

СНИЖЕНИЕ РАЗМЕРНОСТИ ВЕКТОРОВ ПРИЗНАКОВ ПО КРИТЕРИЯМ МУЛЬТИКОЛЛИНЕАРНОСТИ

Н.Е. Козин¹, В.А. Фурсов²

¹ Самарский государственный аэрокосмический университет им. академика С.П. Королева, Самара, Россия,

² Институт систем обработки изображений РАН, Самара, Россия

Аннотация

В работе рассматривается возможность снижения размерности векторов признаков для задачи распознавания образов. Снижение размерности производится за счет отбрасывания менее информативных признаков. В качестве такого критерия информативности предлагается показатель диагонально преобладания. Обсуждается эффективность данного показателя по сравнению с другими мерами мультиколлинеарности с вычислительной точки зрения. Предлагаются алгоритмы, реализующие предложенную идею при решении задачи распознавания изображений.

Ключевые слова: мультиколлинеарность, показатель диагонального преобладания, снижение пространства признаков.

Введение

При построении классификаторов одним из важнейших этапов является выбор системы признаков [1]. Удачный выбор признаков обеспечивает высокое качество распознавания. Обычно признаки стараются выбрать таким образом, чтобы высокое качество распознавания обеспечивалось при минимальной размерности признакового пространства. Это особенно актуально, если задача распознавания должна решаться с высокой оперативностью, возможно, в реальном времени.

Для обеспечения этого требования в число признаков должны отбираться наиболее информативные. Известно два подхода к решению задачи отбора информативных признаков: наращивание числа признаков и редукция (сокращение) числа признаков [2]. В последнем случае задается некоторая, заведомо избыточная система признаков, а затем осуществляется исключение из нее малоинформативных.

При решении задачи распознавания изображений в качестве признаков часто используются отсчеты (значения яркостей точек) самого изображения, а образом является вектор, составленный из отсчетов изображения распознаваемого объекта. В этом случае анализ информативности признаков заключается в выявлении информативных областей (зон семантической значимости) на изображениях.

Интуитивно ясно, что наиболее информативными будут те области на изображениях, в которых имеют место существенные различия, связанные с принадлежностью образов разным классам [3]. Следовательно, алгоритм отбора информативных признаков (пикселей изображения) может состоять в формировании векторов образов, обладающих слабой мультиколлинеарностью (сопряженностью).

В настоящей работе предлагается и исследуется алгоритм сокращения размерности признакового пространства в задаче распознавания изображений. Идея алгоритма заключается в выделении наиболее информативных областей на изображениях путем

вычисления простых достаточных оценок мультиколлинеарности векторов образов.

1. Постановка задачи

Пусть все изображения обучающей выборки имеют одинаковые размеры $N_1 \times N_2$, так что каждое изображение представляется $N_1 N_2 \times 1$ -вектором, компонентами которого являются отсчеты изображений, «развернутые» по строкам или столбцам [4], [5]. Предположим также, что число изображений в обучающей выборке равно M . Обозначим $N = N_1 \times N_2$ и составим информационную $M \times M$ -матрицу Грама

$$\mathbf{C} = \mathbf{U}^T \mathbf{U}, \quad (1)$$

где \mathbf{U} $N \times M$ -матрица, столбцами которой являются $N \times 1$ -векторы образов \mathbf{u}_i , $i = \overline{1, M}$ обучающих объектов.

Поставим задачу из матрицы \mathbf{U} построить матрицу \mathbf{X} меньшей размерности $n \times M$ ($n < N$), содержащую M $n \times 1$ -векторов \mathbf{x}_i , $i = \overline{1, M}$, полученных путем исключения из векторов \mathbf{u}_i , $i = \overline{1, M}$ одноименных малоинформативных компонентов. Технически эта задача сводится к исключению из матрицы \mathbf{U} некоторого (возможно значительного) числа строк, являющихся источником сильной мультиколлинеарности, что по существу означает исключение из всех обучающих изображений неинформативных областей. Заметим, что соответствующая информационная матрица

$$\mathbf{A} = \mathbf{X}^T \mathbf{X} \quad (2)$$

также будет иметь размерность $M \times M$.

Наиболее полной характеристикой мультиколлинеарности векторов являются меры, связанные с собственными значениями определенной в (2) матрицы \mathbf{A} — $\lambda_i(\mathbf{A})$, $i = \overline{1, M}$. В частности, известны следующие меры мультиколлинеарности.

1. Определитель: $\det(\mathbf{A}) = \prod_{i=1}^M \lambda_i$.

2. Спектральное число обусловленности:

$$K(\mathbf{A}) = \frac{\lambda_{\max}}{\lambda_{\min}}.$$

3. Минимальное собственное значение: $\lambda_{\min}(\mathbf{A})$.

4. Показатель диагонального преобладания:

$$\phi = \frac{\text{tr}^2 \mathbf{A}}{\text{tr} \mathbf{A}^2} = \left(\sum_{i=1}^M a_{ii} \right)^2 / \sum_{i,j=1}^M a_{ij}^2 = \left(\sum_{i=1}^M \lambda_i \right)^2 / \sum_{i=1}^M \lambda_i^2.$$

Наиболее простым в вычислительном отношении является мера ϕ – показатель диагонального преобладания. Однако его использование связано с существенными ограничениями. В работе [6] показано, что гарантированные оценки снизу для собственных значений в функции этого показателя существуют лишь в диапазоне:

$$M - 1 < \phi \leq M. \quad (3)$$

Ставится задача построить процедуру, которая бы позволяла с использованием любого из указанных показателей мультиколлинеарности выявлять наиболее информативные области на изображениях с целью формирования (редукции) признакового пространства сравнительно невысокой размерности без существенной потери качества распознавания.

2. Алгоритм определения информативных областей

По предположению (раздел 1) каждое из исходных изображений обучающей выборки имеет размеры $N_1 \times N_2$, т.е. представляется $N \times 1$ -вектором \mathbf{u} ($N = N_1 \times N_2$), а число изображений в обучающей выборке равно M (\mathbf{u}_i , $i = \overline{1, M}$). Предлагается следующий итерационный алгоритм сокращения размерности этих векторов.

1-й шаг. Работа алгоритма начинается с присваивания переменным n_k и m_k начальных значений (при $k=1$) размеров исходных изображений: $n_k = n_1 = N_1$, $m_k = m_1 = N_2$. Далее алгоритм реализуется в виде итерационной схемы, включающей следующую последовательность шагов.

2-й шаг. Значение k увеличивается на единицу. Исходные изображения обучающей выборки разбиваются на k^2 одинаковых прямоугольных фрагментов путем деления каждой стороны на k частей, и вычисляется размерность соответствующего фрагменту изображения вектора:

$$K = n_k \times m_k, \quad (4)$$

где $n_k = n_{k-1}/k$, $m_k = m_{k-1}/k$.

В случае наличия остатков от деления, результат округляется до ближайшего целого.

3-й шаг. Осуществляется проверка условия

$$K > M. \quad (5)$$

Если неравенство выполняется, переход к шагу 4, если нет – к шагу 6.

4-й шаг. Для каждого фрагмента, например, q -го, ($q = \overline{1, M}$) изображения обучающей выборки, сформированных на шаге 2, составляется $K \times M$ -матрица $\mathbf{X}_{q,k}$, с использованием которой формируется соответствующая информационная матрица

$$\mathbf{A}_{k,q} = \mathbf{X}_{k,q}^T \mathbf{X}_{k,q}. \quad (6)$$

5-й шаг. По информационной матрице вычисляется одна из указанных в разделе 2 мер мультиколлинеарности (1-4). Полученное значение присваивается всем отсчетам «своего» фрагмента на поле анализируемого изображения.

6-й шаг. Задается или определяется пороговое значение показателя (мультиколлинеарности). Все отсчеты (пиксели) изображения с равным ему или более высоким, чем заданный порог, значением включаются в число компонент всех векторов \mathbf{x}_i , $i = \overline{1, M}$ обучающих объектов.

Заметим, что схема формирования векторов образов из оставшихся отсчетов изображений не имеет значения. Важно, чтобы эта схема была одинаковой для всех M векторов. В результате реализации описанного выше алгоритма в зависимости от выбранного порога может быть достигнуто существенное снижение размерности признакового пространства.

3. Связь мер мультиколлинеарности, обоснование выбора

Важным, с точки зрения технологии распознавания, является ответ на следующий вопрос: какая мера мультиколлинеарности является предпочтительной. Для этого проведем сравнительные исследования указанных выше 4-х мер мультиколлинеарности. Для того, чтобы они были сравнимы, а результаты сравнения были наглядными, сделаем следующие дополнения.

1. Нормируем информационную матрицу \mathbf{A} таким образом, чтобы все ее диагональные элементы были равны единице. При этом $\text{tr} \mathbf{A} = M$.

2. Вместо указанного выше спектрального числа обусловленности $K(\mathbf{A}) = \lambda_{\max} / \lambda_{\min}$ введем в рассмотрение так называемое обратное число обусловленности:

$$K^{-1}(\mathbf{A}) = \frac{\lambda_{\min}}{\lambda_{\max}}. \quad (7)$$

3. Вместо показателя диагонального преобладания ϕ , учитывая ограничения (3), будем использовать приведенное (к интервалу $[0,1]$) значение показателя

$$\bar{\phi} = \phi - M + 1. \quad (8)$$

С учетом указанных дополнений все четыре рассматриваемые меры мультиколлинеарности оказываются в интервале $[0,1]$. При этом стремление всех

мер к единице соответствует уменьшению степени мультиколлинеарности векторов образов. Это означает, что по мере сокращения размерности векторов признаков следует ожидать увеличения всех (в т.ч. модифицированных) показателей.

Для сравнения эффективности показателей в соответствии с приведенным в разделе 2 алгоритмом проводился эксперимент по снижению размерности векторов признаков. На рисунке 1 приведены графики изменения показателей мультиколлинеарности в зависимости от размерности векторов образов (пороговых значений для отбора информативных данных).

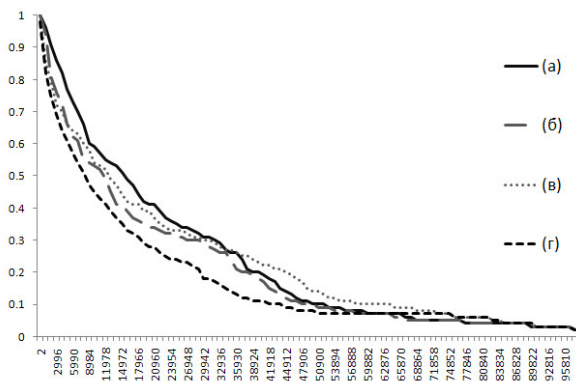


Рис. 1. Зависимость показателей $\det(\mathbf{A})$ (а), $K^{-1}(\mathbf{A})$ (б), $\lambda_{\min}(\mathbf{A})$ (в), $\bar{\phi}$ (г) от размерности векторов \mathbf{x}_i

Из рисунков видно, что характер изменения всех указанных показателей мультиколлинеарности при уменьшении размерности векторов образов \mathbf{x}_i примерно одинаков. На рисунке 2 для сравнения приведены также бинарные изображения полей информативности, полученные с использованием указанных 4-х показателей при достижении одинаковой размерности матрицы \mathbf{X} . Области, соответствующие наиболее информативным признакам, имеют наиболее суммарные значения показателей диагонального преобладания и выделены более светлыми тонами. Как и следовало ожидать, зоны фона, точки которых имеют высокую взаимную корреляцию (на большинстве фотографий фон имеет один цвет), являются наиболее темными, что свидетельствует об их малой информативности. Нетрудно заметить, что вид этих полей отличается также незначительно.

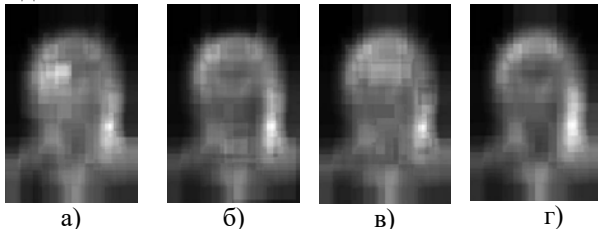


Рис. 2. Поля информативности, полученные с использованием показателей мультиколлинеарности: $\det(\mathbf{A})$ (а), $K^{-1}(\mathbf{A})$ (б), $\lambda_{\min}(\mathbf{A})$ (в), $\bar{\phi}$ (г)

Приведенные результаты исследований показывают, что все четыре показателя мультиколлинеарности дают сходные результаты и близки по эффективности с точки зрения индикации информативных областей изображений. Поэтому в информационных технологиях распознавания изображений предпочтительно использовать простой в вычислительном отношении приведенный показатель диагонального преобладания. При этом важным с точки зрения качества распознавания является выбор порогового значения этого показателя.

В следующем разделе приводятся оценки для границ приведенного показателя диагонального преобладания, а в заключительном разделе – результаты исследования качества распознавания при снижении размеров векторов образов.

4. Границы для приведенного показателя диагонального преобладания

Построим оценку сверху для суммарного значения показателя диагонального преобладания на поле изображения, которое может быть достигнуто в ходе реализации описанного алгоритма.

Максимальное значение приведенного показателя диагонального преобладания для любой нормированной информационной матрицы равно единице. Следовательно, в результате реализации описанного выше алгоритма максимально возможное значение суммарного показателя информативности на поле $\bar{\phi}_{\max}(\mathbf{A})$ может достигнуть следующего предельного значения:

$$\bar{\phi}_{\max}(\mathbf{A}) \leq k_{\max}, \tag{9}$$

где k_{\max} - максимальное число шагов разбиения изображения на фрагменты, определяемое условием (5). Получим соотношение для оценки k_{\max} .

В соответствии со схемой разбиения имеем

$$\begin{aligned} n_1 &= N_1 & m_1 &= N_2 \\ n_2 &= N_1/2 & m_2 &= N_2/2 \\ \dots & & \dots & \\ n_k &= N_1/k & m_k &= N_2/k \end{aligned} \tag{10}$$

Размерность вектора, описывающего фрагмент на k -м шаге в соответствии с (5), (10),

$$K = n_k \times m_k = \frac{N_1 N_2}{k^2}.$$

При $k = k_{\max}$ в соответствии с (5) должно выполняться условие

$$K = \frac{N_1 N_2}{k_{\max}^2} > M$$

или

$$k_{\max} < \sqrt{\frac{N_1 N_2}{M}}. \tag{11}$$

Откуда, с учетом (9), окончательно имеем

$$\bar{\phi}_{\max}(\mathbf{A}) < \sqrt{\frac{N_1 N_2}{M}}$$

Полученные оценки верхней границы для информативности по показателю диагонального преобладания могут быть полезны при сравнительной оценке информативности обучающих выборок.

5. Связь с качеством распознавания

Центральной задачей в общей проблеме распознавания является снижение размерности векторов образов без существенного снижения качества распознавания. В настоящем разделе приводятся эксперименты, иллюстрирующие эти возможности при использовании в качестве меры информативности данных приведенный показатель диагонального преобладания.

Эксперимент

В качестве объектов для распознавания были использованы изображения базы данных лиц FERET [7]. В обучающей выборке для каждого человека имелось одно изображение. Число изображений людей M варьировалось от 2 до 100. Каждое изображение имело размер 384x256 точек. Таким образом, исходное пространство признаков имело размерность $N = 98304$. В соответствии с описанным алгоритмом изображения разбивались на равные части. Для каждой части строилась информационная матрица и вычислялось значение показателя диагонального преобладания. На рисунке 3 показано исходное изображение и первые две итерации его разбиения на четыре и девять равных частей.

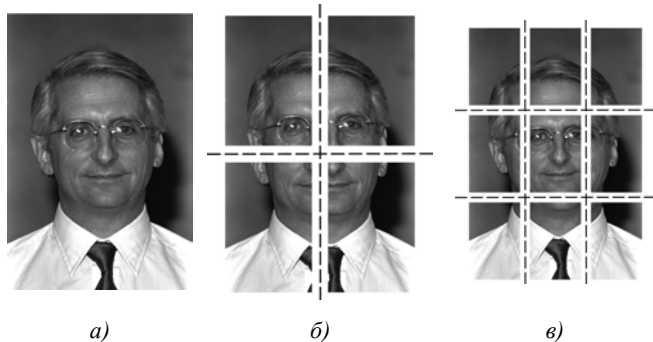


Рис. 3. а) исходное изображение; б) разбиение на 4 части; в) разбиение на 9 частей.

Для снижения размерности исходного пространства признаков из полученной матрицы \mathbf{X} были удалены строки, соответствующие наименьшим значениям информативности. На рисунке 4 представлены зависимости процента распознавания от процента удаленных отсчетов исходного пространства признаков для различных случаев числа изображений M в обучающей выборке.

Из графика видно, что ухудшение качества распознавания при увеличении обучающей выборки (соответственно и числа классов) происходит достаточно медленно. В частности, даже в случае 200 изображений различных людей размерность может

быть снижена до 80% без существенных потерь качества классификации.

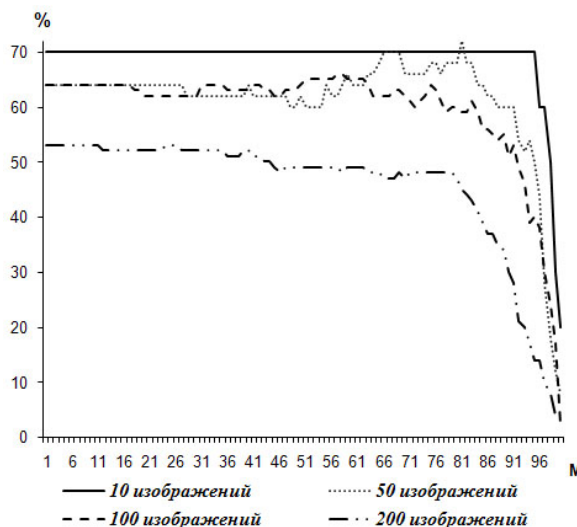


Рис. 4. Зависимость относительного числа правильных распознаваний на контрольной выборке от доли (в %) удаленных (по показателю $\bar{\phi}$) строк

Был проведен также эксперимент по распознаванию трехмерных изображений лиц с использованием базы данных Face Recognition Grand Challenge (FRGC) [8]. Трехмерные данные представлены для 275 человек с общим числом обучающих изображений равным 943 (для каждого человека число изображений варьируется от одного до восьми). В целом, тестовая база данных содержит 4007 изображений. Трехмерный образ имеет разрешение 640x480 точек и содержит как 2D данные текстуры, так и 3D данные формы. В данном эксперименте были использованы только трехмерные данные формы лица, фактически представляющие собой расстояния каждой точки лица от фиксированной позиции сенсора трехмерного сканера. Пример формы трехмерного лица приведен на рисунке 5.

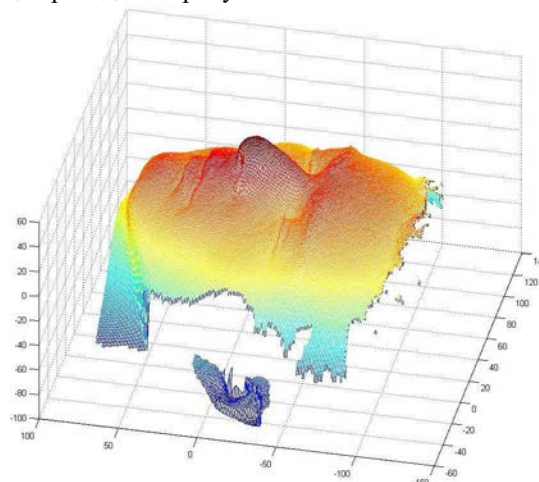


Рис. 5. – Трехмерное изображение лица базы данных FRGC

Приведенное изображение с определенной точки зрения можно считать псевдотрехмерным. Если рас-

смаивать значение z-координаты, отвечающей за глубину формы лица, яркость изображения в соответствующей точке с координатами x и y, то мы имеем аналог двумерного изображения. Таким образом, аналогично эксперименту по распознаванию двумерных изображений лиц, трехмерная модель лица итерационно разбивалась на равные регионы в плоскости xy. На рисунке 6 представлены результаты распознавания, соответствующие трем различным объемам тестовой выборки. Обучающая выборка в каждом случае была одна и та же и содержала 275 изображений.

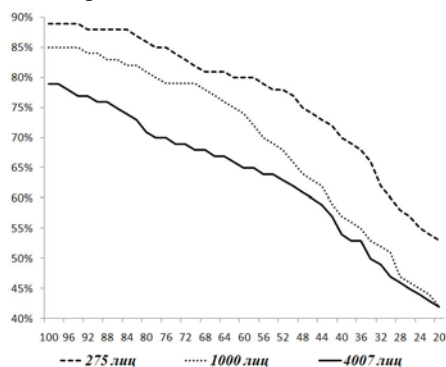


Рис. 6. Зависимость относительного числа правильных распознаваний трехмерных лиц от доли (в %) исходной размерности векторов образов (порога для показателя $\bar{\phi}$)

Как видно из графиков, в отличие от случая распознавания двумерных изображений лиц, при применении трехмерных моделей возможно удаление не более трети исходной области без значительных потерь качества. Это свидетельствует о том, что форма лицевой части головы человека содержит более равномерно распределенную информацию.

Заключение

Подведем итог. Определитель, число обусловленности и минимальное собственное значение являются достаточно полными характеристиками мультиколлинеарности, но их использование связано со значительными вычислительными трудностями. Наиболее подходящим для оценки информативности данных в вычислительном отношении явля-

ется показатель диагонального преобладания. В силу ограничений (3) он не всегда дает гарантированные оценки. Однако в данном случае это несущественно, поскольку задача заключается в отборе наиболее информативных данных, для которых область его существования достигается. Заметим также, что точность вычисления показателя диагонального преобладания (в отличие от других мер) не зависит от степени мультиколлинеарности векторов образов.

Благодарности

Работа выполнена при поддержке российско-американской программы «Фундаментальные исследования и высшее образование» и РФФИ (грант № 06-08-01024).

Литература

1. **Ватанабе, С.** Разложение Карунена-Лоэва и факторный анализ. Теория и приложения, в сборнике переводов «Автоматизированный анализ сложных изображений» / С. Ватанабе: М.:Мир, 1969.- 254-275 с.
2. **Брайан, Д.** The generalized discriminant function: mathematical foundation and computational routine / Д. Брайан // Harvard Educ. Rev., 21, 1951.- 90-95 с.
3. **Theodoridis, S** Pattern Recognition / S. Theodoridis, K. Konstantinos // Academic Press, 2006.
4. **Ту, Дж.** Принципы распознавания образов / Дж. Ту, Р. Гонсалес: М.:Мир, 1978- 416 с.
5. **Дуда, Р.** Распознавание образов и анализ сцен / Р. Дуда, П. Харт: М.:Мир, 1976.- 512 с.
6. **Фурсов, В.А.** Идентификация моделей систем формирования изображений по малому числу наблюдений / В.А. Фурсов: Самарский государственный аэрокосмический университет, 1998.- 220 с.
7. FERET Face Database // <http://www.itl.nist.gov/iad/humanid/colorferet/>
8. **Phillips, P.** Overview of face recognition grand challenge / P. Phillips [et. al.] // Proceedings of CVPR 2005.- 947-54 pp.

FACE RECOGNITION ON THE BASIS OF CONJUGATION INDEXES IN THE SPACE OF SUMMARIZING INVARIANTS

N.E. Kozin¹, V.A. Fursov²

¹*S.P. Korolyov Samara State Aerospace University, Samara, Russia,*

²*Image Processing Systems Institute of the RAS, Samara, Russia*

Abstract

New method for reduction of the feature space in pattern recognition is discussed. The main idea is elimination of the most spurious feature vector components. As the criterion for such informative distinction new feature called diagonal prevalence index is proposed. The efficiency in a sense of both discrimination ability and computational complexity is being discussed in comparison with other multicollinearity features. Finally, we list algorithm for the image recognition based on the diagonal prevalence idea implementation.

Key words: multicollinearity, diagonal prevalence, feature space reduction.

Citation: Kozin NE, Fursov VA. Face recognition on the basis of conjugation indexes in the space of summarizing invariants. *Computer Optics* 2008; 32(3): 307-11.

Acknowledgements: The work was supported by the Russian-American program "Fundamental researches and higher education" and the RFBR (grant № 06-08-01024).

References

- [1] Watanabe S. Karhunen-Loeve expansion and factor analysis. Theory and appendices [In Russian]. Collection of translations "Automated analysis of complex images." Moscow: "Mir" Publisher 1969; 254-275.
- [2] Brayon JG. The generalized discriminant function: mathematical foundation and computational routine. *Harvard Educ Rev* 1951; 21: 90-95
- [3] Theodoridis S, Konstantinos K. *Pattern Recognition*. Academic Press, 2006.
- [4] Tu J., Gonsales R. Principles of pattern recognition [In Russian]. Moscow: "Mir" Publisher 1978; 416.
- [5] Duda R, Hart P. Pattern classification and scene analysis [In Russian]. Moscow: "Mir" Publisher 1976; 512 p.
- [6] Fursov VA. Identification of models of imaging systems for the small number of observations [In Russian]. Samara, SSAU 1998; 220.
- [7] FERET Face Database. URL: <http://www.itl.nist.gov/iad/humanid/colorferet/>
- [8] Phillips P. Overview of face recognition grand challenge. *Proceedings of CVPR 2005*; 947: 54