

Исследование конфигурационного отношения эквивалентности в задаче описания и анализа изображения

Е.А. Дмитриев¹, В.В. Мясников^{1,2}

¹Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

²Институт систем обработки изображений РАН - филиал ФНИЦ «Кристаллография и фотоника» РАН, Молодогвардейская 151, Самара, Россия, 443001

Аннотация. В работе исследуется метод описания изображений, основанный на конфигурационном отношении эквивалентности. В соответствие с этим методом изображение ассоциируется с оптимальной конфигурацией (перестановкой) его составляющих отсчетов, областей или векторов признаков. В работе исследуется способ описания изображения, основанный на конфигурациях областей изображения. Эффективность указанного представления исследуется на примере решения задачи поиска геометрически-искажённых дубликатов.

1. Введение

Одна из ключевых проблем в решении задачи компьютерного зрения – нахождение оптимального описания изображений. С одной стороны, для решения данной задачи требуется, чтобы у найденного представления была инвариантность к одним преобразованиям и в то же время чувствительность к другим. В большинстве случаев для новой прикладной области приходится придумывать новые методы для нахождения оптимального описания изображения. Тем не менее существуют подходы, которые используются для достаточно широкого круга задач [1].

В настоящее время один из наиболее популярных подходов к решению задачи оптимального описания изображений является использование свёрточных нейронных сетей [2]. Тем не менее существуют сильные ограничения по использованию на практике такого подхода: требование “представительной” (достаточного объёма, с отсутствием выбросов и смещённости в данных) выборки, затраты времени на обучение моделей.

Есть и другой подход, который основан на идее, что изображение можно характеризовать набором отношений между его составляющими, например, отсчетами или областями. Данная идея применялась ещё в рамках идеологии семантических сетей [3]. Частным случаем подобного описания в компьютерном зрении являются локальные бинарные шаблоны или LBP [4], знаковые представления изображений [5]. Обобщением представленных методов является подход, описанный в работе [6], где рассматриваются способы конфигурации отношения эквивалентности для решения задачи поиска яркостно-искажённых дубликатов. В настоящей работе предложенный подход [6] используется для решения задачи поиска геометрически-искажённых дубликатов.

Работа построена следующим образом. Во 2 разделе представлена краткая формализация предлагаемого подхода к описанию данных с использованием отношений. В 3 разделе указанная формализация применяется для построения описаний изображений. Этот же раздел содержит описание постановок и результаты экспериментальных исследований, проведенных для демонстрации работоспособности и эффективности предлагаемого подхода.

2. Основные понятия

Рассмотрим множество из N векторов длины n с компонентами, из поля \mathbf{R} :

$$\bar{g}_0, \bar{g}_1, \dots, \bar{g}_{N-1}, \bar{g}_j = \begin{pmatrix} g_{0j} \\ g_{1j} \\ \dots \\ g_{(n-1)j} \end{pmatrix}, j = \overline{0, N-1},$$

формирующее матрицу $G \in \mathbf{R}^{n \times N}$ следующего вида:

$$G \equiv (\bar{g}_0, \bar{g}_1, \dots, \bar{g}_{N-1}) = \begin{pmatrix} g_{00} & g_{01} & \dots & g_{0(N-1)} \\ g_{10} & g_{11} & & g_{1(N-1)} \\ \vdots & & & \vdots \\ g_{(n-1)0} & & \dots & g_{(n-1)(N-1)} \end{pmatrix}.$$

И пусть из каких-либо внешних соображений, формализуемых критерием качества F , для данной конкретной матрицы G определена "наилучшая" перестановка $\sigma_{F,G} \in \Sigma_N$ ее столбцов (и, как следствие, векторов исходного множества), зависящая от компонентов матрицы:

$$\sigma_{F,G} : \mathbf{Z}_{N-1} \xrightarrow{i \rightarrow \sigma_{F,G}(i)} \mathbf{Z}_{N-1}.$$

Здесь \mathbf{Z}_N - множество неотрицательных целых чисел $0, 1 \dots N-1$, а Σ_N - множество всех перестановок для N векторов (объектов). Учитывая, что число возможных перестановок для N объектов составляет $N!$, имеем:

$$|\Sigma_N| = N!.$$

С конкретной перестановкой σ (подиндекс порождающих перестановку критерия F и матрицы G в некоторых случаях будем опускать) свяжем код перестановки - целое неотрицательное число, обозначаемое далее $\chi(\sigma)$ и определяющее ее позицию среди всех возможных $N!$ перестановок, упорядоченных по какому-либо признаку. Учитывая общее количество перестановок, можно записать:

$$\chi : \Sigma_N \xrightarrow{\sigma \mapsto \chi(\sigma)} \mathbf{Z}_{N!}$$

Для формализации понятия "наилучшей" перестановки введем далее формальный критерий качества перестановки в виде функционала:

$$F : \mathbf{R}^{n \times N} \times \Sigma_N \rightarrow \mathbf{R}.$$

который для каждой $n \times N$ матрицы G и перестановки σ из Σ_N , указывает вещественную величину - показатель качества матрицы, вычисленный с учетом перестановки σ столбцов матрицы G . Множество критериев для конкретных параметров n и N матриц обозначим $\mathbf{F}_{n,N}$, а в все возможное множество критериев - \mathbf{F} .

Будем считать далее, что меньшие значения критерия соответствует более высокому качеству. Обозначим оптимальное значение критерия F_{\min}^G , определяя его в виде:

$$F_{\min}^G = \min_{\sigma \in \Sigma_N} F(G, \sigma).$$

Учитывая, что это оптимальное значение может достигаться на практике на некотором подмножестве перестановок, введем дополнительное обозначение для этого непустого подмножества:

$$\Sigma_N^{F,G} = \left\{ \sigma : F(G, \sigma) = F_{\min}^G, \sigma \in \Sigma_N \right\}$$

И, наконец, определим *оптимальную перестановку* $\sigma_{F,G}$ для матрицы G .

Определение 1. *Оптимальной перестановкой* $\sigma_{F,G}$ для матрицы G называется перестановка из $\Sigma_N^{F,G}$ с наименьшим кодом:

$$\sigma_{F,G} = \arg \min_{\sigma \in \Sigma_N^{F,G}} \chi(\sigma)$$

Хэш-код оптимальной перестановки $\kappa(\chi(\sigma_{F,G}))$ полученный однозначным преобразованием $\kappa : \mathbf{Z}_{N!} \rightarrow \mathbf{Z}_K$ ($K \leq N!$) кода оптимальной перестановки, рассматриваются далее как *первичная характеристика-описатель* анализируемых данных (*дескриптор* данных или его составная часть). В простейшем случае, когда κ - тождественное преобразование, хэш-код и код оптимальной перестановки совпадают.

3. Применение метода к описанию изображений

Для применения рассматриваемого подхода к описанию изображений требуется в первую очередь определить способ получения векторов матрицы G . Формально процесс построения дескриптора изображения (или фрагмента изображения) может быть представлен в виде набора этапов.

- Этап 1. Выбор структуры и порядка обхода изображения (порядка формирования векторов).
- Этап 2. Выбор способа задания или расчета компонент векторов.
- Этап 3. Выбор критерия качества перестановки.
- Этап 4. Выбор способа "нормализации" оптимальной перестановки.
- Этап 5. Задание функции расчета хэш-кода оптимальной перестановки.

В случае если для описания изображения используется не один дескриптор, а некоторое множество, то указанные шаги выполняются соответствующее число раз. Агрегирование получаемых хэш-кодов для получения окончательного дескриптора изображения производится исходя из конкретной задачи.

3.1 Применение метода к описанию изображений

Задача поиска дубликатов на изображении [7, 8, 9] (англ.: copy-move forgery detection) состоит в нахождении на изображении таких фрагментов, которые с точностью до допустимого набора преобразований являются неотличимыми. Под допустимыми преобразованиями обычно понимают *геометрические преобразования* типа поворота и масштабирования, а также *яркостные*, включающие монотонно-возрастающие преобразования функции яркости и незначительные шумовые искажения. В данном подразделе предлагаемый метод описания изображения применяется для поиска геометрически-искажённых дубликатов.

Ниже представлено уточнение по всем указанным выше этапам предлагаемого метода. Для определенности полагаем, что анализируемый фрагмент $f(m_1, m_2)$ изображения имеет размер 19x19. Начало координат фрагмента для определенности поместим в его центр: $m_1, m_2 = \overline{-9.9}$.

Этап 1. Структура и порядок обхода ($N=8$) фрагмента изображения представлены на рисунке 1.

Этап 2. Способ задания или расчета компонент $n=6$ -мерного вектора для конкретной позиции продемонстрирован на рисунке 1.

Этап 3. Критерий качества перестановки выберем в виде:

$$F(G, \sigma) = \sum_{j=0}^{N-2} \left\| \bar{g}_{\sigma(j+1)} - \bar{g}_{\sigma(j)} \right\|_2.$$

Этап 4. "Нормализацию" оптимальной перестановки с учётом циклического сдвига определим следующим образом:

$$\sigma_{F,G}^* = \arg \min_{\sigma \in \Sigma_N^{F,G}} \sigma_{F,G}((i+k^*) \bmod N),$$

где

$$k^* = \arg \min_{k=0, N-1} \sigma_{F,G}((i+k^*) \bmod N).$$

Этап 5. Формирование числа-дескриптора (16-битного) происходит следующим образом:

$$\chi(\sigma_{F,G}) \in \mathbf{N}_{5!} (8!=40320).$$

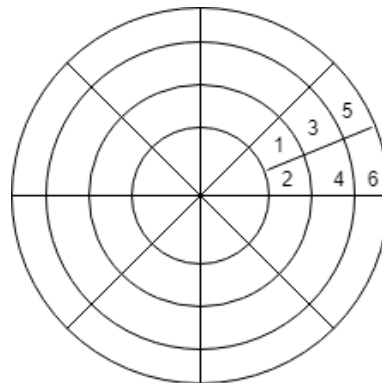


Рисунок 1. Иллюстрация к порядку и способу формирования векторов описания анализируемого фрагмента.

Число-дескриптор, характеризующее фрагмент изображения в позиции (m1,m2) помещается в отсчет выходного изображения-образа дескриптора. Результаты обнаружения дубликатов (повернутых на 0, 45 и 90 градусов) с использованием указанных дескрипторов приведены, соответственно, на рис.2.

4. Заключение

В данной статье представлен метод получения описания изображения для решения задачи поиска геометрически-искажённых дубликатов с помощью конфигурационного отношения эквивалентности.

Проведены эксперименты, в ходе которых выяснено, что разработанный метод демонстрирует хорошую эффективность в рассматриваемой задаче.

5. Литература

- [1] Гашников, М.В. Методы компьютерной обработки изображений / М.В. Гашников, Н.И. Глумов, Н.Ю. Ильясова, В.В. Мясников, С.Б. Попов, В.В. Сергеев, В.А. Сойфер, А.Г. Храмов, А.В. Чернов, В.М. Чернов, М.А. Чичева, В.А. Фурсов; под ред. В.А. Сойфера. – М.: Физматлит, 2003. – 784 с.
- [2] Schroff, F. Facenet: A unified embedding for face recognition and clustering / D. Kalenichenko, J. Philbin, F. Schroff // Proc. CVPR, 2015.
- [3] Ballard, D.H. Computer vision / D.H. Ballard, C.M. Brown // Prentice-Hall Inc., Englewood Cliffs. – New Jersey, 1982. – 547 p
- [4] Ojala, T. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions / T. Ojala, M. Pietikäinen, D. Harwood // Proceedings of the 12th IAPR International Conference on Pattern Recognition (ICPR). – 1994. – Vol. 1. – P. 582-585.

- [5] Goncharov, A.V. Investigation of the properties of images sign representation in the pattern recognition problems / A.V. Goncharov // Izvestiya S. Fed. U. Engineering Sciences. – 2009 – p. 178-188. (in Russian).
- [6] Мясников, В.В. Описание изображений с использованием конфигурационного отношения эквивалентности / В.В. Мясников // Компьютерная оптика. – 2018. – Т. 42, № 6. – С. 998-1007. DOI: 10.18287/2412-6179-2018-42-6-998-1007.
- [7] Kuznetsov, A. A copy-move detection algorithm using binary gradient contours / A Kuznetsov, V. Myasnikov // Proc. International Conference on Image Analysis and Recognition. – 2016. – P. 349-357.
- [8] Мясников, В.В. Локальное порядковое преобразование цифровых изображений / В.В. Мясников // Компьютерная оптика. – 2015. – Т. 39, № 3. – С. 397-405. DOI: 10.18287/0134-2452-2015-39-3-397-405.
- [9] Kuznetsov, A.V. A fast plain copy-move detection algorithm based on structural pattern and 2D Rabin-Karp rolling hash / A.V. Kuznetsov, V.V. Myasnikov // Proceedings of the 11th International Conference, ICIAR. – Lecture Notes in Computer Science. – 2014. – Vol. 8814(1). – P. 461-468.

Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 18-01-00748-а, № 17-29-03190-офи-м.

Equivalence relation configuration study in the image description and analysis problem

Е.А. Dmitriev¹, V.V. Myasnikov^{1,2}

¹Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

²Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

Abstract. This paper presents image description method based on equivalence relation configuration. According to this approach analyzed image is matched to optimal configuration of its attributes like pixel values, regions or feature vectors. Image description using regions configuration is considered in this paper. Results of experiments on copy-move forgery detection task showed effectiveness of analyzed approach.