

Нормализация изображений

Л.В. Рябова, ассистент кафедры средств защиты информации института информационно-диагностических систем Национального авиационного университета, lubanau@ukr.net

Обработка изображений с целью их распознавания является одной из центральных и практически важных задач при создании систем искусственного интеллекта. Проблема носит явно выраженный комплексный иерархический характер и включает ряд основных этапов: восприятие поля зрения, сегментация, нормализация выделенных объектов, распознавание. Такой важный обязательный этап как понимание (интерпретация) изображений включается частично в этап сегментации и окончательно решается на этапе распознавания. Основным элементом любой задачи распознавания изображений является ответ на вопрос: относятся ли данные (входные) изображения к классу изображений, который представляет данный эталон? Казалось бы, ответ можно получить, сравнивая непосредственно изображение с эталонами (или их признаки). Однако возникает ряд трудностей и проблем, специфических, в особенности, при создании систем технического зрения (СТЗ):

1. Изображения предъявляются на сложном фоне.
2. Изображения эталона и входные изображения отличаются положением в поле зрения.
3. Входные изображения не совпадают с эталонами за счет случайных помех.
4. Отличия входных и эталонных изображений возникает за счет изменения освещенности, подсветки, локальных помех.
5. Эталоны и изображения могут отличать геометрические преобразования, включая такие сложные как аффинные и проективные.

Для решения задачи в целом и на отдельных ее этапах применяются различные методы сегментации, нормализации и распознавания. Операция предобработки применяется практически всегда после снятия информации с видеодатчика и преследует цель снижения помех на изображении, возникших в результате дискретизации и квантования, а также подавления внешних шумов. Как правило, это операции усреднения и выравнивания гистограмм.

Методы нормализации при распознавании занимают промежуточное место между сигнальными корреляционными и признаковыми алгоритмами. В отличие от признаковых при нормализации изображение не “теряется”, а только замещается изображением того же класса эквивалентности. В то же время, в отличие от корреляционных методов, множество входных изображений заменяется множеством нормализованных изображений.

Каждое нормализованное изображение, в общем случае, находится гораздо ближе к своему эталону (с позиции групповых преобразований), что значительно сокращает количество корреляций на завершающем этапе распознавания.

Суть нормализации заключается в автоматическом вычислении неизвестных параметров преобразований, которым подвергнуты входные изображения, и последующем приведении их к эталонному виду. Процедура преобразований производится с помощью операторов нормализации (нормализаторов), а вычисление параметров выполняется функционалами, действующими на множестве изображений. В монографии [3] подробно излагаются параллельные и последовательные, параметрические и следящие нормализаторы, которые нашли эффективное применение для базовых преобразований: смещений, поворотов, растяжений, косых сдвигов и некоторых их комбинаций. Остается открытым вопрос о поиске универсальных и надежных нормализаторов для сложных групп преобразований – аффинных и проективных.

Параллельная нормализация методом сечений не всегда приводит к однозначному определению параметров аффинных или проективных преобразований, несмотря на кажущуюся универсальность. В самом деле, всякий раз для нового изображения возникает вопрос о рациональном разбиении на градации яркости полутоновых изображений. Это требует постоянного присутствия оператора и интерактивного режима работы. Кроме того изображения типа “силуэт”, а также малоградационные вообще не поддаются нормализации этим методом.

Метод полиномиальных коэффициентов обладает значительной общностью и универсальностью, однако имеет чрезвычайно высокую вычислительную трудоемкость и невысокую точность в связи с накоплением ошибок вычислений.

Наибольший интерес на данном этапе развития теории нормализации представляют последовательные методы, основанные на поэтапном вычислении параметров сложных преобразований и применении частичных нормализаторов на каждом этапе. Последовательные методы предполагают возможность разложения сложных групп на более простые подгруппы. Например, аффинную группу преобразований G_a можно представить в виде суперпозиции центроаффинной G_p и группы параллельных смещений: $G_a = G_p G_c$. Это разложение позволяет ставить вопрос о последовательной нормализации с помощью суперпозиции частичных: $F_a = F_p F_c$, где F_c – нормализатор центрирования (смещений), а F_p – нормализатор центроаффинной группы, которая однозначно определяется матрицей $A = (a_{ij})$, $i, j = 1, 2, \dots$

В свою очередь, чтобы применить частичные нормализаторы к центроаффинным преобразованиям изображений, необходимо уметь разлагать матрицу A на составляющие. При этом не все разложения равноценны с точки зрения практической реализации. Общеизвестно представление квадратурной матрицы в виде комбинации самосопряженной

и ортогональной. В свою очередь самоопряженную матрицу можно выразить произведением ортогональных из собственных векторов, диагональной и обратной ортогональной. Это разложение, в принципе, можно использовать для построения суперпозиции соответствующих частичных нормализаторов аффинной группы [3]. Однако, практическое применение указанного разложения вызывает большие трудности. В частности, возникает неоднозначность при равенстве масштабных коэффициентов матрицы Λ (теряет смысл второй поворот). Кроме того, возникают существенные ограничения на величины углов поворотов, которые можно определить (не более $\pi/4$).

Другие разложения, полученные в последнее время, именно $A=\Lambda XU$, $A=Y\Lambda X$, $A=\Lambda YU$, $A=\Lambda YX$, где X, Y - соответственно матрицы косых сдвигов вдоль оси x и вдоль оси y , U -матрица поворота, оказываются практически более удобными и не имеют ограничений, отмеченных выше. В этой связи значительный интерес представляет подход, основанный на инвариантных прямых при аффинных преобразованиях. Речь идет о фундаментальном свойстве аффинной группы преобразований: прямые переходят в прямые независимо от конкретных параметров преобразований. Доказана возможность перехода от двумерного анализа изображения к анализу его линейных сечений, что позволяет не только упростить вычисления, но и достичь прогресса в расширении типов допустимых входных изображений. Этому в решающей степени способствуют новые разложения центроаффинной группы, приведенные выше.

Определенный интерес представляет применение инвариантных процедур по полю зрения [4]. Так применение преобразований по модулю Фурье позволяет произвести автоматическое центрирование изображений. Модуль преобразования Меллина инвариантен к изменению масштабов по осям. На базе оператора Гаммерштейна строятся инварианты к центроаффинным преобразованиям. Такие операторы позволяют применять последовательную нормализацию при сложных преобразованиях вплоть до проективных. Однако, эксперименты показывают, что их эффективное использование возможно только для хорошо сегментированных изображений. При этом объемы вычислений в реальных системах часто выходят за рамки допустимых.

Нормализация проективных преобразований представляет собой трудную и еще не решенную до конца задачу. Прогресс достигнут по отдельным подгруппам проективной группы, которая характеризуется в общем 8 независимыми параметрами. Так для нормализации ортогональной проективной подгруппы, которая характеризуется 3 параметрами, получены функционалы общего вида:

$$\Phi_1 = \iint_D \frac{x B(x, y)}{(x^2 + y^2 + 1)} dx dy, \Phi_2 = \iint_D \frac{y B(x, y)}{(x^2 + y^2 + 1)} dx dy, \Phi_3 = \iint_D \frac{B(x, y)}{(x^2 + y^2 + 1)} dx dy,$$

где $V(x,y)$ -функция яркости изображения, D -поле зрения. Для простейших проективных преобразований типа перспективы $x \rightarrow \frac{x}{1-hy}$, $y \rightarrow \frac{y}{1-hx}$, получены подходящие нелинейные функционалы типа отношения моментов. Однако, ядра моментов содержат отрицательные степени координат x , y , что ограничивает область анализа изображений полуплоскостями или даже отдельными квадрантами. Учитывая ограниченные возможности методов сечений и коэффициентов, универсальные и практически пригодные методы нормализации проективных преобразований следует искать на пути применения частичных операторов. Можно показать, что любое проективное преобразование можно разложить на аффинные и простейшие проективные, например, $x \rightarrow \frac{x}{y}$, $y \rightarrow \frac{1}{y}$ или преобразование перспективы. Это дает принципиальную возможность применить суперпозицию частичных нормализаторов. Однако, условия синтеза нормализаторов требуют поиска таких, которые обладают свойствами перестановочности и устойчивости [3]. Дополнительные вопросы вносят многочисленные условия и ограничения на коэффициенты проективных преобразований, возникающие в реальных условиях. Однако, несмотря на приведенные трудности, нет принципиальных ограничений на построение нормализаторов проективных преобразований, которые лежат в основе монокулярного восприятия окружающего мира органом зрения человека и задача эта будет в конце концов решена. Это будет этапным моментом в решении проблемы распознавания изображений.

Заключение

Процесс распознавания изображений является сложной многоэтапной процедурой. Многоэтапность (иерархичность) обусловлена тем, что различные задачи обработки на самом деле тесно связаны и качество решения одной из них влияет на выбор метода решения остальных. Так выбор метода распознавания зависит от конкретных условий предъявления входных изображений, в том числе характера фона, других изображений, помеховой обстановки и связан с выбором методов предобработки, сегментации, фильтрации.

Однако, для типичных ситуаций, на основе проведенного в статье анализа, можно предложить универсальную иерархическую структуру распознавания. На первом этапе применяются наименее трудоемкие признаковые алгоритмы для решения задачи о неэквивалентности входных изображений и эталонов. Входное множество изображений при этом существенно сокращается. На втором этапе оставшиеся изображения подвергаются нормализации. На третьем - нормализованные изображения классифицируются одним из конструктивных способов, например корреляционным. При такой структуре распознавания время решения задач сокращается в сотни раз.

Литература:

1. Handbook of pattern recognition and computer vision / Chen C.H., Rau L.F. and Wang P.S.P.(eds.). – Singapore-New Jersey-London-Hong Kong: World Scientific Publishing Co. Pte. Ltd., 1995. - 984 p.
2. Shalkoff R.J. Digital image processing and computer vision. – New York-Chichester-Brisbane-Toronto-Singapore: John Wiley & Sons, Inc., 1989. - 489 p.
3. Путятин Е.П., Аверин С.И. Обработка изображений в робототехнике. М: Машиностроение, 1990. 320 с.
4. Гиренко А.В., Ляшенко В.В., Машталир В.П., Путятин Е.П. Методы корреляционного обнаружения объектов. Харьков: АО “БизнесИнформ”, 1996. 112 с.
5. Вестник Национального Технического Университета “Харьковский политехнический институт” Выпуск 114.- Харьков: НТУ “ХПИ”, 2001. – 128с.
6. Проблемы бионики. Всеукраинский межведомственный сборник. Выпуск 50.- Харьков: “ХГТУРЭ”, 1999. – 217с.
7. Рябова Л.В. Базовые операторы для обработки изображений радужной оболочки глаза . – Наука і молодь: Зб. наук. пр. – К.: НАУ, 2006. – Вип.6.– С. 45-48.