

УДК 004.932.2:004.93'1

Н.В. Власенко

Харьковский национальный университет радиоэлектроники, Харьков

ИССЛЕДОВАНИЕ СВОЙСТВ ТРАНСФОРМИРОВАННЫХ ИНВАРИАНТНЫХ ПРИЗНАКОВЫХ ОПИСАНИЙ НА БАЗЕ ОРТОГОНАЛЬНЫХ ПРЕОБРАЗОВАНИЙ

Работа посвящена исследованию свойства инвариантности при применении дискретных ортогональных преобразований для структурных описаний изображений в задачах распознавания. Приводятся результаты экспериментальных исследований при сопоставлении описаний в исходном и трансформированном пространствах признаков с точностью до порогового значения.

Ключевые слова: характерные признаки, структурное описание, распознавание, SURF, дескрипторы, преобразование Уолша-Адамара, инвариантность.

Введение

Характерные признаки (ХП – это ключевые точки (дескрипторы), которые отражают свойства отдельных участков) изображения. Множество ключевых точек (векторов) образует структурное описание изображения [1 – 7].

Основным современным инструментом формирования структурных описаний является метод SURF, который решает две задачи – поиск особых точек изображения и создание их дескрипторов-векторов, инвариантных к смещениям, масштабу и вращению [6].

Применение трансформаций признаков описаний [2 – 4] вызвано необходимостью снижения размерности признакового пространства для увеличения скорости распознавания или сокращения объемов информации из-за избыточности описаний или наличия пересечения описаний объектов из разных классов. При этом сохранение свойства инвариантности трансформированного описания к геометрическим преобразованиям анализируемого видео-объекта является основополагающим при распознавании.

Целью работы является исследование свойств инвариантности описаний, полученных с использованием детектора SURF и подверженных трансформациям на основе использования ортогонального преобразования Уолша-Адамара [2].

Задачи исследования включают теоретический анализ моделей преобразования, изучение их свойств применительно к описаниям изображений с использованием необходимого параметрического обеспечения и практическое применение трансформированных описаний в задаче классификации в базах видео-информации.

Пространство описаний R_1^n , сформированное методом SURF

Метод SURF формирует описание $Z \subset R_1^n$ из множества R_1^n векторов $z = (z_1, \dots, z_n)$, $z \in Z$

векторного пространства R^n , $R_1^n \subset R^n$. Множество R_1^n определено в виде $R_1^n = \{z \mid z_i \in R^1, \|z\|=1, z_i \in [-1, 1]\}$, т.е. как множество всех n -мерных векторов с вещественными компонентами, причем значения компонентов ограничены отрезком $[-1, 1]$, а евклидова норма векторов из R_1^n равна 1: $\|z\| = \sqrt{\sum_{i=1}^n z_i^2} = 1$. Можно считать, что векторы в n -мерном координатном пространстве R_1^n исходят из начала координат, в результате R_1^n образует поверхность многомерного шара радиусом $R = 1$ (рис. 1).

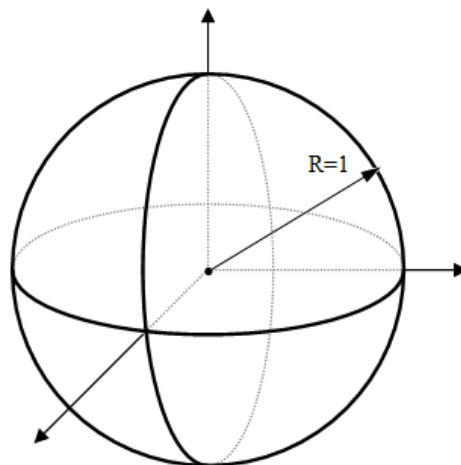


Рис. 1. Схематическое изображение пространства описаний R_1^n

Модели применения преобразования Уолша-Адамара

Дискретное преобразование Уолша-Адамара (ПУА), для которого матрица преобразования состоит из дискретных функций Уолша (ФУ), упорядоченных по Адамару, т.е. по частоте [2], можно считать линейным преобразованием L вектора $z \in R_1^n$ к вектору $u = (u_1, \dots, u_n)$ из нового

множества U_1^n векторов $U_1^n \subset R^n$, т.е. $L: R_1^n \rightarrow U_1^n$. Запишем это преобразование в развернутом виде:

$$U = \begin{bmatrix} u_1^1 \dots u_n^1 \\ u_1^2 \dots u_n^2 \\ \dots \\ u_1^s \dots u_n^s \end{bmatrix} = ZA = \begin{bmatrix} z_1^1 \dots z_n^1 \\ z_1^2 \dots z_n^2 \\ \dots \\ z_1^s \dots z_n^s \end{bmatrix} \times \frac{1}{n} \begin{bmatrix} a_{11} \dots a_{1n} \\ a_{21} \dots a_{2n} \\ \dots \\ a_{n1} \dots a_{nn} \end{bmatrix}. \quad (1)$$

Выражение (1) реализуется путем умножения (строка на столбец) матрицы исходного описания Z , включающей s векторов ХП в виде строк, на ортогональную матрицу Адамара A размером $n \times n$ с нормировочным коэффициентом $1/n$. Выражение $U = ZA$ из-за симметричности A можно считать обобщенной записью традиционного для алгебры линейного преобразования вида $u = Az$, где $u \in R^n, z \in R^n$ – n -мерные вектора-столбцы [1], т.к. любому изменению базиса, связанному с переходом к новому базису, соответствует невырожденное линейное преобразование координат каждого вектора.

Матрица A есть матрицей линейного преобразования в координатном базисе векторов из R_1^n . Каждое невырожденное линейное преобразование, как известно, имеет единственное обратное линейное преобразование, матрица которого есть обратной к матрице A [1]. Для преобразования (1) исходное описание может быть получено из U в соответствии с выражением $Z = UA^{-1}$ путем умножения на обратную матрицу [2]. Учитывая, что для ПУА обратная матрица из-за свойства ортогональности совпадает с A за исключением коэффициента $1/n$, то описание Z может быть вычислено по формуле обратного преобразования $Z = UA$.

Для конкретного вектора $z \in R_1^n$ взаимосвязь с его спектром $u \in U_1^n$ в соответствии с (1) определяется парой дискретных ПУА

$$\begin{cases} u_i = \frac{1}{n} \sum_{k=1}^n z_k a_{ki}, \\ z_i = \sum_{k=1}^n u_k a_{ki}, \end{cases} \quad (2)$$

где a_{ki} – дискретная ФУ (столбец матрицы A) с порядковым номером i . Традиционно в теории обработки сигналов отсчеты векторов z, u изменяются в пределах от 0 до $n-1$ [2–4].

Ортогональная матрица – это квадратная матрица с вещественными элементами, для которой ее обратная матрица равна транспонированной, т.е. $A^{-1} = A^T$. Для матрицы Адамара выполняется соотношение $AA^T = nE$, где E – единичная матрица размера n . Столбцы и строки ортогональной мат-

рицы образуют системы ортонормированных векторов, для которых скалярное произведение строки (столбца) на себя равно 1, а на любую другую – 0. Ортогональные матрицы соответствуют линейным операторам, переводящим ортонормированный базис линейного пространства в ортонормированный.

Строгое в математическом плане определение ортонормированного ПУА записывается в виде соотношения [4]:

$$\begin{cases} u_i = \frac{1}{\sqrt{2^r}} \sum_{k=1}^n z_k a_{ki}, \\ z_i = \frac{1}{\sqrt{2^r}} \sum_{k=1}^n u_k a_{ki}, \end{cases} \quad (3)$$

которое отличается от (2) нормировочными коэффициентами. Значение $r = \log_2 n$ наиболее просто определяется в случае, если n есть степень двойки. Для этого случая матрица Адамара может быть построена по общей формуле $A_{2n} = \frac{1}{\sqrt{2}} \begin{bmatrix} A_n & A_n \\ A_n & -A_n \end{bmatrix}$,

начиная с $A_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ [4]. Конкретно, например, при $n=64$ значение $\sqrt{2^r}$ равно 8, а при $n=128$ величина $\sqrt{2^r} = 8\sqrt{2}$. ПУА в виде (3) сохраняет скалярное произведение между векторами, т.е. выполнено условие $(Ax, Ay) = (x, y)$, где (x, y) – скалярное произведение, $x \in R_1^n, y \in R_1^n$,

$Ax \in U_1^n, Ay \in U_1^n$. Другими словами, норма и угол между векторами сохраняются, т.е. $\|Ax\|^2 = \|x\|^2$ и $\cos(\widehat{Ax}, \widehat{Ay}) = \cos(\widehat{x}, \widehat{y})$. Это позволяет сделать вывод, что отображение $L: R_1^n \rightarrow U_1^n$ в ситуации (3) переводит пространство R_1^n в себя, т.е. $U_1^n = R_1^n$.

При этом U_1^n так же, как и R_1^n , есть множество векторов на многомерном единичном шаре, т.е. с нормой 1 (рис. 1). В практическом плане это дает возможность использовать одни и те же значения параметра порога при установлении эквивалентности в пространствах исходных и трансформированных векторов. В то же время недостатком (3) в сравнении с (2) можно считать необходимость выполнения дополнительной операции деления для каждого из элементов.

В интерпретации теории фильтрации сигналов [2–4] представление в базисе Уолша можно рассматривать как применение к исходному вектору ХП конечного банка фильтров, отражающего спектральные свойства вектора описания. Построение сокращенного базиса на основе использования неполного кортежа ФУ, отобранного по некоторому

критерию, позволяет получить компрессированное (компактное) представление, обладающее высокими показателями быстродействия при распознавании.

Важным моментом представляется изучение инвариантных свойств матрицы U , т.к. описание Z , сформированное детектором SURF для видео-объекта, есть множество векторов, обладающих свойством инвариантности к геометрическим преобразованиям анализируемого видео-объекта из некоторой группы преобразований [6].

Пусть G – группа геометрических преобразований, $g \in G$ – ее элементы, B – изображение видео-объекта, gB – преобразованное изображение. Описание $Z(B)$ называют инвариантным [7] относительно G , если выполнено условие

$$\forall g \in G : Z(B) = Z(gB), \quad (4)$$

где равенство понимается в смысле эквивалентности множеств или векторов.

Учитывая уже проанализированные особенности ПУА, необходимо определить наличие свойства (4) для преобразованного описания U , т.е. проверить выполнение условия $\forall g \in G : U(B) = U(gB)$. Именно свойство инвариантности наряду с компактностью описания, способностью адаптации и автоматического анализа является ключевым для современных прикладных систем распознавания, основанных на интеллектуальном анализе данных [3 – 6].

Результаты экспериментов

Подтверждение свойства инвариантности для трансформированных описаний осуществлялось в процессе компьютерного моделирования. Компьютерные эксперименты проведены для видео-базы домашних животных [5], пример изображений с отмеченными координатами вычисленного детектором SURF множества ХП показан на рис. 2.

Заметим, что выполнение свойства инвариантности применительно к сопоставлению элементов двух описаний с точностью до порогового значения δ_z даже для традиционного голосования признаков SURF в R_1^n значительным образом зависит от этого порога.

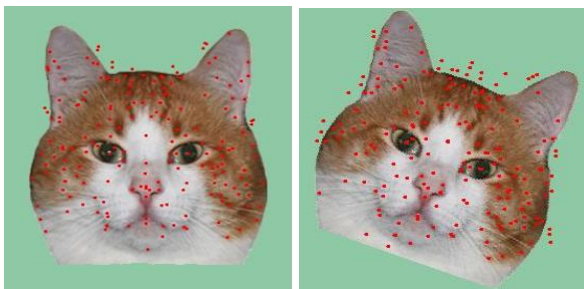


Рис. 2. Эталонное и преобразованное изображения видео-базы

Табл. 1 содержит экспериментальные зависимости числа эквивалентных ХП для 1-го и 2-го эталонов базы (187 и 268 ХП соответственно) при сопоставлении их со своими же эталонами, трансформированными поворотом в 20° (210 и 280 точек). Как видим, лишь при пороге $\delta_z = 0,8$ в процессе сравнения описаний определяются все соответствия точек преобразованного изображения эталону, что характеризует выполнение свойства инвариантности в полном объеме, т.е. для всех точек описания. При меньшем значении порога $\delta_z < 0,8$ лишь для части точек описания соблюдается это свойство.

Таблица 1

Зависимость числа эквивалентных ХП эталона и изображения от порога δ_z

Значение δ_z	1-й эталон	2-й эталон
0,08	13	8
0,16	61	71
0,24	92	108
0,32	105	137
0,40	126	179
0,48	143	216
0,64	176	255
0,80	187	265

В то же время наши экспериментальные исследования показывают, что с увеличением порога δ_z , в частности, при $\delta_z = 0,8$, изображения базы при голосовании уже не различаются между собой в пространстве R_1^n , т.к. практически для всех эталонов число сформированных голосов при сравнении преобразованного объекта с базой равно числу эталонных голосов. Этот характерный факт присущ и другим базам. Из этого можно сделать вывод о существовании для конкретной базы некоторого оптимального значения из рассмотренного диапазона $0,08 < \delta_z < 0,8$, при котором описания базы с точки зрения их элементов различимы, и в то же время число точек, представляющих объект, достаточно значимо для осуществления распознавания. Таким образом, значение порога δ_z можно считать ключевым для обеспечения нужного качества распознавания.

Экспериментальные исследования показали, что оптимальное с этой точки зрения значение порога для базы кошек находится в диапазоне 0,32–0,4. С другой стороны, величину порога δ_z также можно использовать для управления сокращением объема описания, опираясь в целях

сокращения объема вычислений при классификации лишь на его часть.

Эксперименты, проведенные с применением ПУА в виде (3) показывают, что число эквивалентных точек трансформированных описаний при тех же порогах практически совпадает со значениями табл. 1, что можно считать экспериментальным подтверждением свойства инвариантности при применении ПУА. Аналогичная ситуация увеличения сходства между разными описаниями и в результате – ухудшение различимости наблюдается в процессе возрастания величины порога. Так же, как и для традиционного голосования ХП в R_1^n , оптимальное значение порога δ_u при сопоставлении трансформированных описаний в U_1^n можно считать принадлежащим интервалу $\delta_u \in [0,32-0,4]$. Аналогичный вывод можно сделать также относительно преобразования масштаба.

В целом степень влияния преобразований или искажений может быть оценена на основании значений матрицы взаимных расстояний (подобий) между анализируемыми объектами, в случае компьютерной симуляции – это матрица $\{h_{i,j}\}$ числа голосов, отданных объектом за правильный (диагональ $h_{i,i}$) и за все остальные эталоны. Превышение диагонального элемента над остальными в столбце (строке) отражает событие правильного распознавания. Для исследуемой видео-базы из 12 изображений и для обоих признаков пространств R_1^n , U_1^n вычислены матрицы взаимного подобия описаний эталонов с применением голосования ХП и их трансформаций.

Для обоих пространств характерна высокая степень подобия между изображениями с номерами 5 и 10. Значения матрицы $\{h_{i,j}\}$ при $\delta_z = 0,32$ содержатся в табл. 2.

Таблица 2

Матрица взаимных голосов для эталонов базы

	1	2	3	4	5	6	7	8	9	10	11	12
1	103	23	3	0	22	1	1	31	1	27	2	12
2		137	12	2	27	8	7	33	1	35	3	13
3			50	0	2	1	0	7	1	4	2	3
4				185	0	15	4	1	2	0	7	0
5					98	1	0	53	0	59	0	17
6						147	5	2	3	1	12	1
7							113	1	1	0	2	0
8								121	0	51	3	12
9									101	0	1	0
10										124	3	16
11											118	1
12												63

Дальнейшие эксперименты были связаны со сравнительной оценкой помехозащищенности традиционного и модифицированного ПУА методов. Для этого в каждой точке моделировалось воздействие на входное изображение аддитивного шума с матожиданием 0 и дисперсией σ^2 (СКО σ). Эксперименты показали, что действие шума, например, при $\sigma=0,001$, можно считать аналогичным действию геометрических преобразований в плане уменьшения диагональных и увеличения остальных элементов матрицы $\{h_{i,j}\}$.

До определенного уровня шума $\sigma=0,1$ матрица голосов для традиционного и модифицированного методов практически совпадает. Однако при значительных уровнях шума $\sigma=1$ начинают сказываться

интегральные свойства преобразование Уолша-Адамара и диагональные элементы матрицы $\{h_{i,j}\}$ для модифицированного подхода имеют более устойчивый характер.

Табл. 3 содержит зависимость числа найденных эталонных точек на преобразованном (поворот 20°) изображении в зависимости от уровня аддитивной помехи.

Таблица 3

Число эквивалентных точек в зависимости от уровня помехи

σ	0	0,01	0,1	1	10	50	100
Число точек	187	164	164	163	121	34	11

Как видим, даже при очень высоком уровне шума $\sigma=100$, что приблизительно соответствует соотношению сигнал-шум 1, исследуемые методы устойчиво определяют 6% точек своего описания, что часто бывает достаточным для решения задачи распознавания объектов.

Выводы

В данном исследовании впервые доказано, что эффективное представление структурных признаков изображения с использованием преобразование Уолша-Адамара позволяет сформировать компактное информативное описание, сохраняющее свойство инвариантности к геометрическим преобразованиям изображений.

Ключевым фактором обеспечения инвариантности является пороговое значение δ_z , которое должно оптимальным образом одновременно обеспечивать как сходство преобразованного объекта с эталоном, так и различие между описаниями в процессе распознавания. При этом пороговые значения, используемые при классификации в пространстве исходного векторного описания и в пространстве, полученном путем применения ПУА, являются идентичными.

Результаты компьютерного моделирования для конкретной видео-базы показали, что существует оптимальное значение порога из диапазона $0,08 < \delta_z < 0,8$, при котором определен такой набор точек описания, который обеспечивает достаточную достоверность классификации.

Практическая ценность работы состоит в экспериментальном подтверждении эффективности представления характерных признаков путем при-

менения ортогонального преобразования на прикладной базе видео-данных домашних животных.

Перспектива дальнейших исследований состоит в разработке критериев для оптимального вычисления порога δ_z эквивалентности элементов описания.

Список литературы

1. Ефимов Н.В. *Линейная алгебра и многомерная геометрия* / Н.В. Ефимов, Э.Р. Розендорн. – М.: ФИЗМАТЛИТ, 2004. – 464 с.
2. Ахмед Н. *Ортогональные преобразования при обработке цифровых сигналов: пер. с англ.* / Н. Ахмед, К.Р. Рао. – М.: Связь, 1980. – 248 с.
3. Ересько Ю.Н. *Локализация изображений в автоматических визирах* / Ю.Н. Ересько. – М.: Компания Спутник+, 2002. – 357 с.
4. Pratt W.K. *Digital Image Processing* / W.K. Pratt. – New York: John Wiley and Sons Inc., 2001. – 723 p.
5. Власенко Н.В. *Анализ корреляций наиболее значимых соответствий между описаниями объектов в задачах метрического распознавания* / Н.В. Власенко, В.П. Маиталир // *Реєстрація, зберігання і обробка даних.* – 2012. – Т. 14, № 3. – С. 20-32.
6. Bay H. *Surf: Speeded up robust features* / H. Bay, T. Tuytelaars, L. Van Gool // *European Conference on Computer Vision.* – 2006. – P. 404-417.
7. Путьтин Е.П. *Обработка изображений в робототехнике* / Е.П. Путьтин, С.И. Аверин. – М.: Машиностроение, 1990. – 320 с.

Поступила в редколлегию 28.11.2012

Рецензент: д-р техн. наук, проф. Е.П. Путьтин, Харьковский национальный университет радиоэлектроники, Харьков.

ДОСЛІДЖЕННЯ ВЛАСТИВОСТЕЙ ТРАНСФОРМОВАНИХ ІНВАРІАНТНИХ ОЗНАКОВИХ ОПИСІВ НА БАЗІ ОРТОГОНАЛЬНИХ ПЕРЕТВОРЕНЬ

Н.В. Власенко

Робота присвячена дослідженню властивості інваріантності при використанні дискретних ортогональних перетворень для структурних описів зображень в задачах розпізнавання. Наведені результати експериментальних досліджень при зіставленні описів у вихідному та трансформованому просторах ознак з точністю до порогового значення.

Ключові слова: характерні ознаки, структурний опис, розпізнавання, SURF, дескриптори, перетворення Уолша-Адамара, інваріантність.

INVESTIGATION OF TRANSFORMED INVARIANT ATTRIBUTIVE DESCRIPTIONS BASED ORTHOGONAL TRANSFORMATION

N.V. Vlasenko

Work is devoted to the application of the invariance of discrete orthogonal transforms for structural descriptions of images in pattern recognition. The results of experimental studies in comparing of descriptions in the original and transformed spaces signs up to a threshold value.

Keywords: characteristic features, structural description, identification, SURF, descriptors, the Walsh-Hadamard invariance.