Санкт-Петербургский политехнический университет Петра Великого

Институт машиностроения, материалов и транспорта

Высшая школа автоматизации и робототехники

ОТЧЁТ ПО БИЛЕТУ

По дисциплине: Системы технического зрения

Выполнил А.М.Орехов  
студент группы №3331506/00401

Принял В.В.Титов

«\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_2024 г

Санкт-Петербург

2024

# Билет №5. Оценка аффинных преобразований и гомографии.

Гомографией называют проективное преобразование одной плоскости в другую. В случае компьютерного зрения, за плоскости принимают изображения или плоские поверхности в трехмерном пространстве. К применениям гомографии относят:

* Ректификация (Выпрямление) изображений
* Сшивание панорам из отдельных изображений
* Деформация текстур в трехмерной графике
* Определение положения камеры
* Многовидовая геометрия

По существу, гомография H отображает точки плоскости (в однородных координатах) согласно следующему уравнению:

|  |  |  |
| --- | --- | --- |
|  | или |  |

Однородные координаты – система координат, использующаяся в теме проективной геометрии. Причиной этому является свойство, гарантирующее что точка, представленная в однородных координатах, при умножении на ненулевой масштабный коэффициент, не изменяется. Также, зачастую производится нормировка: полагают, что и с помощью этого получается однозначно представлять точки изображения только координатами x и y.

Таким образом, получается, что гомография H определена с точностью до масштабного коэффициента w и имеет восемь независимых степеней свободы.

Физический смысл параметров матрицы проективного преобразования следующий:

* Параметры a, b, d, e являются матрицей поворота, а дополнительное домножение параметров a и e, находящихся на главной диагонали матрицы дополнительно масштабирует точки.
* Параметры c и f задают параллельный перенос
* Параметры g и h задают изменение перспективы

Параметры проективного преобразования между двумя изображениями можно определить непосредственно, зная соответственные точки двух изображений. В общем случае у проективного преобразования 8 степеней свободы. Каждая найденная пара соответствующих точек дает возможность составить два уравнения. Соответственно, в минимальном варианте для оценки гомографии требуется всего 4 пары точек. Впрочем, для большей стойкости алгоритма к возможным неточностям определения точек, использование минимально возможного количества нежелательно.

Переписав матричную операцию, можно получить уравнения для компонент x и y:

Запишем уравнения в следующей форме:

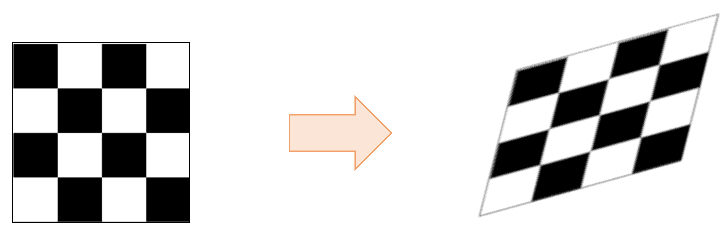
Где i – индекс пары соответствующих точек. Теперь, если переписать уравнение проективного преобразования, подставив в него полученные уравнения, можно получить матрицу вида:

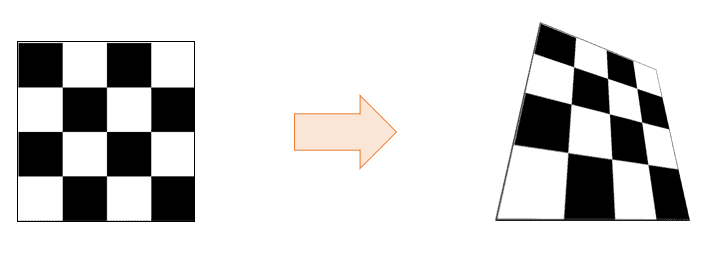
Уравнение подобного вида является основой метода решения путем прямого линейного преобразования. Далее по алгоритму из этого уравнения потребуется найти нетривиальное значение h – векторизованной матрицы H:

Кроме гомографии, также существует несколько важных частных случаев проективных преобразований. Одним из таких является ***аффинное преобразование***.

|  |  |  |
| --- | --- | --- |
|  | или |  |

Данный вид преобразований отличается от гомографии отсутствием параметров перспективных изменений, но сохраняют преобразования поворота, масштабирования и параллельного переноса. При аффинном преобразовании, на двух изображениях сохраняется параллельность прямых. В случае проективного преобразования сохраняется только прямолинейность отдельных линий (Рисунок 1).





*Рисунок 1 – Аффинное и проективное преобразование*

Для вычисления параметров аффинного преобразования можно применить метод прямого линейного преобразования, приравняв последние два элемента к нулю () Аналогично гомографии, поскольку у аффинного преобразования только шесть степеней свободы, для решения уравнения потребуется уже три пары соответствующих точек.

Таким образом, были рассмотрены основы прямого линейного преобразования для гомографии и аффинного преобразования. Была получена система уравнений в матричной форме, которая описывает соответствие между соответствующими точками на исходном и преобразованном изображении. Следующим шагом является нахождение вектора параметров ***h*** и его последующее преобразование обратно в матрицу **H**. Для идеального случая с отсутствием шумов и идеально выбранным парам соответствующих точек возможно напрямую найти значение матрицы. Однако, зачастую составленная система будет иметь количество уравнений больше, чем число неизвестных, а, следовательно, не существует такого вектора h, который бы являлся решением сразу для всех уравнений.

В реальных условиях нахождение решения только по четырем парам точек может быть усложнено из-за шумов, ошибок измерения, выбросов и ошибочно сопоставленных точек, не являющихся на самом деле соответствующими. Для решения этой проблемы требуется выбирать количество соответствующих точек большее, чем требуется для выбранной модели преобразования, улучшая результат за счет избыточности. В таком случае, для нахождения решения уравнения потребуется применение методов, предполагающих нахождение оптимального решения, пользуясь определенной метрикой, например, метод наименьших квадратов. Он, впрочем, не всегда может давать корректные результаты, в случаях, когда кроме «правильных» точек в набор были выбраны и ложные соответствия. Из-за подобной особенности как правило применяются робастные методы, позволяющие не учитывать шумы и ложные соответствия, например, RANSAC.

# Билет №33. Эпохи решения задачи классификации.

Задача классификации является одной из ключевых проблем в области машинного зрения, нацеленной на разделение объектов, сцен, или событий на определенные категории на основе формализованных визуальных признаков. На протяжении нескольких десятилетий эта область исследований пережила несколько эпох, каждая из которых отмечена появлением новых методов, технологий и подходов, улучшавших точность и эффективность классификации.

В начале 1960-х компьютерное зрение зарождалось в университетах, изучающих вопросы искусственного интеллекта и одними из первых решаемых в этой сфере задач, была задача классификации геометрических объектов. Одним из препятствий для создания алгоритма классификации объектов является непосредственно разнообразие различных классов объектов. Объекты могут иметь различные формы, размеры, текстуры и цвета, что делает сложной задачу их точного выделения и классификации. На заре сферы компьютерного зрения, данную проблема решалась уменьшением разнообразия классов для распознавания и упрощением их признаков. Так, в 1963 году вышла диссертационная работа, посвященная классификации трехмерных объектов на изображении.

Изображение выглядит как текст, диаграмма, зарисовка, дизайн

Автоматически созданное описание

Примененный метод на полученном изображении выделял острые углы, находя их с помощью дифференциального оператора, после чего применялся многоэтапный процесс по выделению линий. Первым шагом по изображению проходило окно 4х4, находившее точки, потенциально принадлежащие линиям. После чего над изображением проводилось 4 операции корреляции с целью найти линии со строго определенными направлениями – 0, 1, и -1. Далее, если прямые соприкасались друг с другом в окне пикселей 4х4, они соединялись и на место соединения добавлялась точка. Все линии без соединений удалялись. На собранных формах выполнялись операции по поиску фигур, сначала собирая линии в двумерные примитивы, а после сопоставляя их форму с формами трехмерных примитивов.

Позже, в 1966 группа ученых для работы над задачей инициировали «летнюю школу зрения» как эксперимент по созданию программы, которая позволит компьютеру «объяснять то, что он видит». Первичной целью, сформулированной авторами, было разделить область кадра на объекты, фон и «хаос». В качестве плана-максимум планировалось решить задачу идентификации объектов в кадре по предварительно известной библиотеке объектов, в качестве которых были выбраны геометрические примитивы – шары, кубики с одинаковыми и разными цветами граней и цилиндры.

• Early 1990s: invariants, appearance-based methods

• Mid-late 1990s: sliding window approaches

• Late 1990s: feature-based methods

• Early 2000s: parts-and-shape models

• 2000 – 2010: bags of features

• 2010-2015: combination of local and global methods, modeling context, integrating recognition and segmentation

• 2015 – present: machine learning, artificial neural network