

ФОРМИРОВАНИЕ ОТНОШЕНИЙ АФФИННЫХ БАЗИСОВ ПРИ РАСПОЗНАВАНИИ ПО СТРУКТУРНЫМ ПРИЗНАКАМ

Гороховатский В.А., Долбня Н.Е., Русаков А.Ю.

Харьковский национальный университет радиоэлектроники

E-mail: gorohovatsky-v@ Rambler.ru; тел. (057) 702-14-19

This work is dedicated to the solving of such problem as comparing visual objects based on different characteristics. Affine-transformed images with defined features had been taken as investigated objects. The modification of geometric hashing method was suggested. The number of experiments were held for definition of the speed of working and the results accuracy. The optimal limit of comparing with the hash-table was reached in the empirical way.

Введение. Распознавание изображений на базе структурных признаков (СП) по сравнению с интегральными методами обладает таким важным преимуществом, как возможность работы со структурными представлениями объектов [1,2]. Наряду с существенным сокращением объема анализируемой информации, применение аппарата СП позволяет осуществлять распознавание с учетом частичных изменений объекта.

СП чаще всего представляют собой конечный набор значений локальных фильтров. Примером могут быть признаки Харриса, локальные потоки [2], дескрипторы метода SIFT [3]. Анализ СП происходит в условиях действия аффинных искажений, поэтому при принятии решения должна обеспечиваться инвариантность. Примеры изображений с выделенными СП в условиях преобразований показаны на рис.1.

Сущность. Предположим, что имеются база данных моделей и анализируемый объект в виде множеств СП. Известно, что этот объект может быть представлен посредством аффинного преобразования одной из моделей. Требуется принять решение о соответствии объекта некоторой модели и найти параметры преобразования, совмещающего модель и изображение объекта.

Будем полагать, что модель M_i задана в форме множества признаков точек $m_{ik} \in M_i, k = 1, \dots, s(i)$. Любое подмножество трех неколлинеарных (в плане пространственных координат) точек $w = \{e_{00}, e_{01}, e_{10}\}$ модели $M_i, w \subset M_i$, определяет аффинный базис, который задает на M_i некоторую систему координат [1].



Рисунок 1 – Изображения с выделенными СП

Суть метода аффинных инвариантов [1] состоит в представлении точек модели и объекта $x \in X$ в некотором аффинном базисе посредством параметров (ξ, η) в виде

$$x = \xi(e_{10} - e_{00}) + \eta(e_{01} - e_{00}) + e_{00}, \quad (1)$$

где (e_{00}, e_{01}, e_{10}) - базис. Применяя аффинное преобразование T к точке x , получим

$$Tx = \xi(Te_{10} - Te_{00}) + \eta(Te_{01} - Te_{00}) + Te_{00}. \quad (2)$$

Из соотношения (2) следует, что Tx имеет те же аффинные координаты (ξ, η) относительно базиса $(Te_{00}, Te_{01}, Te_{10})$, что и точка x относительно базиса (e_{00}, e_{01}, e_{10}) . На основе этих рассуждений применим координаты (ξ, η) в качестве признаков для распознавания по множеству СП.

Для реализации метода на этапе обучения строим хеш-таблицу для эталонных моделей (табл. 1).

Таблица 1 – Хеш-таблица эталонных моделей изображений

ξ	η	Модель M_i	Базис		
			e_{00}	e_{01}	e_{10}

Распознавание будем осуществлять путем подсчета и последующего поиска максимума для количества голосов, отданных каждым СП $x \in X$ объекта при фиксированном базисе $w \in W$ за соответствующую модель M_i , W - множество базисов. Формирование гистограммы голосов традиционно происходит путем полного перебора и проверки возможных соответствий множества представлений объекта со всеми элементами таблицы моделей. Сравнение величин аффинных координат (ξ, η) с табличными значениями реализуется путем оценки отклонения в метрике модуля разности. Голос модели фиксируется, если выполнено условие

$$\rho([\xi, \eta], [\xi_0, \eta_0]) = |\xi - \xi_0| + |\eta - \eta_0| < \varepsilon, \quad (3)$$

где ξ_0, η_0 - табличные значения, ε - допустимое отклонение.

Модификация метода. Предлагаемое усовершенствование метода состоит в следующем. На множестве W базисов объекта формируем и анализируем бинарное отношение $R_W \subset W \times W$, заключающееся в том, что параметры (аффинные координаты (ξ, η)) представления точки x в двух разных базисах относятся к одной и той же модели из хеш-таблицы. За счет введения отношений и анализа множества пар базисов надежность сопоставления в пространстве структурных признаков значительно улучшается по отношению к одиночным базисам [1], так как вероятность случайного совпадения признаков при этом становится существенно ниже [2].

Заметим, что количество пар базисов из R_W существенно больше числа одиночных базисов, т.е. мощности множеств W, R_W соотносятся как $\#W \square \#R_W$ ($\#$ - мощность). Соответственно увеличивается и необходимый объем вычислений.

Чтобы в некоторой степени сократить возросший объем вычислений при сопоставлении СП для пар базисов, целенаправленно использован усеченный набор

$R_W^* \subset R_W$. Процедура усечения $R_W^* = U(R_W)$ состоит в построении конечной цепочки базисов в соответствии со сформировавшейся в процессе обработки нумерацией СП объекта. Например, пусть точка 1 (СП) берется в качестве анализируемой, а точки 2,3,4 выступают в качестве первого элемента пары базисов. Вторым элементом пары можно выбрать точки 3,4,5. Как видим, построенные таким способом базисы пересекаются на множестве точек 3,4. При условии истинности отношения, если базисы голосуют за одну и ту же модель, фиксируется голос пары за эту модель. Относительно рассматриваемой в данный момент точки выбирается следующая пара базисов: точки 3,4,5 и 4,5,6 из множества СП. Процедура продолжается, пока закончится возможность выбора пар базисов из цепочки. Количество используемых пар базисов объекта в данной конкретной ситуации оценивается комбинаторно по формуле

$$C_n = C_{n-1}^3 \cdot n = \frac{n!}{6(n-4)!}, \quad (4)$$

где $n > 4$ – количество точек во множестве СП объекта.

Как видим, в предложенной модификации каждая пара СП участвует в анализе двух базисов. Понятно, что здесь может быть применен и любой другой путь формирования последовательности пар базисов, приводящий к сокращению количества сопоставлений. При введении процедуры усечения U количество полученных голосов должно быть дополнительно нормировано путем деления на максимально возможное число пар базисов для конкретного типа усечения, определяемое соотношением вида (4).

Компьютерное моделирование. Эксперименты проводились для монохромных изображений букв шрифта «Десо» (см. рис.1), множества СП моделей формировались в интерактивном режиме с числом признаков в пределах 8-14. Эмпирическим путем определен порог \mathcal{E} для отклонения признаков при сравнении с хэш-таблицей. Аффинные преобразования и распознавание осуществлялись на множестве имеющихся моделей. Эксперименты показали, что вероятность распознавания по множеству СП в условиях аддитивных помех (незначительное отклонение координат СП) не хуже, чем для традиционного метода. В то же время быстроедействие модифицированного метода в некоторых случаях превосходит классический подход примерно в 3-4 раза. Выигрыш растет с увеличением размерности анализируемого множества СП, поэтому для букв с большим числом СП, например, букв «Ж», «Ш» этот выигрыш выше. Устойчивость метода к действию локальных искажений в пределах 30% (пропадание истинных и появление ложных точек) примерно такая же, как и для классического подхода.

Выводы. Разработанная модификация позволяет без существенного снижения надежности значительно повысить быстроедействие распознавания при применении метода аффинных инвариантов в структурных методах распознавания изображений. Путем введения согласованной процедуры формирования и анализа ограниченного множества аффинных базисов можно добиться оптимального для конкретных условий соотношения между надежностью и быстроедействием распознавания в фиксированной базе моделей. Процедурный путь построения базисов позволяет не только расширять или сужать объем используемой информации для обеспечения требуемого уровня надежности, но и учитывать структурные особенности распознаваемых объектов.

Литература

1. Шатино Л., Стокман Дж. Компьютерное зрение. Пер. с англ. – М.: Бином, 2006.–752 с.
2. Путьатін Є.П., Гороховатський В.О., Матам О.О. Методи та алгоритми комп'ютерного зору: Навч. посібник. – Харків: ТОВ «Компанія СМІТ», 2006. –236с.
3. Lowe D.G. Distinctive Image Features from Scale-Invariant Keypoints.- International Journal of Computer Vision, 60, 2, 2004, p. 91-110.
4. Wallraven C., Caputo B., Graf A. Recognition with local features: the kernel recipe. – In IEEE International Conference on Computer Vision (ICCV'2003). P. 257-264.