрических соответствий есть однозначность оценок величин параметров преобразований, которая устанавливается системой координат и последовательностью их выполнения, так как в общем случае отсутствует коммутативность преобразований, например, между смещениями и поворотом. По этой причине порядок анализа последовательности преобразований (поворот, масштаб, смещения) должен быть фиксированным.

Отметим, что анализ соответствий также непосредственно приводит к расширению используемого пространства признаков. Например, при анализе пар (θ_a, θ_b) возможное количество рассматриваемых вариантов равно числу сочетаний $t = C_m^2$. Вычисление параметров на базе пары (θ_a, θ_b) реализуется в виде отображения $\Psi : \theta \times \theta \rightarrow G$, где G — множество значений параметров преобразований, $\theta \times \theta$ — декартово произведение. В компьютерном зрении множество G трактуют как многопараметрическую группу [9].

Пусть $g = (g_1, g_2, g_3, g_4)$, $g \in G$ — вектор значений преобразований, полученный при анализе пары (θ_a, θ_b) : g_1 — поворот, g_2 — масштаб (однородный), g_3, g_4 — смещения. При этом вначале определяются параметры смещения, а затем — поворот и масштаб. При формировании объекта применяется обратный порядок последовательности преобразований. Матрица преобразований в однородных координатах имеет вид:

$$A_g = \begin{pmatrix} g_2 \cos(g_1) - g_2 \sin(g_1) & g_2 g_3 \cos(g_1) - g_2 g_4 \sin(g_1) \\ g_2 \sin(g_1) & g_2 \cos(g_1) & g_2 g_4 \cos(g_1) + g_2 g_3 \sin(g_1) \\ 0 & 0 & 1 \end{pmatrix}.$$

В целях упрощения дальнейшего анализа необходимо выполнить нормировку компонент вектора g , приведя значения признаков к одинаковому диапазону, например, к отрезку $[0,1]$. Один из вариантов: $g_k^* = (g_k - \min_k g_k) / (\max_k g_k - \min_k g_k)$, где g_k^* — нормированные значения. Второй вариант нормировки выглядит как $g_k^* = g_k / \|g\|$, где $\|g\|$ — норма вектора преобразований, например,

$\|g\| = \sqrt{\sum_{k=1}^4 g_k^2}$ — евклидова норма. В этом случае получаем нормированное векторное пространство. Третий вариант нормировки имеет вид $g_k^* = g_k / \max |g_k|$ и приводит значения к отрезку $[-1,1]$ и т.д. Далее все величины g_k считаем нормированными.

Вторым актуальным моментом есть мера для оценки близости двух векторов g^i и g^j , $g^i = \{g_k^i\}_{k=1}^4$, $g^j = \{g_k^j\}_{k=1}^4$. Сравнение можно осуществить на основе метрики $\rho(g^i, g^j)$, меры сходства $d(g^i, g^j)$ или нормы $\|g^i\|$. Наряду с классическими метриками для векторов типа евклидовой, можно применить меру вида $\rho(g_k^i, g_k^j)$, связанную с анализом отдельных компонент, что точнее учитывает отклонения разных типов преобразований. Например, условие близости на основе такой меры может выглядеть как $\forall k = \overline{1,4} : \rho(g_k^i, g_k^j) \leq \delta$ или $\sum_{k=1}^4 \rho(g_k^i, g_k^j) \leq \delta$.

Кластеризация на множестве преобразований

Рассмотрим некоторые пути формирования кластеров в виде групп близких значений преобразований, полученных по множеству соответствий.

1. Группировка по величине нормы на основе медианы. Сведем значение вектора g к его норме $\|g\|$. Расположим значения норм для m полученных соответствий $\|g^1\|, \dots, \|g^m\|$ в порядке возрастания значений $\|g^1\| \leq \|g^2\| \leq \dots \leq \|g^m\|$ и определим на основе данной последовательности медиану g^* (центральный элемент). Далее введем порог $\delta_g > 0$ для установления эквивалентности двух значений норм и подсчитаем количество m^* таких элементов из ранжированной последовательности, значения которых близки к медиане в некотором смысле: $\| \|g^i\| - \|g^*\| \| \leq \delta_g$. Полученная оценка для m^* , например, в виде отношения m^*/m , может быть основой для принятия итогового решения о соответствии объекта и эталона. Описанный вариант обработки соответствий основан на величине нормы вектора g , что в целом упрощает анализ за счет перехода к одномерной задаче, зато приводит к стиранию различия между значениями отдельных параметров g_k . Кроме того, в общем случае существенно различающиеся векторы преобразований могут иметь близкие нормы.

2. Группировка на основе просеивания (решета) [5]. Вводится парное сравнение объектов g^i из выборки $\{g^i\}$ и составляется для каждого объекта g^i кортеж $g^{N(i)} = \{g^{i1}, \dots, g^{in}\}$ «похожих» на него объектов. Далее формируется некоторая оценка $q(g^{N(i)}, g^i)$ «разброса» элементов кортежа относительно g^i . Затем выборка $\{g^i\}$ упорядочивается, считая, что $g^i \leq g^j$, если $q(g^{N(i)}, g^i) \leq q(g^{N(j)}, g^j)$. Классификация по принципу решета: первый объект из отсортированного списка вместе со своим кортежем формирует первый кластер, после чего элементы кор-

тежа удаляются из выборки; первый из оставшихся объектов упорядоченной выборки и его кортеж формируют второй таксон; элементы выборки, которые не были удалены на предыдущем шаге, удаляются из выборки; процедура классификации продолжается до тех пор, пока вся выборка или наперед заданная ее доля не будет расклассифицирована. Такая классификация приводит к покрытию классами, а не к разбиению, так как таксоны, несмотря на последовательное исключение элементов в процессе анализа могут пересекаться из-за того, что вначале все элементы участвуют в упорядочении выборки, и порядок после их исключения не пересматривается. Распространенными на практике вариантами подхода, не связанными с конкретным видом распределений классов, есть алгоритм Уишарта [5], построенный на основе правила ближайших соседей, когда кортеж формируется из фиксированного числа n ближайших элементов, а в качестве $q(g^{N(i)}, g^i)$ используется расстояние $\rho(g^i, n)$ от элемента g^i до самого дальнего из n его ближайших соседей. Параметрами семейства алгоритмов просеивания являются: способ формирования кортежа, мера сходства на множестве объектов, функция q оценки разброса.

3. Иерархическая классификация [5]. Она реализует разбиение выборки $\{g^i\}$, $g^i \in G$, на непересекающиеся группы и построена на некоторой метрике (или мере сходства) $\rho(G_1, G_2)$ для кластеров $G_1 \subset G$, $G_2 \subset G$. Например, иерархическая агломеративная кластеризация состоит в объединении выбранной пары кластеров G_1^*, G_2^* на каждом шаге процедуры в соответствии с выражением $(G_1^*, G_2^*) = \arg \min_{G_1 \neq G_2} \rho(G_1, G_2)$, где G_1, G_2 — классы предыдущего шага. В целях ускорения вычислений в процедурах иерархической классификации часто применяют систему порогов, позволяющую объединять классы, мера близости которых не превосходит заданное на данном уровне значение порога. Разработана серия критериев для оценки однородности сформированных классов. Одним из наиболее универсальных показателей есть статистический разброс $q(S) = \sum_{X \in S} \|X - Z\|^2$ класса S , где $Z = [\mu(S)]^{-1} \sum_{X \in S} X$ — центр класса, $\mu(S)$ — число элементов, $\|X - Z\|^2$ — квадрат евклидова расстояния. Мерой сходства между классами G_1, G_2 может быть приращение статистического разброса при их объединении $\rho(G_1, G_2) = q(G_1 \cup G_2) - q(G_1) - q(G_2)$, что является метрикой. На основе таких критериев строятся функционалы качества разбиения (например, корреляционное отношение), применяемые как условие завершения процедуры кластеризации. Другими критериями могут быть $\min_{G_1 \neq G_2} \rho(G_1, G_2) > \delta$, где δ — некоторый порог, а также достижение заданного количества групп [5]. Подробный анализ методов, критериев, функционалов качества для разнообразия подходов кластеризации приведен в [8].

Близкий к кластеризации подход можно реализовать на основе градуировки (дискретизации) значений компонент вектора g . Этот путь соответствует построению кластеров на основе преобразования Хафа [7]. На каждом шаге такой

процедуры определяется и инкрементируется дискретная точка четырехмерного пространства преобразований, к которой относится анализируемый вектор g^i . В результате в аккумуляторе накапливаются количества значений преобразований, максимум среди которых дает возможность оценить параметры. Кластеризация обладает большей точностью по сравнению с оценкой на основе преобразования Хафа, однако, уступает ему по быстродействию. Снижение точности может повлиять на качество распознавания.

Все рассмотренные подходы формируют некоторое результирующее число отфильтрованных соответствий $0 < m^* \leq m$. Понятно, что величина m^* должна быть значимой, т.е. удовлетворять неравенству $m^* > \delta_m$, где δ_m — некоторый порог, который задается априорно либо вычисляется в процессе фильтрации адаптивно к используемым мерам. Как правило, δ_m означает минимальное число (или долю) соответствий, допустимое для принятия решения о сходстве объектов. Особенностью применения кластеризации на множестве соответствий есть анализ пар элементов, поэтому в результате удаляются точки, которые не вошли в отфильтрованное множество, образующее максимальный кластер. Отметим, что кластеризацию можно применить также отдельно для множества эталонов в целях фильтрации одинаковых КТ.

Эксперименты

В целях качественного анализа обсуждаемого подхода для практических наборов признаков проведены компьютерные эксперименты на реальных изображениях. Моделирование осуществлялось в системах программирования Matlab и Visual C++. В качестве компонент использовались множества КТ, полученные применением варианта детектора SIFT [2] (вектор размерностью 128) для базы данных, содержащей 30 полутоновых изображений аквариумных рыбок размером 100×100 пикселей [6]. На рис. 2 показано изображение одного из эталонов при преобразовании поворота в 45° , размещенные в оверлейном режиме дескрипторы (темные точки) и соответствия между ними в виде отрезков. В данном случае все 42 дескриптора изображения образовали кластер с параметром 45° , и ложных соответствий (в отличие от рис. 1) не имеется. Незначительное число ложных соответствий (3–5) наблюдается при сопоставлении разных эталонов из базы. Ситуация существенного числа ложных соответствий в большей степени характерна для помех фона. Например, на рис. 1 число таких соответствий равно 12. Эксперименты показали, что все они устраняются применением иерархической кластеризации при пороге $\delta \approx 0,2$.

В сравнительном плане оценивались свойства классического метода SIFT и модификации с применением иерархической кластеризации на основе евклидовой метрики. Оба метода в идеальном случае (без помех) обеспечивают безошибочное распознавание для базы аквариумных рыбок в условиях геометрических преобразований (поворот: $0 \dots 180^\circ$, масштаб: $0,5 \dots 8$, смещения — произвольные в поле зрения). Однако модифицированный метод намного увереннее принимает решение о классе. Для классического подхода доля полученных соответствий по отношению к числу соответствий со «своим» эталоном на множестве изображений ба-

зы колеблется в пределах 0,4–1,0; для модификации с применением кластеризации — 0,13–1,0. В то же время отношение ближайшего локального максимума числа соответствий к величине глобального максимума достигает: для классического подхода — 0,9; для модификации — только 0,4, что говорит о существенно большей достоверности решений модифицированного подхода. Это подтверждают также и эксперименты по оценке помехозащищенности.

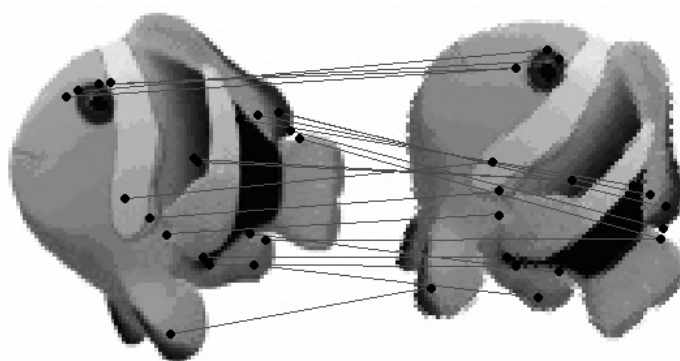


Рис. 2. Соответствия КТ при геометрических преобразованиях поворота

Сравнительная оценка помехозащищенности показывает значительные преимущества методов, включающих кластеризацию геометрических данных, по отношению к методам распознавания на основе классического детектора, использующего одиночные соответствия дескрипторов. Если классический детектор обеспечивает высокую вероятность распознавания $P = 0,98$ лишь при энергетическом соотношении сигнал–шум $\eta \leq 5$, то подход на основе иерархической кластеризации обеспечивает эту же вероятность при значительно большем уровне помех $\eta \leq 8$. Отметим, что при уровне $\eta = 8$ вероятность распознавания для классического метода не превышает значения $P = 0,6$, что неприемлемо для большинства практических задач.

Выводы

Фильтрация соответствий характерных признаков, полученных при сопоставлении визуальных объектов, позволяет улучшить достоверность установления их связи, и как результат — увеличить вероятность распознавания объектов в ситуациях влияния фона и локальных помех. Основой такого усовершенствования является геометрическая информация, заключающаяся в пространственном расположении признаков.

Впервые показано, что фильтрация соответствий ключевых точек изображений является эффективным способом устранения ложных элементов, связанных с локальными помехами и фоном. Изучены особенности и свойства методов кластеризации применительно к задаче фильтрации в пространстве геометрических преобразований.

Практически важным является сравнение различных принципов кластеризации, получение конкретных характеристик достоверности и помехозащищенности

на экспериментальной базе изображений, что подтверждает целесообразность применения предложенных подходов в задачах компьютерного зрения.

Перспективы исследования состоят в разработке схем по выбору оптимальных значений параметров кластеризации, а также оценка характеристик методов при распознавании объектов в условиях одновременного действия помех разного типа.

1. *Шапиро Л.* Компьютерное зрение / Л. Шапиро, Дж. Стокман; пер. с англ. — М.: Бином, 2006. — 752 с.

2. *Lowe D.G.* Distinctive Image Features from Scale-Invariant Keypoints / D.G. Lowe // International Journal of Computer Vision. — 2004. — 60, N 2. — P. 91–110.

3. *Путятин Е.П.* Распознавание изображений в пространстве инвариантных локальных признаков / Е.П. Путятин, В.А. Гороховатский, С.В. Кузьмин // Радиоэлектроника и информатика. — 2006. — № 1(32). — С. 69–73.

4. *Гороховатский В.А.* Иерархия пространственных отношений структурных признаков в задачах сопоставления визуальных объектов / В.А. Гороховатский // Системи управління, навігації та зв'язку. — К.: Центральний наук.-досл. інститут навігації і управління, 2008. — Вип. 3(7). — С. 85–89.

5. Прикладная статистика: Классификация и снижение размерности: Справ. изд. / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин; под ред. С.А. Айвазяна. — М.: Финансы и статистика, 1989. — 607 с.

6. *Гороховатский В.А.* Структурное распознавание изображений на основе моделей голосования признаков характерных точек / В.А. Гороховатский, Е.П. Путятин // Реєстрація, зберігання і оброб. даних. — 2008. — Т. 10, № 4. — С. 75–85.

7. *Duda R.O.* Pattern Classification / R.O. Duda, P.E. Hart, D.G. Stork. — 2 ed. — Wiley, 2000. — 738 p.

8. *Мандель И.Д.* Кластерный анализ / И.Д. Мендель. — М.: Финансы и статистика, 1988. — 176 с.

9. *Путятин Е.П.* Методи та алгоритми комп'ютерного зору: навч. посіб. / Е.П. Путятін, В.О. Гороховатський, О.О. Матат. — Харків: ТОВ «Компанія СМІТ», 2006. — 236 с.

10. <http://www1.cs.columbia.edu/CAVE/software/softlib/coil-20.php>

Поступила в редакцію 09.01.2009